

Научная статья

УДК 8'33

DOI: <https://doi.org/10.18721/JHSS.14105>



МАШИННЫЙ ПЕРЕВОД В ЭПОХУ ЦИФРОВИЗАЦИИ: НОВЫЕ ПРАКТИКИ, ПРОЦЕДУРЫ И РЕСУРСЫ

О.Н. Камшилова  , Л.Н. Беляева 

Российский государственный педагогический университет им. А.И. Герцена,
Санкт-Петербург, Российская Федерация

 onkamshilova@gmail.com

Аннотация. В российской лингвистике к условиям цифровизации традиционно относят применение математических и компьютерных методов для решения различных задач, в основном связанных с проблемами обработки текста в различных автоматизированных системах. В статье анализируется влияние процесса цифровизации на назначение и использование систем машинного перевода (МП) в современных условиях, описываются новые практики использования продуктов МП как в профессиональной переводческой деятельности, так и в рамках решения частных задач пользователей таких систем. Отмечаются объективные преимущества и недостатки МП с точки зрения практикующих переводчиков-профессионалов и простых пользователей. Рассматриваются новые условия работы переводчиков, их новые роли и навыки, определяемые влиянием цифровизации на работу с текстом. Специальное внимание уделяется постредактированию продуктов МП как новой области профессиональной деятельности переводчика, необходимой для обеспечения качественного перевода, извлечению корректной информации. Определяется объем необходимого и достаточного постредактирования при решении частных задач пользователями-непрофессионалами. Анализируются доступные процедуры и лингвистические ресурсы, способные оптимизировать работу с системами МП.

Ключевые слова: цифровизация, машинный перевод (МП), системы МП, постредактирование, переводческие практики, лингвистические ресурсы.

Для цитирования: Камшилова О.Н., Беляева Л.Н. Машинный перевод в эпоху цифровизации: новые практики, процедуры и ресурсы // Terra Linguistica. 2023. Т. 14. № 1. С. 41–56. DOI: 10.18721/JHSS.14105



MACHINE TRANSLATION IN THE AGE OF DIGITALIZATION: NEW PRACTICES, PROCEDURES AND RESOURCES

O.N. Kamshilova  , L.N. Beliaeva 

Herzen State Pedagogical University of Russia,
St. Petersburg, Russian Federation

 onkamshilova@gmail.com

Abstract. In Russian linguistics digitalization is traditionally associated with the use of mathematical and computer methods applied mainly to text processing problems in various automated systems. The article analyzes the impact of digitalization on the use and purpose of machine translation (MT) systems in modern conditions. It describes new practices of using MT products both by professional translators and by a general MT system user for their individual purpose. It highlights the objective advantages and disadvantages of MT application from the point of view of practicing professional translators and ordinary users as well. The article considers a translator's new working conditions, their new roles and skills determined by the impact of digitalization on working with text. It pays special attention to post-editing MT products as a translator's new professional activity, which is needed to ensure high-quality translation and to extract correct information. It also describes the necessary and sufficient post-editing procedures to be performed by non-professional users while pursuing their own goals through MT application. Finally, the research focuses on the analysis of procedures and available linguistic resources that can optimize working with MT systems.

Keywords: digitalization, machine translation (MT), MT systems, post-editing, translation practices, linguistic resources.

Citation: O.N. Kamshilova, L.N. Beliaeva, Machine translation in the age of digitalization: new practices, procedures and resources, *Terra Linguistica*, 14 (1) (2023) 41–56. DOI: 10.18721/JHSS.14105

Введение

В современном цифроориентированном мире, благодаря изменению геополитической и социокультурной ситуаций, возникновению все новых возможностей представления информации в цифровой среде, разработке телекоммуникационных систем и технологий, новых средств связи, систем автоматического анализа и обработки информации, кардинально меняется как само представление о тексте, так и представление о способах и возможностях работы с текстами на естественных и искусственных языках.

Процесс цифровизации (что бы конкретно не понималось под этим процессом, ср., например, [1]) и особенности его реализации определяют необходимость учета целого комплекса составляющих: технических, связанных с выбором соответствующих технологий, а также собственно лингвистических, определяющих особенности лингвистического обеспечения всего применяемого комплекса методов и технологий.

В российской лингвистике к условиям цифровизации традиционно относят применение математических и компьютерных методов для решения различных задач, в основном связанных с проблемами обработки текста в различных автоматизированных системах или, вообще, к приложению лингвистических знаний в других сферах деятельности. В связи с задачами настоящей статьи выделим особо сферу машинного перевода (МП) и, в частности, работу с продуктами МП.

Технические составляющие процесса цифровизации ограничиваются и определяются учетом оснащенности различных учреждений компьютерами, способными поддерживать процесс диа-



лога и дистанционного обмена данными. Собственно *лингвистические составляющие* включают использование специальных предметно-ориентированных словарных систем и баз данных, локализацию программного обеспечения специализированных систем поддержки обучения, формирование специализированных корпусов текстов, применение систем МП в процессе локализации и лексикографирования.

Широкое развитие систем МП, их полная доступность делают бессмысленной борьбу с их использованием как в работе переводчиков и специалистов в различных областях знаний, так и при обучении иностранным языкам и при подготовке переводчиков. Более целесообразным и актуальным представляется включение работы с системами МП и постредактирования продуктов МП в требования к профессиональной подготовке (ср. [2–3]). Задачей настоящего исследования является анализ лингвистических процедур и доступных ресурсов работы с системами МП, способных обеспечить техническую и лингвистическую составляющие функционала специалистов по работе с текстом.

Методология

В основе исследования лежит критический анализ назначения и использования систем МП в современных условиях, обзор специальных условий, необходимых как для работы с системами МП, так и для работы по редактированию (постредактированию) результатов МП, сравнительный анализ ресурсов и процедур работы с системами МП, обеспечивающих успешное использование МП для решения профессиональных и пользовательских задач разного рода (исследовательских, образовательных, информационных).

Результаты исследования

Назначение и использование систем МП в современных практиках перевода

Основным назначением систем машинного перевода (МП) является оперативный перевод специальных (научных, научно-технических, информационных) текстов. Качество результатов работы такой системы определяется требованиями его потребителей: специалистов в конкретной области знаний и переводчиков. Целесообразность использования систем МП установлена давно и не требует новых доказательств. Применение систем МП в рамках научного и технического перевода обеспечивает, во-первых, выигрыш в продуктивности за счет значительного сокращения временных затрат; во-вторых, выигрыш в качестве за счет согласованности и корректности перевода терминологии при использовании предметно-ориентированных систем; в-третьих, выигрыш в удобстве организации работ за счет возможности контроля над переводом потоков текстов и возможности разделения работы между несколькими переводчиками.

Очевидные успехи в области разработки систем МП привели к тому, что сегодня можно наблюдать серьезные изменения в их назначении и использовании: если *исторически* машинный перевод был ориентирован на специальные тексты, а пользователями этих систем были переводчики или специалисты отдельных предметных областей, то сегодня в практике МП складывается ряд *новых направлений*. В частности, перевод литературных произведений и субтитров, что не удивительно, поскольку применение МП в этой сфере дает 36% увеличения производительности труда переводчиков в соотношении «количество переведенных слов на единицу времени» [4–5]. Пользователем систем МП в данном случае по-прежнему остается переводчик, однако приходится констатировать низкое качество перевода, отсутствие редактирования результатов МП, что можно объяснить большим спросом на массовую продукцию (массовая литература, кино и сериалы на многочисленных частных каналах и т.п.), в результате потребитель получает некачественные переводы, демонстрирующие пренебрежение элементарными правилами перевода вроде всем известных «ложных друзей» переводчика:



Алые губы подчеркивали пронзительную сексуальность и безупречную красоту актрисы, а в сочетании с роскошными драгоценностями, меховыми манто и открывающими плечи сатиновыми платьями заявляли всему миру о ее славе, успехе и богатстве.¹

Еще одно направление и канал использования и распространения некачественного перевода – новостные агрегаторы, сетевые издания и каналы на платформах Telegram, YouTube и т.п, типа сайта AKKet.com:

*Многие россияне воспользовались возможностью и купили товары от компании IKEA, получив **такие** в пункте выдачи заказов или с доставкой на дом. Получать **такие** в свои руки покупатели будут до самого конца сентября <...> В рамках последней распродажи многие россияне купили себе **различные самые разные** товары, которые были доступны для покупки **до середины** августа <...> При этом магазины в какой-то момент также заработают, потому что иначе смысла сохранять **такие** и поддерживать их в **актуальном** состоянии **бы попросту не было.**²*

Об использовании МП в данном тексте свидетельствуют не только ошибки в предложном управлении, лексической сочетаемости, порядке слов, но и анафорическое употребление местоимения *такие* (ср. *such, these*), в совершенно не характерной для русского языка конструкции и, напротив, естественной для английского. Подобные неотредактированные тексты можно встретить на многих официальных каналах иностранных новостных агентств, публикующих материалы на русском языке. Все это свидетельствует о том, что редактирование переводных материалов приносится в жертву скорости публикации.

При этом нельзя не отметить, что качество результатов систем нейронного машинного перевода в последнее время повысилось кардинально, выявляя в рамках переводческих техник новое противопоставление: *machine translation* ↔ *translation from scratch*, т.е. противопоставление полного или частичного машинного перевода переводу ручному от начала до конца. При этом последний становится «штучным товаром», предполагая высочайшую квалификацию переводчика, а первый одинаково популярен как у переводческих компаний, выполняющих огромный объем работы, так и у широкого круга лиц – от переводчиков-фрилансеров до специалистов-исследователей, студентов, решающих академические задачи, и просто пользователей сети, предлагающей доступные инструменты типа Гугл-переводчика. Обращение широкого круга пользователей к применению результатов МП определяет третье направление в изменении характера практик МП.

Анализ материалов научных журналов, конференций, студенческих исследовательских работ позволяет заключить следующее: системы МП широко используются специалистами в различных областях знаний для анализа и перевода статей, извлекаемых из множества источников, в первую очередь из материалов конференций, представленных в системе Интернет, а также для перевода заглавий, аннотаций, ключевых слов и собственно текстов, написанных на родном языке автора, для публикации в отечественных и зарубежных журналах:

In article the heuristic potential categories “rationality” is considered, the basic approaches to its studying are shown. The author allocates initial characteristics and genetic roots of rationality, analyzes the reasons of a crisis state of idea “ratio” in modern humanitarian knowledge. The further prospects of theoretical judgement of a phenomenon of “rationality” are discussed.³

При этом «новые пользователи» демонстрируют следующие типичные недостатки в работе с результатами МП:

- не знают процедур работы с такими системами;
- не осознают необходимости постредактирования и возможности минимизации объема постредактирования результатов МП;

¹ Rachel Felder. RED LIPSTICK – An Ode to a Beauty Icon, перевод «Секретное оружие : история красной помады / Рейчел Фелдер ; [перевод с английского А. А. Джапаридзе]: Эксмо; Москва; 2021, Научный редактор Анна Жуковская

² <https://akket.com> – публикация от 08.09.22

³ Аннотация к научной статье, поданной в рецензируемый журнал



- не понимают связи объемов и методов постредактирования с полнотой и точностью автоматического словаря соответствующей системы.

При всех определенных успехах современных систем МП отношение к ним весьма неоднозначное как у переводчиков-профессионалов, так и в академической среде.

Более 38% практикующих переводчиков продолжают опасаться использования МП как возможного конкурента, который сократит объем их работы или вообще заменит их [6]. Кроме того, объем ежедневного постредактирования, выполняемого переводчиками, в среднем составляет приблизительно 5 000 слов, этот процесс вызывает неприятие переводчиков и отрицательное отношение к результатам МП в целом. Новые характеристики самого процесса перевода и требований к его качеству и скорости получения результатов определяются еще и тем, что он включен в так называемую индустрию локализации (*Localization Industry*) [7–8], которая охватывает не только собственно перевод, но и адаптацию его результата к культуре принимающего языка, решение маркетинговых и технологических задач. При этом следует учитывать, что результаты обработки текстов на разных языках являются базой решения различных научных и практических задач, поэтому использование систем МП является важной частью перевода как технологического процесса.

В академической сфере борьба с использованием МП при подготовке студентами материалов на иностранном языке, в т.ч. на русском, если речь идет об обучении иностранных студентов и аспирантов, носит принципиальный характер. Преподаватели иностранных языков рассматривают использование таких систем студентами как совершенно недопустимое и неэтичное. Почти половина тех, кто изучает или преподаёт романские языки (испанский, французский, итальянский и португальский), считает применение МП непродуктивным, более того, рассматривают это как мошенничество [9]. При этом нельзя отрицать, что практичность, простота использования и бесплатный доступ к сетевым системам МП сделали эти инструменты очень популярными среди изучающих язык. Предваривший наше исследование опрос студентов и аспирантов филологического факультета РГПУ им. А.И. Герцена показал, что большинство использующих МП для перевода на английский язык предпочитают систему Google Translate и оценивают результат работы с этим инструментом достаточно высоко. Поэтому сегодня в академической среде наряду с этическим запретом к использованию МП наблюдается новая тенденция – поиск эффективных способов использования таких систем в обучении языкам [10].

Наш собственный опыт использования систем МП в обучении позволяет утверждать, что оценка эффективности и целесообразности использования МП зависят от типа изучаемого языка, а приемлемость и процедуры использования систем МП преподаватель должен оценивать сам, исходя из возможностей образовательной среды и уровня подготовки студентов.

Постредактирование как новый вид деятельности переводчика

Переводчик должен обладать знаниями в сфере владения терминологией предметной области, входного языка и языка перевода, межкультурных отношений, поиска информации, технологических инструментов и программ, а также в сфере поставки услуг переводчиков и др. Постредактирование как особый вид деятельности переводчика предполагает умения выполнять корректуру и редактировать перевод, а также умение устанавливать и контролировать стандарты качества. Однако, общая структура цифровизации сегодня предполагает изменение не только деятельности постредактора, но и процесса редактирования в целом.

Традиционно редактор рассматривается как специалист, основной задачей которого является *редактирование*, т.е. установление меры ценности готовящегося к изданию произведения и приведение его к стандартной форме, соответствующей стилистическим требованиям (ср. [11]). Последовательность работы редактора в той форме, в которой она определяется полученными в вузовском обучении знаниями, можно описать следующим образом: определение способа изложения (функционально-смыслового типа речи: описания, повествования, рассуждения) и его разновидности по конкретным признакам с опорой на их знание; установка



соответствия требованиям к способу построения, выбору языковых средств, поиск возможных недостатков и типичных ошибок [12]. Под литературным редактированием понимается «совершенствование редактором формы литературного произведения (композиции, языка, стилистических качеств)» [13, с. 9].

Активно развивающийся практически во всех сферах деятельности процесс цифровизации изменил традиционный формат работы с текстом, делая обязательным освоение новых платформ, работу с интернет-версиями изданий и принятие форматов мультимедийной редакции. В современных условиях работа редактора предполагает учет этих новых возможностей, платформ и технологий, поскольку аудитория перешла в киберпространство, т.е. можно утверждать, что происходят изменения профессиональных компетенций редактора под воздействием как внешних, так и внутренних факторов [14]. Новый подход к созданию и ведению сайтов, к созданию журналистских текстов потребовал от редактора освоения новых ролей и навыков, не входивших ранее в сферу его базовых компетенций. У общей профессии редактора появились новые варианты: администратор сайта, контент-менеджер, аккаунт-менеджер, seo- и smm-специалисты (*social media marketing – smm*), редактор пользовательского контента, мультимедийный редактор, редактор-модератор (фасилитатор), редактор-рерайтер и множество других. Редактор-модератор организует коммуникацию с аудиторией, целью его работы является повышение уровня доверия к определённому СМИ. В функции редактора-рерайтера входит поддержание сайта конкретного СМИ в актуальном состоянии, постоянное обновление контента, оптимальное мультимедийное наполнение, а также интерактивное взаимодействие с аудиторией [15], см. также Программу переподготовки редакторов на факультете журналистики МГУ⁴. Следовательно, подготовка переводчиков, равно как и лингвистов (в широком смысле, включая подготовку многих профессий, предполагающих работу с текстом [16]) должна кардинально измениться.

Независимо от полноты автоматического словаря системы и его привязки к предметной области, результат МП требует постредактирования на уровне синтаксической структуры всего предложения, на лексическом уровне для уточнения и/или изменения переводов отдельных слов и словосочетаний, введения переводов для незарегистрированных в словаре единиц, а также для изменения морфологических характеристик рода, числа, падежа, уточнения форм времени и залога, введения корректной пунктуации. При оценке трудоемкости этого процесса внесение стилистических изменений обычно не рассматривается. Проведенные исследования [17–18] показали, что отказ от включения МП в процесс перевода с последующим трудоемким редактированием больше свойственен профессиональным переводчикам, чем тем, кто еще только получает эту профессию. Возможно, это связано еще и с уровнем компьютерной грамотности опрошенных, а также с небольшим опытом перевода. Многолетний опыт собственной работы авторов показывает, что работа с постредактированием результатов МП оставляет простор для решения творческих и лингвистических задач. Обучение постредактированию результатов МП должно составлять обязательную часть подготовки современных специалистов и не только переводчиков.

Постредактирование результатов МП в решении частных задач

Использование систем МП в исследовательской работе требует особого подхода к получаемому результату и к объему необходимого и достаточного постредактирования, при этом особое внимание должно уделяться переводу слов, которые не зарегистрированы в автоматическом словаре используемой системы, являются неологизмами и могут включаться в словосочетания, чаще всего терминологические, называющие новые реалии – именованные сущности. Постредактирование на лексическом уровне требует уточнения и изменения переводов конкретных лексических единиц, на синтаксическом – преобразования структуры предложения. Например, в случаях перевода с английского языка на русский необходимы проверка согласования по роду, числу

⁴ Программа профессиональной переподготовки «Редактор текстов для СМИ». Факультет журналистики МГУ им. М.В. Ломоносова (сайт). [Электронный ресурс]. URL: <http://www.journ.msu.ru/education/extra/editor/>



и падежу, уточнение места подлежащего, иногда полная перестройка предложения или переход к непрямой структуре типа *we have ^мы имеем *у нас есть*.

Если учесть, что при переводе научного или технического текста на выбор переводного эквивалента конкретного термина профессионалом-переводчиком затрачивается до 75% времени, необходимого для перевода текста в целом [19–21], то трудоемкий процесс обращения с иноязычной терминологией, формирования и описания переводных эквивалентов новых, вводимых в научный обиход терминов, оказывается исключен из процесса обучения студентов и аспирантов академическому письму, что при достаточно высоком уровне владения иностранным языком, позволяющем специалисту в определенной предметной области читать и понимать иноязычный научный текст даже при наличии беспереводных терминологических единиц, приводит к созданию научных текстов на безумной смеси языков, новом научном воляпюке. Вряд ли пример приведенного ниже фрагмента научного текста можно рассматривать как корректный текст на русском языке:

Это справедливо, например, для энвайронментализма, который, в свою очередь, был тесно связан с другими арт-практиками, в частности, ленд-артом, экологическим искусством, арте повера, био-артом, art&science.

Постредактирование результатов МП и получение окончательного варианта перевода текста требуют обращения к словарным и энциклопедическим базам данных. Сложности работы на этом уровне определяются тем, что использование систем машинного перевода специалистами в различных областях знаний, особенно использование предметно-ориентированных практических систем, действительно дает возможность понимания общего содержания текста, но не вводит непереверденные системой лексические единицы в систему лексики языка перевода, русского языка в частности, что определяет объем заимствований при создании текстов на русском языке в соответствующей предметной области. Тем самым создается база для неоправданного введения в русский язык массы новых слов, в частности англицизмов. Поскольку эти слова встречаются недостаточно устойчиво, их значение определено нечетко, а сами слова не введены в терминологические базы данных, их можно рассматривать как молодые, не полностью освоенные лексикой русского языка – лексические недопески – то ли терминоиды, то ли сленгизмы [22], то ли единицы общеупотребительной лексики.

Совершенно очевидно, что перевод стандартизированной терминологии и поиск эквивалентов для новых терминологических единиц остаются одной из самых трудоемких задач, решаемых и в процессе перевода, и в постредактировании результатов МП.

Ресурсы и процедуры для оптимизации работы с системами МП

В общепринятом сценарии постредактирования результатов машинного перевода человек-постредактор исправляет перевод, созданный системой МП. Так, например, множество систем переводческой памяти (Trados, MemoQ, OmegaT, и др.) обеспечивают доступ к системам МП, результаты работы которых могут быть объединены в совокупность переводов систем переводческой памяти. Следует иметь в виду, что в больших переводческих проектах часто возникает необходимость разделения текста большого объема между несколькими переводчиками. В этих условиях именно система МП, настроенная на конкретную предметную область, позволяет унифицировать перевод терминологии.

Недавно появилась возможность активного машинного обучения и взаимодействия [23], при котором система МП переупорядочивает сегменты исходного языка, которые должны переводиться и подвергаться постредактированию, чтобы максимизировать продуктивность и эффект от накопленного опыта. Вместо представления текста в первоначальной последовательности, система сортирует сегменты согласно степени доверия к полученным ранее результатам МП, так что она может научиться более корректно решать задачу перевода на основе исправлений, сделанных человеком. Это сопровождается новыми концептуализациями последовательности выполнения



работ по переводу и постредактированию, которые связывают эти новые возможности с инновационным использованием краудсорсинга.

Чтобы полностью использовать потенциал краудсорсинга, необходимо найти новые способы для того, чтобы делить связные тексты на фрагменты, которые независимо друг от друга несколько переводчиков могут переводить и редактировать одновременно. Под действием увеличившегося спроса на производительность перевода и на то, чтобы цикл перевода стал короче, известные инструментальные средства перевода (Wordbee, Trados, MATECAT) предлагают объединенные функциональные возможности.

Некоторые компании (Unbabel1, MotaWord2), которые используют технологии LSP, позволяющие подсчитывать и ограничивать трафик, регулировать скорость загрузки и приоритеты, фильтровать и распределять контент, ищут возможности экспериментировать с более динамичным подходом к совместному переводу, при котором документ разбивается на меньшие единицы. Компания MotaWord, например, объявила, что разработала «самую быструю в мире платформу для ручного перевода», основанную на совместной облачной платформе, «эффективно координируемой через интеллектуальный внутренний интерфейс», в котором участвует более 9 000 переводчиков [24, с.11]. Это возможно только в том случае, если большие документы раздроблены на маленькие сегменты, чтобы постредактировать ограниченное количество меньших фрагментов текста с введением в действие «толпы» (crowd) переводчиков.

Однако, неясно, как переводчики справляются с ситуацией, в которой меньшие сегменты, возможно выбранные из разных частей одного и того же документа, представлены в отрыве от контекста. Влияние того, что переводчики переводят сегменты в порядке, не соответствующем исходному, на процесс перевода ранее не исследовалось.

Для совершенствования работы постредактора и оптимизации ее результата современные исследования процесса постредактирования предлагают методы слежения за положением глаз или зрачков постредактора и регистрацию сигналов от нажатий человеком клавиш на клавиатуре компьютера, а также метрики, принятые в области исследования процесса перевода.

Практическая работа переводчика с системой машинного перевода предусматривает:

- ручное предредактирование текста – подготовку исходного текста (массива текстов) к переводу;
- ручное постредактирование переводов – редактирование результатов работы системы МП;
- ведение собственного (пользовательского) словаря, фиксирующего результаты МП и определяющего настройку системы МП на задачи конкретного переводчика.

Предредактирование текста необходимо для установления единства используемой терминологии, например, в системах извлечения данных (*data mining systems*), при работе с которыми некорректные результаты часто возникают из-за расхождения между данными, извлекаемыми из текста, и номинацией соответствующих объектов в словарном обеспечении (базах данных или онтологиях). В задачи предредактирования должно включаться исправление ошибок в текстах на глобальном английском языке, в частности, в излишней пассивизации текста, и упрощение текста в связи с решением задач перевода и инженерии знаний.

Постредактирование результатов МП и получение окончательного варианта перевода текста требует обращения к словарным и энциклопедическим базам данных, корпусам текстов, заранее выбранным переводчиком. При решении вопроса о выборе перевода конкретной терминологической единицы необходимо привлечение миниконкорданса. В результате такой работы на этапе собственно перевода должен формироваться пользовательский словарь, характеризующий терминологические особенности конкретного текста. Этот словарь на этапе поддержки системы добавляется в ее лингвистические ресурсы.

В процессе перевода в режиме реального времени текст может предварительно обрабатываться, при этом:



- получение частотного словаря и миниконкорданса по конкретному тексту помогает выявить основную терминологию и установить ее контекст,
- предварительное редактирование текста позволяет снять его стилистические несообразности, устранить сверхдлинные предложения и т.д.
- использование системы машинного перевода, выбранной и настроенной на необходимую предметную область, дает вариант перевода, который требует анализа и постредактирования.

Собственно процесс редактирования перевода может в этом случае осуществляться либо в стандартном двухоконном интерфейсе Windows, когда исходный текст и перевод представлены в двух параллельных окнах, что позволяет редактировать и заново переводить оба текста, либо в режиме работы только с текстом перевода. Опыт практической работы с системами МП позволяет рекомендовать именно первый вариант как наиболее удобный.

Таким образом, после завершения перевода конкретного текста должна происходить перенастройка лингвистических ресурсов: пополняется корпус параллельных текстов за счет исходного текста и его перевода, формируется и/или пополняется пользовательский словарь, включающий терминологию, выявленную и проверенную переводчиком, пополняется база словарей.

Примером специально разработанной для решения исследовательских задач в области характеристик результатов систем МП является электронная платформа E-NBU, созданная в Лаборатории языковых технологий Нового болгарского университета (New Bulgarian University). Эта платформа изначально разрабатывалась как инструмент для преподавания языка и представляла собой генератор онлайн-упражнений, извлекаемых из аннотированных корпусов с экспортом в систему Moodle или другие образовательные платформы. Платформа E-NBU в настоящее время расширена дополнительными модулями и функциональными возможностями, позволяющими проводить исследования в области перевода и анализа ошибок, а также поддерживать лексикографические проекты.

Другой тип электронной платформы – это достаточно давно разрабатываемая в РГПУ им. А.И. Герцена модель автоматизированного рабочего места (АРМ) преподавателя и переводчика [25–26]. АРМ преподавателя, включает

- 1) среду для создания, организации и ведения электронных текстовых архивов и извлечения корпусов текстов;
- 2) модули для проведения лингвистического анализа: лемматизатор, анализатор частей речи, приписывающий словоформам или словосочетаниям соответствующие теги; анализатор терминов; морфологический анализатор, синтаксический анализатор; анализатор многокомпонентных лексических единиц (включая сложные термины, аналитические формы, фразеологические единицы); средство выравнивания параллельных текстов; конкордансер;
- 3) лингвистическую базу данных, позволяющую работать с корпусом без потери информации;
- 4) модули, предназначенные для создания и редактирования онлайн-упражнений.

Специальные средства для ведения электронного текстового архива позволяют формировать различные метаданные, которые могут, по отдельности или в комбинациях, составлять основу для построения корпусов текстов. После проведения лингвистического анализа, необходимого для создания корпуса текстов, из него могут извлекаться вторичные («виртуальные») корпуса: списки предложений, содержащих определенную единицу – лемму, словоформу, словосочетание, тег или комбинацию тегов. Применяемая архитектура (см. рис. 1) позволяет параллельно использовать несколько систем предварительной обработки и сравнивать их результаты, что превращает такое АРМ в удобную среду для проведения экспериментов и исследований [27], которая может подключаться к системе Moodle.

Большое число лингвистических ресурсов сегодня представляет собой базы и банки данных, а также корпусы текстов, предназначенные в том числе для обучения языкам и подготовки переводчиков. К числу таких ресурсов относятся:

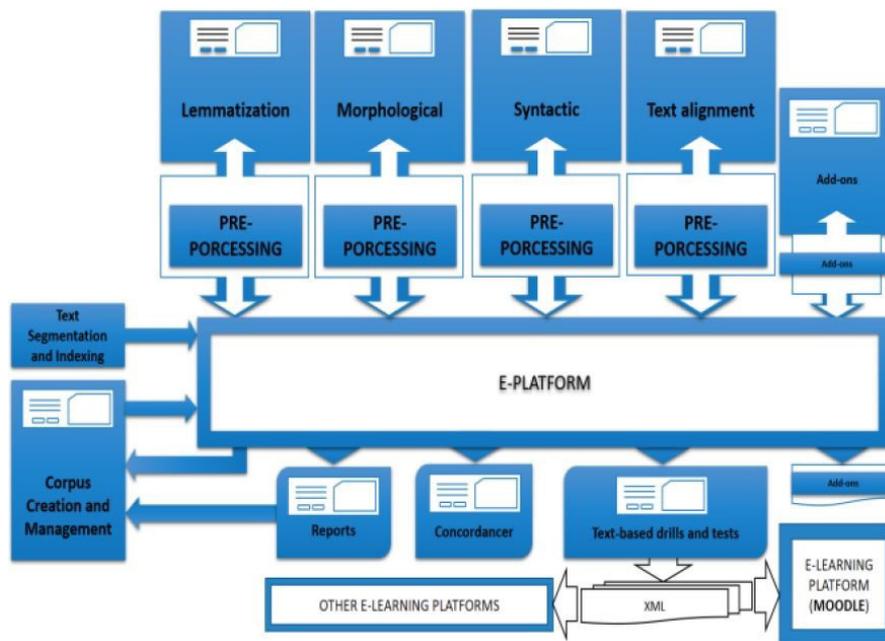


Рис. 1. Архитектура платформы E-Platform [27, с.100]

Fig. 1. Architecture of the E-Platform [27, p. 100]

- Ресурс *TAUS Data Cloud* (<https://www.taus.net/data/taus-data-cloud>) – крупнейшее хранилище переводческой памяти, объем которого превышает 79 миллиардов слов в более чем 2200 языковых пар. Этот ресурс в основном используется для обучения системам МТ как в промышленности, так и в научных и учебных кругах. Пользователи могут извлекать данные, если у них достаточно кредитов, которые можно либо купить, либо заработать, загрузив собственные переводческие материалы. Данные доступны для поиска на основе относительно небольшого набора меток: язык источника и язык перевода, предметная область, тип контента, владелец данных и поставщик данных.

- Ресурс *My Memory* (<http://mymemory.translated.net>), управляемый компанией LSP Translated, включает в себя программное обеспечение для управления терминологическими данными Европейского союза и Организации Объединенных Наций, а также данные, полученные с многоязычных веб-сайтов. Его основными пользователями являются те, чья работа связана с лингвистическими технологиями, те, кто занимается обработкой текстов на естественных языках, и переводчики. Загрузка массивов переводческой памяти является бесплатной. Поиск на основе метаданных ограничен языковыми парами и предметными областями.

- Ресурс *European Parliament Proceedings Parallel Corpus* (Параллельный корпус слушаний в Европейском парламенте [28]) характеризуется большим размером, многоязычностью и выравниванием по предложениям. Однако в него включен ограниченный круг предметных областей и типов текста. Метаданные очень скудны (языки и направления перевода), что ограничивает его полезность для учебных целей.

- Европейская комиссия предоставляет свободный доступ к ряду своих крупных многоязычных систем переводческой памяти и корпусов из областей политики, права и экономики (<https://ec.europa.eu/jrc/en/language-technologies>). Эти ресурсы часто используются для обучения систем МП и для поддержки систем переводческой памяти в конкретных предметных областях. Метаданные также скудны.

- Ресурс *OPUS* [29] представляет собой бесплатную, постоянно растущую коллекцию уже существующих, автоматически обогащаемых (например, тегами частей речи) корпусов и систем



переводческой памяти, извлеченных из Интернета. Его основными преимуществами являются большой размер, множество языковых пар, текстовое разнообразие с точки зрения предметных областей и типов текстов, разнообразие способов перевода (письменный перевод, локализация и субтитры), а также его открытый характер. С другой стороны, метаданных в нем мало: они предоставляют различные, но небольшие наборы меток в зависимости от соответствующего корпуса, доступ к которому осуществляется через интерфейс OPUS, например, корпус EuroParl. Благодаря своему размеру, вариативности и бесплатной доступности он может быть полезным для учебных целей.

- Ресурс *Translational English Corpus* (<http://www.alc.manchester.ac.uk/translation-and-intercultural-studies/research/projects/translational-englishcorpus-tec/>) включает в себя несколько субкорпусов, состоящих из различных видов текстов, и сообщает множество метаданных об экстралингвистических параметрах текстов, включая данные о переводчиках, эти тексты создавших. В отличие от других ресурсов, этот корпус является не параллельным, а одноязычным сопоставимым корпусом, противопоставляющим оригинальный текст на английском языке английскому языку перевода. Использование его данных оказало влияние на теорию перевода, особенно в том, что касается исследований переводческих и стилистических вариаций. Несомненна его значимость и для развития навыка постредактирования.

- Ресурс *MeLLANGE Learner Translator Corpus* (<http://www.alc.manchester.ac.uk/translation-and-intercultural-studies/research/projects/translational-englishcorpus-tec/>) является многоязычным учебным корпусом и включает не профессиональные переводы, а переводы, выполненные переводчиками-стажерами. Несмотря на сравнительно небольшой размер, он предоставляет множество специфичных для перевода метаданных на различных уровнях, включая информацию о профилях переводчиков и процессах перевода (например, затраченное время и тип используемых средств перевода). Он лучше всего подходит для изучения дидактических аспектов перевода, качества перевода и предоставляет обширный материал для работы с постредактированием.

- Ресурс *Human Language Project* [30] представляет собой попытку создания универсального корпуса для всех языков мира для изучения языковых универсалий и документации на естественных языках.

К сожалению, не существует универсального хранилища параллельных переводческих данных, подходящего для максимально широкого спектра задач, связанных с обучением языкам, переводу и постредактированию.

Заключение

МП сегодня становится регулярной и неотъемлемой частью переводческих практик. Кроме того, доступные в сети системы МП активно используются для решения частных задач – информационных, исследовательских, образовательных.

В самой практике МП намечаются новые направления его применения: от перевода сугубо специальных текстов научно-технического характера к переводу текстов художественных, рекламных, новостных.

Новые условия работы переводчика и лингвиста, а также требования к их профессиональным компетенциям, продиктованные влиянием цифровизации на работу с текстом, делают работу с системами МП неизбежной. Не пользовательские, но профессиональные навыки обращения с продуктами МП, их постредактированием должны стать основой работы специалиста.

Ресурсы и процедуры работы с системами МП, проанализированные в этом исследовании, способны обеспечить техническую и лингвистическую составляющие процесса использования систем МП, прежде всего для переводчиков и редакторов, а также исследователей, работающих с иноязычными источниками и участвующих в международном научном обмене.



Сегодня спектр систем МП постоянно расширяется и совершенствуется, и с точки зрения доступных пар языков, и по качеству результатов. Поэтому позиция неприятия результатов их работы и неучета их использования непродуктивна. Простое запрещение применения МП неэффективно, поскольку пользователи продолжают обращаться к системам МП независимо от вводимых запретов.

Материалы статьи были представлены на IV Международной конференции по инженерной и прикладной лингвистике «Пиотровские Чтения – 2022», посвященной 100-летию со дня рождения профессора Р.Г. Пиотровского в РГПУ им. А.И. Герцена 22 ноября 2022 г.

СПИСОК ИСТОЧНИКОВ

1. **Стариченко Б.Е.** Цифровизация образования: иллюзии и ожидания // Педагогическое образование в России. 2020. № 3. С. 49–586. URL: <https://cyberleninka.ru/article/n/tsifrovizatsiya-obrazovaniya-illyuzii-i-ozhidaniya?ysclid=ldvzz4y1cs241893073>
2. **Абросимова Н.А., Щелокова Е.А.** Пред- и постредактирование машинного перевода медицинских текстов // Мир науки, культуры, образования. 2022. № 4 (95). С. 178–181. URL: <https://cyberleninka.ru/article/n/pred-i-postredaktirovanie-mashinnogo-perevoda-meditsinskih-tekstov?ysclid=ldw44lkycw553753824>
3. **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. 64–73. URL: <https://cyberleninka.ru/article/n/postredaktirovanie-mashinnogo-perevoda-kak-aktualnoe-napravlenie-podgotovki-perevodchikov-v-vuzah/viewer>
4. **Toral A., Castilho S., Hu K., Way A.** Attaining the Unattainable? Reassessing Claims of Human Parity in Neural Machine Translation, 2018. URL: <https://arxiv.org/abs/1808.10432>
5. **Toral A., Oliver A., Ribas-Bellestín P.** Machine Translation of Novels in the Age of Transformer // Maschinelle Übersetzung für Übersetzungsprofis, edited by Jörg Porsiel. Berlin: BDÜ Fachverlag, 2020. Pp. 276–296. URL: <https://www.researchgate.net/publication/346557116>
6. **de Almeida G., O'Brien S.** Analysing Post-Editing Performance: Correlations with Years of Translation Experience // Proceedings of the 14th Annual conference of the European Association for Machine Translation, May 27-28, 2010, Saint Raphael, France – European Association for Machine Translation, 2010. URL: <https://www.aclweb.org/anthology/2010.eamt-1.19>
7. **Ачкасов А.В.** Индустрия локализации и подготовка переводчиков // Материалы VIII Международной научной конференции «Индустрия перевода». Пермь, 2016. С. 13–19. URL: <https://elibrary.ru/item.asp?id=26181052&ysclid=ldwxzskzkc395978920>
8. **Беляева Л.Н.** Лингвистические технологии в современном сетевом пространстве: language worker в индустрии локализации. СПб.: Книжный дом, 2016. 134 с. URL: <https://search.rsl.ru/ru/record/01008793133?ysclid=ldw1rssz1875247094>
9. **Koponen M., Sulubacak U., Vitikainen K., Tiedemann J.** MT for subtitling: User evaluation of post-editing productivity // EAMT-2020, Proceedings of the 22nd Annual Conference of the European Association for Machine Translation. Lisboa, Portugal, 2020. Pp.115–124. URL: <https://www.researchgate.net/publication/348688005>
10. **Alves N., Marques M., Guimaraes P., Almeida J.A, Canario R.** Education and Training Courses in Portugal: Another kind of schooling? 2021. URL: <https://www.researchgate.net/publication/351050875>
11. **Стефанов С.И.** Реклама и полиграфия: опыт словаря-справочника. М.: Гелла принт, 2004. 340 с. URL: <https://search.rsl.ru/ru/record/01002486074?ysclid=ldw3y0ofmq501451819>
12. **Колесникова О.И.** Обучение студентов литературному редактированию в контексте развития художественно-языковой компетенции // Ярославский педагогический вестник. 2016. № 2. С. 60–63. URL: http://vestnik.yspu.org/releases/2016_2/14.pdf
13. **Накорякова К.М.** Справочник по литературному редактированию для работников массовой информации. М.: Флинта: Наука, 2010. 200 с. URL: <https://search.rsl.ru/ru/record/01004396714?ysclid=ldw3w40ial722575012>
14. **Хлопунова О.В., Цаканян А.А.** Профессиональные компетенции современного редактора: лидер или менеджер // Вестник Волжского университета имени В.Н. Татищева. 2019. № 4 (1).



С. 1–10. URL: <https://cyberleninka.ru/article/n/professionalnye-kompetentsii-sovremennogo-redaktoera-lider-ili-menedzher?ysclid=ldw1g8crls909581657>

15. **Фролова В.И.** О меняющейся роли редактора в эпоху обновления интернет-коммуникации // Вестник Волжского университета им. В.Н. Татищева. 2015. № 4 (19). С. 51–56. URL: <https://cyberleninka.ru/article/n/o-menyayuscheysya-rol-i-redaktora-v-epohu-obnovleniya-internet-kommunikatsii?ysclid=ldw11592hr40120245>

16. **Beliaeva L.N., Kamshilova O.N.** Problems and Perspectives of Language Worker Professional Training // International Journal of Open Information Technologies. 2018. Vol. 6. № 3. Pp. 35–42. URL: <http://injoit.org/index.php/j1/article/view/641>

17. **Moorkens J., O'Brien S.** Post-Editing Evaluations: Trade-offs between Novice and Professional Participants // EAMT 2015. Proceedings of the 18th Annual Conference of the European Association for Machine Translation. Antalya, Turkey, May 11-13, 2015. Pp. 75–81. URL: <https://www.researchgate.net/publication/275031846>

18. **Zaretskaya A.** The Use of Machine Translation among Professional Translators // Proceedings of the EXPERT Scientific and Technological Workshop, Malaga, 26th and 27th June 2015: Editions Tradulex, Geneva, 2015. Pp. 1–13. URL: <https://www.researchgate.net/publication/283667234>

19. **Кудашев И.С.** Проектирование переводческих словарей специальной лексики. Helsinki: Univ. of Helsinki, Dep. of transl. studies, 2007. 443 с. URL: <https://helda.helsinki.fi/bitstream/handle/10138/19272/proektir.pdf>

20. **Vasiljevs A., Pinnis M., Gornostay T.** Service model for semi-automatic generation of multilingual terminology resources // Terminology and Knowledge Engineering. 2014. Pp. 67–76. URL: <https://www.researchgate.net/publication/266565145>

21. **Vieira L.N.** Post-Editing of Machine Translation // M. O'Hagan (Ed.), The Routledge Handbook of Translation and Technology. Routledge, 2019. Pp. 319–335. URL: <http://www.routledge.com/9781138232846>

22. **Фельде О.В.** Эффективное речевое общение (базовые компетенции). Словарь-справочник под редакцией А.П. Сковородникова. Красноярск: Сибирский федеральный университет, 2012. URL: <http://elib.sfu-kras.ru/handle/2311/63722>

23. **Ortiz-Martinez D.** Online Learning for Statistical Machine Translation // Computational Linguistics. 42 (1). 2016. Pp. 121–161. URL: http://dx.doi.org/10.1162/COLI_a_00244

24. **Báez C.T., Schaeffer M., Carl M.** Experiments in Non-Coherent Post-editing // The Proceedings of the First Workshop on Human-Informed Translation and Interpreting Technology (HiT-IT). Varna, Bulgaria, Sept 7, 2017. Pp. 11–20. URL: <https://www.researchgate.net/publication/322031657>

25. **Беляева Л.Н., Джепа Т.Л., Зак Г.Н., Камшилова О.Н., Нымм В.Р., Разумова В.В.** Автоматизированное рабочее место филолога в структуре образовательного пространства современного вуза / коллективная монография. СПб.: Книжный дом, 2013. 127 с. URL: <https://search.rsl.ru/ru/record/01006589603?ysclid=ldw33c4dsz595025912>

26. **Беляева Л.Н., Джепа Т.Л.** Автоматизированное рабочее место переводчика: лингвистические ресурсы и технологии // Структурная и прикладная лингвистика. 2012. № 9. С. 109–128. URL: <https://www.elibrary.ru/item.asp?id=20351467&ysclid=ldw3as3jgx839950183>

27. **Stambolieva M.** Corpus Linguistics, Translation and Error Analysis // Proceedings of the Human-Informed Translation and Interpreting Technology Workshop (HiT-IT 2019)/ Varna, Bulgaria. 2019. Pp. 98–104. URL: <https://www.researchgate.net/publication/337854111>

28. **Koehn P.** Shared task: Statistical machine translation for European languages // ACL Workshop on Parallel Texts. 2005. Pp. 79–86. URL: <https://www.researchgate.net/publication/234795504>

29. **Tiedemann J.** Parallel Data, Tools and Interfaces in OPUS // Proceedings of the Eighth International Conference on Language Resources and Evaluation (LREC'12). 2012. Pp. 2214–2218. URL: http://www.lrec-conf.org/proceedings/lrec2012/pdf/463_Paper.pdf

30. **Abney S., Bird S.** The Human language Project: building a universal corpus of world's languages. In: Proceedings of the 48th ACL. 2010. Pp. 88–97. URL: <https://www.researchgate.net/publication/220873435>



REFERENCES

- [1] **B.Ye. Starichenko**, Tsifrovizatsiya obrazovaniya: illyuzii i ozhidaniya [Digitalization of education: illusions and expectations], *Pedagogicheskoye obrazovaniye v Rossii*, 3 (2020) 49–586. Available at: <https://cyberleninka.ru/article/n/tsifrovizatsiya-obrazovaniya-illyuzii-i-ozhidaniya?ysclid=ldvzz4y1cs241893073>
- [2] **N.A. Abrosimova, Ye.A. Shchelokova**, Pred- i postredaktirovaniye mashinnogo perevoda meditsinskikh tekstov [Pre- and post-editing of machine translation of medical texts], *Mir nauki, kultury, obrazovaniya*. 4 (95) (2022) 178–181. Available at: <https://cyberleninka.ru/article/n/pred-i-postredaktirovanie-mashinnogo-perevoda-meditsinskih-tekstov?ysclid=ldw44lkycw553753824>
- [3] **N.V. Nechayeva, S.Yu. Svetova**, Postredaktirovaniye mashinnogo perevoda kak aktualnoye napravleniye podgotovki perevodchikov v vuzakh [Post-editing machine translation as a new activity for teaching translation at universities], *Voprosy metodiki prepodavaniya v vuze*. 7 (25) (2018) 64–73. Available at: <https://cyberleninka.ru/article/n/postredaktirovanie-mashinnogo-perevoda-kak-aktualnoe-napravlenie-podgotovki-perevodchikov-v-vuzah/viewer>
- [4] **A. Toral, S. Castilho, K. Hu, A. Way**, Attaining the Unattainable? Reassessing Claims of Human Parity in Neural Machine Translation, (2018). Available at: <https://arxiv.org/abs/1808.10432>
- [5] **A. Toral, P. Oliver**, Ribas-Bellestín, Machine Translation of Novels in the Age of Transformer, *Maschinelle Übersetzung für Übersetzungsprofis*, edited by Jörg Porsiel. Berlin: BDÜ Fachverlag, 2020. Pp. 276–296. Available at: <https://www.researchgate.net/publication/346557116>
- [6] **G. de Almeida, S. O'Brien**, Analysing Post-Editing Performance: Correlations with Years of Translation Experience, *Proc. of the 14th Annual conference of the European Association for Machine Translation*, May 27–28, 2010, Saint Raphael, France, 2010. Available at: <https://www.aclweb.org/anthology/2010.eamt-1.19>
- [7] **A.V. Achkasov**, Industriya lokalizatsii i podgotovka perevodchikov [Localization industry and translator training], *Materialy VIII Mezhdunarodnoy nauchnoy konferentsii «Industriya perevoda»* [“Translation Industry”]. Perm, 2016, pp. 13–19. Available at: <https://elibrary.ru/item.asp?id=26181052&ysclid=ldwxzskzkc395978920>
- [8] **L.N. Belyayeva**, Lingvisticheskiye tekhnologii v sovremennom setevom prostranstve: language worker v industrii lokalizatsii [Linguistic technologies in modern network space: language worker in localization industry], *Knizhnyy dom*, SPb, 2016. Available at: <https://search.rsl.ru/ru/record/01008793133?ysclid=ldw1rssh1875247094>
- [9] **M. Koponen, U. Sulubacak, K. Vitikainen, J. Tiedemann**, MT for subtitling: User evaluation of post-editing productivity, *EAMT-2020, Proc. of the 22nd Annual Conference of the European Association for Machine Translation*. Lisboa, Portugal, 2020, pp. 115–124. Available at: <https://www.researchgate.net/publication/348688005>
- [10] **N. Alves, M. Marques, P. Guimaraes, J.A. Almeida, R. Canario**, Education and Training Courses in Portugal: Another kind of schooling? 2021. Available at: <https://www.researchgate.net/publication/351050875>
- [11] **S.I. Stefanov**, Reklama i poligrafiya: opyt slovarya-spravochnika [Advertising and printing: the experience of a reference dictionary], *Gella print*, M., 2004. Available at: <https://search.rsl.ru/ru/record/01002486074?ysclid=ldw3y0ofmq501451819>
- [12] **O.I. Kolesnikova**, Obucheniye studentov literaturnomu redaktirovaniyu v kontekste razvitiya khudozhestvenno-yazykovoy kompetentsii [Students’ Training to the Literary Editing in the Context of the Artistic-Language Competence Development], *Yaroslavskiy pedagogicheskiy vestnik*. 2 (2016) 60–63. Available at: http://vestnik.yspu.org/releases/2016_2/14.pdf
- [13] **K.M. Nakoryakova**, *Spravochnik po literaturnomu redaktirovaniyu dlya rabotnikov massovoy informatsii* [A Handbook of Literary Editing for Media Professionals], Flinta: Nauka, M., 2010. Available at: <https://search.rsl.ru/ru/record/01004396714?ysclid=ldw3w40ial722575012>
- [14] **O.V. Khlopunova, A.A. Tsakanyan**, Professionalnyye kompetentsii sovremennogo redaktora: lider ili menedzher [Professional competencies of a modern editor: leader or manager], *Vestnik Volzhskogo universiteta imeni V.N. Tatishcheva*. 4 (1) (2019) 1–10. Available at: <https://cyberleninka.ru/article/n/professionalnye-kompetentsii-sovremennogo-redaktora-lider-ili-menedzher?ysclid=ldw1g8crls909581657>
- [15] **V.I. Frolova**, O menyayushcheysya roli redaktora v epokhu obnovleniya internet-kommunikatsii [On the changing role of editor in the age of internet communication renewal], *Vestnik Volzhskogo universi-*



teta im. V. N. Tatishcheva. 4 (19) (2015) 51–56. Available at: <https://cyberleninka.ru/article/n/o-menyay-uscheyasya-rol-i-redaktora-v-epohu-obnovleniya-internet-kommunikatsii?ysclid=ldw11592hr40120245>

[16] **L.N. Beliaeva, O.N. Kamshilova**, Problems and Perspectives of Language Worker Professional Training, *International Journal of Open Information Technologies*. 6 (3) (2018) 35–42. Available at: <http://injoit.org/index.php/j1/article/view/641>

[17] **J. Moorkens, S. O'Brien**, Post-Editing Evaluations: Trade-offs between Novice and Professional Participants, *EAMT 2015. Proc. of the 18th Annual Conference of the European Association for Machine Translation*. Antalya, Turkey, 2015, pp. 75–81. Available at: <https://www.researchgate.net/publication/275031846>

[18] **A. Zaretskaya**, The Use of Machine Translation among Professional Translators, *Proc. of the EXPERT Scientific and Technological Workshop*, Malaga: Editions Tradulex, Geneva, 2015, pp. 1–13. Available at: <https://www.researchgate.net/publication/283667234>

[19] **I.S. Kudashev**, *Proyektirovaniye perevodcheskikh slovary spetsialnoy leksiki [Designing LSP Dictionaries for Translators]*, Univ. of Helsinki, Dep. of transl. studies, Helsinki, 2007. Available at: <https://helda.helsinki.fi/bitstream/handle/10138/19272/proektir.pdf>

[20] **A. Vasiljevs, M. Pinnis, T. Gornostay**, Service model for semi-automatic generation of multilingual terminology resources, *Terminology and Knowledge Engineering*, 2014, pp. 67–76. Available at: <https://www.researchgate.net/publication/266565145>

[21] **L.N. Vieira**, Post-Editing of Machine Translation, *The Routledge Handbook of Translation and Technology*. Routledge, 2019, pp. 319–335. Available at: <http://www.routledge.com/9781138232846>

[22] **O.V. Felde**, *Effektivnoye rechevoye obshcheniye (bazovyye kompetentsii). Slovar-spravochnik pod redaktsiyey A.P. Skovorodnikova [Effective verbal communication (basic competencies): a reference dictionary, A.P. Skovorodnikov – ed.]*, Krasnoyarsk, 2012. Available at: <http://elib.sfu-kras.ru/handle/2311/63722>

[23] **D. Ortiz-Martinez**, Online Learning for Statistical Machine Translation, *Computational Linguistics*. 42(1) (2016) 121–161. Available at: http://dx.doi.org/10.1162/COLI_a_00244

[24] **C.T. Báez, M. Schaeffer, M. Carl**, Experiments in Non-Coherent Post-editing, *Proc. of the First Workshop on Human-Informed Translation and Interpreting Technology (HiT-IT)*. Varna, Bulgaria, 2017, pp. 11–20. Available at: <https://www.researchgate.net/publication/322031657>

[25] *Avtomatizirovannoye rabocheye mesto filologa v strukture obrazovatel'nogo prostranstva sovremennogo vuza [Automated workplace of a philologist in the structure of university educational space] / L.N. Belyayeva, T.L. Dzhepa, G.N. Zak, O.N. Kamshilova, V.R. Nymm, V.V. Razumova, Knizhnyy dom, SPb., 2013. Available at: <https://search.rsl.ru/ru/record/01006589603?ysclid=ldw33c4dsz595025912>*

[26] **L.N. Belyayeva, T.L. Dzhepa**, *Avtomatizirovannoye rabocheye mesto perevodchika: lingvisticheskiye resursy i tekhnologii [Translator's automated workplace: linguistic resources and technologies]*, *Strukturnaya i prikladnaya lingvistika*. 9 (2012) 109–128. Available at: <https://www.elibrary.ru/item.asp?id=20351467&ysclid=ldw3as3jgx839950183>

[27] **M. Stambolieva**, *Corpus Linguistics, Translation and Error Analysis*, *Proc. of the Human-Informed Translation and Interpreting Technology Workshop (HiT-IT 2019)* Varna, Bulgaria, 2019, pp. 98–104. Available at: <https://www.researchgate.net/publication/337854111>

[28] **P. Koehn**, Shared task: Statistical machine translation for European languages, *ACL Workshop on Parallel Texts*, 2005, pp. 79–86. Available at: <https://www.researchgate.net/publication/234795504>

[29] **J. Tiedemann**, **Parallel Data**, *Tools and Interfaces in OPUS*, *Proc. of the Eighth International Conference on Language Resources and Evaluation (LREC'12)*, 2012, pp. 2214–2218. Available at: http://www.lrec-conf.org/proceedings/lrec2012/pdf/463_Paper.pdf

[30] **S. Abney, S. Bird**, The Human language Project: building a universal corpus of world's languages, *Proc. of the 48th ACL*, 2010, pp. 88–97. Available at: <https://www.researchgate.net/publication/220873435>

СВЕДЕНИЯ ОБ АВТОРАХ / INFORMATION ABOUT AUTHORS

Камшилова Ольга Николаевна

Olga N. Kamshilova

E-mail: onkamshilova@gmail.com

ORCID: <https://orcid.org/0000-0002-1488-2206>



Беляева Лариса Николаевна

Larisa N. Beliaeva

E-mail: lauranbel@gmail.com

ORCID: <https://orcid.org/0000-0002-8622-4595>

Поступила: 16.01.2023; Одобрена: 27.02.2023; Принята: 17.03.2023.

Submitted: 16.01.2023; Approved: 27.02.2023; Accepted: 17.03.2023.