

УДК 81'322.4

Волкова Ирина Дмитриевна
кандидат филологических наук,
доцент кафедры теории и практики
перевода
Волгоградский государственный
университет
г. Волгоград, Россия

Irina Volkova
PhD in Philology,
Associate Professor of the Translation
Studies Department
Volgograd State University
Volgograd, Russia
romadina-irina@mail.ru

Новожилова Анна Алексеевна
кандидат филологических наук,
доцент кафедры теории и практики
перевода
Волгоградский государственный
университет
г. Волгоград, Россия

Anna Novozhilova
PhD in Philology,
Associate Professor of the Translation
Studies Department
Volgograd State University
Volgograd, Russia
novozilova@yandex.ru

Митягина Вера Александровна
доктор филологических наук,
профессор кафедры теории и практики
перевода
Волгоградский государственный
университет
г. Волгоград, Россия

Vera Mityagina
PhD in Philology,
Professor of the Translation Studies
Department
Volgograd State University
Volgograd, Russia
mityagina@mail.ru

Никитенко Татьяна Васильевна
кандидат филологических наук,
заведующий кафедрой теории
и практики перевода
Белорусский государственный
университет
г. Минск, Беларусь

Tatyana Nikitenko
PhD in Philology,
Head of the Translation Studies
Department
Belarusian State University
Minsk, Belarus
tatnikita@gmail.com

СПЕЦИФИКА ПРЕДРЕДАКТИРОВАНИЯ ПРЕЗЕНТАЦИОННЫХ ТЕКСТОВ ДЛЯ МАШИННОГО ПЕРЕВОДА (на материале английского и немецкого языков)

Статья посвящена исследованию особенностей процесса предредактирования презентационных текстов, размещаемых на сайтах различных организаций, для последующего перевода в системах МП. Актуальность данного исследования обусловлена частотностью практики использования сервисов МП для перевода вербального контента сайтов, как разработчиками сайтов, так и пользователями. Исследование проведено на материале контента одноязычных сайтов медицинских и культурно-досуговых учреждений Российской Федерации. С целью выявления технологии предредактирования использована система машинного перевода Google Translate, и проанализирован перевод

презентационных текстов, выполненный данной системой на английский и немецкий языки. В результате проведенного анализа были выявлены ошибки, допускаемые системой Google Translate при работе с языковыми парами русский-английский, русский-немецкий, и предложены техники редактирования исходного текста с учетом данных ошибок.

Ключевые слова: предредактирование; презентационный текст; машинный перевод; многоязычный сайт; постредактирование; исходный текст.

PRE-EDITING OF PRESENTATIONAL TEXTS FOR MACHINE TRANSLATION (Based on the English and German Languages)

The article explores the pre-editing of presentational texts published on the websites of various organizations, intended for subsequent translation using machine translation (MT) systems. The relevance of this study stems from the frequent use of MT services for translating website content, both by website developers and users. The research is based on the content of monolingual websites of medical and cultural-leisure institutions of the Russian Federation. To identify pre-editing techniques, the study employed the Google Translate system and analyzed the translation of presentational texts into English and German. The analysis revealed errors made by Google Translate when dealing with Russian-English and Russian-German language pairs. The techniques for editing the source text to account for these errors have been proposed.

Key words: pre-editing; presentational text; machine translation; multilingual website; post-editing; source text.

Исследование выполнено при финансовой поддержке Российского научного фонда в рамках научного проекта № 25-28-00223 «Переводоведческое обоснование многоязычия презентационного контента в парадигме ИИ» / The research was supported by Russian Science Foundation (project No. 25-28-00223) “Translatological substantiation of multilinguality of presentational content in the AI paradigm”. (<https://rscf.ru/en/project/25-28-00223/>)

Процесс совершенствования алгоритмов осуществления машинного перевода характеризуется очень высокой скоростью, чем и обусловлена постановка проблем теоретического осмысления и оценки возможностей новейших систем машинного перевода (МП). Несмотря на постоянное развитие данных систем, роль специалиста-переводчика в создании переводного контента является очень важной и зачастую заключается в выполнении постредактирования машинного перевода. На сегодняшний день постредактирование является неотъемлемой составляющей процесса перевода, поскольку возможности современных программ МП позволяют специалистам значительно сократить сроки выполнения перевода, и выполнение последующего редактирования занимает гораздо меньше времени по сравнению с переводом «с нуля».

Однако наряду с распространенной практикой постредактирования актуальным также представляется процесс «предредактирования» (от англ. *pre-editing*), который, согласно своему названию, осуществляется в отличие

от постредактирования перед загрузкой исходного текста в программу машинного перевода. В исследованиях, посвященных автоматизации процесса перевода, отмечается, что предредактирование во многом способствует повышению эффективности и оптимизации перевода в целом и последующего редактирования в частности [1; 2; 3]. Многие переводоведы отмечают, что выполнение предредактирования особенно важно при работе со специальными текстами, насыщенными специфической терминологией и сложным синтаксисом [4; 5]. Кроме того, переводчикам в последнее время зачастую приходится работать с изначально «дефектными» текстами, в которых содержатся ошибки и неточности и которые нуждаются в обязательном предредактировании. Как отмечают У. А. Андреевкова и О. Г. Шевченко, «причиной этому служат экономия на подготовке документации, привлечение не носителей языка, отсутствие контроля, низкая компетентность авторов, а также массовое использование автоматизированных систем перевода» [6, с. 105–106].

Цель настоящего исследования заключается в определении эффективной технологии предредактирования презентационных текстов, позволяющей учитывать необходимость последующего перевода на иностранные языки уже на этапе создания исходного текста. Актуальность данного исследования обусловлена частотностью практики использования сервисов МП для перевода вербального контента сайтов, как разработчиками сайтов, так и пользователями.

Предредактирование направлено на «улучшение» исходного текста, на повышение его качества, что позволяет сократить или даже полностью исключить необходимость постредактирования. Что же в данном контексте следует понимать под качеством исходного текста? Процесс предредактирования предполагает подробный анализ исходного текста, выполняемый переводчиком или редактором с точки зрения функционирования систем МП и прогнозирования потенциальных ошибок. Таким образом, повысить качество исходного текста в процессе предредактирования означает обеспечить высокую степень его переводимости в программах МП, что, в зависимости от типа и жанра текста, может подразумевать сокращение длины предложений, обеспечение единообразия терминологии, пояснение или опущение культурно-специфичных выражений и т. д.

Актуальным данное направление деятельности переводчиков представляется для работы с многоязычными презентационными материалами, размещаемыми на сайтах различных организаций, что позволяет оптимизировать и унифицировать общий процесс создания переводного контента при минимальном постредактировании. Кроме того, распространенной практикой Интернет-пользователей является также использование систем МП при просмотре страниц иноязычных сайтов, на которых отсутствует версия на языке пользователя, в связи с чем важно обеспечить высокую степень переводимости презентационных гипертекстов в данных программах.

Безусловно для создания «модели» презентационного текста, отличающегося высокой степенью переводимости в системах МП необходимо учитывать характеристики презентационных текстов и ошибки, допускаемые машинными переводчиками. Материалом исследования моделирования и технологии предредактирования презентационных текстов послужил вербальный контент сайтов российских государственных учреждений, включая медицинские учреждения – Волгоградский областной клинический перинатальный центр № 2 (г. Волгоград), Московский многопрофильный научно-клинический центр им. С. П. Боткина (г. Москва), Национальный медицинский исследовательский центр здоровья детей (г. Москва) и их переводы на английский и немецкий языки, выполненные системой машинного перевода Google Translate. Общий объем проанализированного материала составляет около 30 000 п.з. Для анализа использован вербальный контент соответствующих характеристикам презентационного дискурса веб-страниц («О центре», «О перинатальном центре», «О ММНКЦ им. С. П. Боткина» и др.).

При переводе презентационных гипертекстов медицинской тематики, размещенных на сайтах указанных учреждений с русского на английский и немецкий язык с помощью онлайн-сервиса Google Translate были выявлены следующие ошибки. Во-первых, следует отметить вариативность перевода названия медицинского учреждения ММНКЦ им. С. П. Боткина, которая во многом связана с исходной вариативностью формулировок:

Исходный текст [7]	Машинный перевод
<p>ММНКЦ им. С.П. Боткина – крупнейшее многопрофильное лечебное и научно-образовательное учреждение Москвы.</p>	<p>англ. <i>MMNCC named after S.P. Botkin is the largest multidisciplinary medical and scientific-educational institution in Moscow.</i></p>
	<p>нем. <i>Das S.P. Botkin MMNCC ist die größte multidisziplinäre medizinische und wissenschaftlich-pädagogische Einrichtung in Moskau.</i></p>
<p>ММНКЦ имени С.П. Боткина – один из 5 опорных центров онкологической помощи в Москве.</p>	<p>англ. <i>SP Botkin MMNCC is one of the 5 reference centers for oncological care in Moscow.</i></p>
	<p>нем. <i>Das S.P. Botkin MMNCC ist eines der fünf Referenzzentren für onkologische Versorgung in Moskau.</i></p>
<p>В Боткинской ежегодно принимают порядка 1 миллиона амбулаторных пациентов.</p>	<p>англ. <i>Botkinskaya annually receives about 1 million outpatients.</i></p>
	<p>нем. <i>Das Botkinskaya-Krankenhaus empfängt jährlich rund eine Million ambulante Patienten.</i></p>

<p><i>В 2017 году Боткинская больница признана лучшим многопрофильным стационаром Москвы.</i></p>	<p>англ. <i>In 2017, Botkinskaya Hospital was recognized as the best multidisciplinary hospital in Moscow.</i></p>
	<p>нем. <i>2017 wurde das Botkinskaya-Krankenhaus als bestes multidisziplinäres Krankenhaus Moskaus ausgezeichnet.</i></p>
<p><i>Каждая восьмая операция в Боткинской относится к категории высокотехнологичной медицинской помощи.</i></p>	<p>англ. <i>Every eighth operation at Botkin falls into the category of high-tech medical care.</i></p>
	<p>нем. <i>Jede achte Operation im Botkinskaya-Krankenhaus fällt in die Kategorie der hochtechnologischen medizinischen Versorgung.</i></p>
<p><i>С 2019 года правительство Москвы осуществляет в Боткинской больнице масштабную программу комплексного развития.</i></p>	<p>англ. <i>Since 2019, the Moscow government has been implementing a large-scale comprehensive development program at Botkin Hospital.</i></p>
	<p>нем. <i>Seit 2019 führt die Moskauer Regierung ein umfassendes Entwicklungsprogramm für das Botkinskaya-Krankenhaus durch.</i></p>

Как показывают приведенные примеры, многие из переведенных названий на английском и немецком языках являются неадекватными и сложными для восприятия. Официальное название на английском – *Botkin Hospital*, но оно встречается в переводном тексте только один раз. При переводе на немецкий машинный переводчик использовал два варианта – транслитерацию для аббревиатуры *ММНКЦ*, и имени собственного *Боткинская* и дословный перевод слова *больница*, при этом ни один, ни другой вариант не является правильным, поскольку даже транслитерация была осуществлена по правилам английского языка. В данном случае при создании исходного русскоязычного текста необходимо использовать единообразный вариант названия учреждения (*Боткинская больница*), без употребления аббревиатур (*ММНКЦ*) и сокращений (*Боткинская*), что позволит значительно снизить вариативность номинации. Удачный пример подобной номинации демонстрируют тексты, размещенные на сайте Национального медицинского исследовательского центра здоровья детей: используется краткое обозначение *Центр*, которое находит соответствующее отражение в переводной версии *'The Center' / 'das Zentrum'*: *В штате Центра – высококвалифицированные специалисты – 'The Center employs highly qualified specialists' / 'Das Zentrum beschäftigt hochqualifizierte Fachkräfte'*; *Все виды исследований в Центре являются общедоступными – 'All types of research at the Center are publicly available' / 'Alle Forschungsarbeiten des Zentrums sind öffentlich zugänglich'*;

Все подразделения Центра имеют современное научное и медицинское оборудование – ‘All departments of the Center have modern scientific and medical equipment’ / ‘Alle Abteilungen des Zentrums verfügen über moderne wissenschaftliche und medizinische Ausstattung’ и др. [8].

Во-вторых, использование сложных синтаксических конструкций и длинных высказываний также создает сложности для восприятия переводного текста. Для презентационных текстов на английском и немецком языках нехарактерны длинные предложения, однако, как показывает анализ переводного текста, Google Translate не членит высказывания, сохраняя исходную конструкцию. Ниже представлены примеры исходного и отредактированного фрагментов оригинального и переводных текстов.

Пример 1

Исходный текст [7]	Машинный перевод
<p><i>Принимают пациентов Межокружной нефрологический центр (в отреставрированном историческом корпусе № 11), самый красивый и просторный Центр амбулаторной онкологической помощи (ЦАОП) для пациентов Западного округа, хирургический корпус № 7, где размещен Городской центр компетенций по гнойной травматологии и хирургии и хирургический стационар кратковременного пребывания (СКП).</i></p>	<p>англ. <i>The Interdistrict Nephrology Center (in the restored historical building No. 11), the most beautiful and spacious Outpatient Oncology Care Center (O OCC) for patients of the Western District, Surgical Building No. 7, where the City Competence Center for Purulent Traumatology and Surgery and the Short-Stay Surgical Hospital (SSH) are located, accept patients.</i></p> <p>нем. <i>Das Interdistrikte Nephrologische Zentrum (im restaurierten historischen Gebäude Nr. 11), das schönste und geräumigste Ambulante Onkologische Versorgungszentrum (O OCC) für Patienten des Westbezirks, das Chirurgiegebäude Nr. 7, in dem sich das Städtische Kompetenzzentrum für eitrige Traumatologie und Chirurgie und das Kurzzeitchirurgische Krankenhaus (SSH) befinden, nimmt Patienten auf.</i></p>
Отредактированный текст	Машинный перевод
<p><i>Межокружной нефрологический центр принимает пациентов в отреставрированном историческом корпусе № 11. Это самый красивый и просторный Центр амбулаторной онкологической помощи (ЦАОП) для пациентов Западного округа. Хирургический корпус № 7 также открыт для</i></p>	<p>англ. <i>The Interdistrict Nephrology Center accepts patients in the restored historical building No. 11. This is the most beautiful and spacious Outpatient Oncology Care Center (O OCC) for patients in the Western District. Surgical building No. 7 is also open to patients. It houses the City Competence Center for Purulent Traumatology and Surgery and the short-stay surgical hospital (SSH).</i></p>

<p><i>пациентов. В нем размещен Городской центр компетенций по гнойной травматологии и хирургии и хирургический стационар кратковременного пребывания (СКП).</i></p>	<p>нем. <i>Das Interdistrikt-Nephrologiezentrum empfängt Patienten im restaurierten historischen Gebäude Nr. 11. Es ist das schönste und geräumigste ambulante onkologische Versorgungszentrum (OOC) für Patienten im Westbezirk. Auch das chirurgische Gebäude Nr. 7 steht Patienten offen. Es beherbergt das Städtische Kompetenzzentrum für eitrige Traumatologie und Chirurgie sowie das Kurzzeit-Chirurgiekrankenhaus (SSH).</i></p>
--	--

В переводе исходного предложения подлежащее и сказуемое разделены большим количеством дополнительной информации, что нарушает связность и цельность конструкции. В предложенной отредактированной версии вместо одного предложения представлены четыре, что позволяет машинному переводчику создать адекватный и удобный для восприятия перевод на обоих языках.

Пример 2

Исходный текст [9]	Машинный перевод
<p><i>Основной задачей центра является оказание специализированной медицинской помощи женщинам репродуктивного возраста, беременным с высокой степенью риска развития осложнений в период беременности, родов и послеродовом периоде, роженицам и новорожденным, недоношенным детям, особенно новорожденным с критически низкой массой тела при рождении, родившимся как в перинатальном центре, так и других родовспомогательных учреждениях Волгоградской области и переведенных в него, в том числе оказание высокотехнологичной помощи.</i></p>	<p>англ. <i>The main objective of the center is to provide specialized medical care to women of reproductive age, pregnant women with a high risk of complications during pregnancy, childbirth and the postpartum period, women in labor and newborns, premature babies, especially newborns with critically low birth weight, born both in the perinatal center and other maternity institutions of the Volgograd region and transferred to it, including the provision of high-tech care.</i></p> <p>нем. <i>Das Hauptziel des Zentrums besteht darin, Frauen im gebärfähigen Alter, schwangeren Frauen mit einem hohen Risiko von Komplikationen während der Schwangerschaft, der Geburt und im Wochenbett, Frauen in den Wehen und Neugeborenen, Frühgeborenen, insbesondere Neugeborenen mit kritisch niedrigem Geburtsgewicht, die sowohl im Perinatalzentrum als auch in anderen Entbindungseinrichtungen der Region Volgograd geboren und dorthin verlegt werden, eine spezialisierte medizinische Versorgung zu bieten, einschließlich der Bereitstellung einer High-Tech-Versorgung.</i></p>

Отредактированный текст	Машинный перевод
<p><i>Основной задачей центра является оказание специализированной высокотехнологичной медицинской помощи <...>.</i></p>	<p>англ. <i>The main objective of the center is to provide specialized high-tech medical care_...>.</i></p> <p>нем. <i>Das Hauptziel des Zentrums ist die Bereitstellung spezialisierter medizinischer Hightech-Versorgung ...>.</i></p>

В примере 2 также наблюдается разделение связанных смысловых элементов большим объемом информации. В начале предложения указано *Основной задачей центра является оказание специализированной медицинской помощи... >*, далее идет перечисление категорий пациентов, которым оказывается помощь, и в самом конце предложения – *<... в том числе оказание высокотехнологичной помощи*. В предлагаемом отредактированном варианте фраза *<... в том числе оказание высокотехнологичной помощи* встроена в начало предложение путем использования прилагательного: *Основной задачей центра является оказание специализированной высокотехнологичной медицинской помощи... >*, что позволяет добиться адекватного перевода на английский и немецкий языки, который будет намного легче восприниматься получателями.

Пример 3

Исходный текст [8]	Машинный перевод
<p><i>Центр объединяет детских врачей страны для совместной клинической, научной и учебной деятельности: ежегодно проводятся конгрессы педиатров России, тематические научно-практические конференции, школы для специалистов.</i></p>	<p>англ. <i>The center unites pediatricians of the country for joint clinical, scientific and educational activities: congresses of pediatricians of Russia, thematic scientific and practical conferences, schools for specialists are held annually.</i></p> <p>нем. <i>Das Zentrum vereint Kinderärzte des Landes für gemeinsame klinische, wissenschaftliche und pädagogische Aktivitäten: Jährlich finden Kongresse der Kinderärzte Russlands, thematische wissenschaftliche und praktische Konferenzen sowie Schulen für Spezialisten statt.</i></p>

Отредактированный текст	Машинный перевод
<p><i>Центр объединяет детских врачей страны для совместной клинической, научной и учебной деятельности. Ежегодно проводятся конгрессы педиатров России, тематические научно-практические конференции, школы для специалистов.</i></p>	<p>англ. <i>The center unites pediatricians of the country for joint clinical, scientific and educational activities. Congresses of pediatricians of Russia, thematic scientific and practical conferences, schools for specialists are held annually.</i></p> <p>нем. <i>Das Zentrum vereint Kinderärzte des Landes für gemeinsame klinische, wissenschaftliche und pädagogische Aktivitäten. Jährlich finden Kongresse russischer Kinderärzte, thematische wissenschaftliche und praktische Konferenzen sowie Schulungen für Fachkräfte statt.</i></p>

В примере 3 также показано изменение структуры исходного предложения в процессе предредактирования путем членения на более короткие синтагмы.

Стоит обратить внимание на использование в исходном тексте аббревиатур, характерных для медицинского дискурса. В тексте о ММНКЦ им. С. П. Боткина их достаточно много, и соответственно, все они сохранены машинным переводчиком: *высокотехнологичная медицинская помощь (ВМП) – ‘high-tech medical care (HMC)’ / ‘Hightech-Medizinische Versorgung (HMC)’; Центр амбулаторной онкологической помощи (ЦАОП) – ‘the Center for Outpatient Oncological Care (CAOC)’ / ‘Zentrum für Ambulante Onkologische Versorgung (CAOC)’; стационар кратковременного пребывания (СКП) – ‘Short-Stay Surgical Hospital (SSH)’ / ‘Kurzzeitklinik (SSH)’; Медицинский симуляционный центр (МЦЦ) – ‘Medical Simulation Center (MSC)’ / ‘Medizinisches Simulationszentrum (MSC)’ [7].* Учитывая тот факт, что многие аббревиатуры являются локальными и образованы от переводных калек с русского языка, то в переводе в них нет необходимости, они перегружают текст, и в процессе предредактирования текста для машинного перевода их необходимо опустить. Кроме того, при переводе данных аббревиатур на немецкий язык Google использует англоязычный вариант, который не совпадает с первыми буквами компонентов, входящих в состав полных наименований на немецком языке, что является абсолютно некорректным и недопустимым без дополнительных пояснений и может служить доказательством тому, что в данном случае система не переводит напрямую, а привлекает английский как язык посредник.

Среди других параметров, которые следует учитывать при предредактировании или составлении презентационных текстов, которые будут переводиться системами машинного перевода на английский и/или немецкий языки,

можно выделить также следующие: грамматическая корректность (четкая структура предложений, минимизация деепричастных и длинных причастных оборотов, правильное согласование рода, числа и падежей и т. д.); конкретность и четкость (например, избегание местоимений с неясной референцией, учет наличия артиклей в английском и немецком языках); отказ от морфологической сложности (например, длинные цепочки родительного падежа, сложные глагольные формы); правильная пунктуация.

Таким образом, на основе проведенного анализа можно сделать вывод о языковых характеристиках, лежащих в основе модели презентационного текста на русском языке, отличающегося высокой степенью переводимости в системах машинного перевода. К таким характеристикам относится использование коротких предложений, единообразие лексической номинации, отсутствие аббревиатур, грамматическая и пунктуационная корректность. Решение научно-исследовательской проблемы разработки принципов предредактирования текста для машинного перевода и моделирования презентационного текста позволяет оптимизировать процессы создания многоязычного контента с использованием систем МП в рамках Интернет-дискурса и повысить его качество.

ЛИТЕРАТУРА

1. Hiraoka Y., Yamada M. Pre-editing plus neural machine translation for subtitling: effective pre-editing rules for subtitling of TED Talks // Proceedings of MT. Dublin MT Summit XVII. 2019. Vol. 2. P. 64–74.
2. Мэнсюе Ч. Влияние предредактирования на эффективность машинного перевода исторической литературы на примере перевода "Истории науки и техники Китая" // Вопросы истории. 2023. № 9–1. С. 198–207. DOI: 10.31166/VoprosyIstorii202309Statyi26.
3. Taufik A. Pre-editing of Google neural machine translation [Electronic resource] // Journal of English Language and Culture. 2020. Vol. 10. № 2. P. 64–74. DOI: <http://dx.doi.org/10.30813/jelc.v10i2.2137>.
4. Животова А. А., Бердонос В. Д. Оптимизационное предредактирование узкоспециальных русскоязычных текстов для их машинного перевода на английский язык // Информационные и математические технологии в науке и управлении. 2024. № 2 (34). С. 169–182. DOI: 10.25729/ESI.2024.34.2.016.
5. Коканова Е. С., Пак Н. С. Предредактирование текстов для машинного перевода как метод исследования английских конструкций с причастием в медицинском дискурсе // Когнитивные исследования языка. 2022. № 3 (50). С. 676–680.
6. Андреевкова У. А., Шевченко О. Г. Влияние аномалий дефектных текстов на машинный перевод и предредактирование дефектных текстов (на материале англоязычных инструкций) // Социальные и гуманитарные науки в условиях вызовов современности. Материалы III Всероссийской научно-практической конференции молодых ученых с международным участием. В 2-х частях. Комсомольск-на-Амуре, 2025. С. 105–107.

7. ММНКЦ им. С. П. Боткина – Московский многопрофильный научно-клинический центр им. С. П. Боткина [Электронный ресурс]. URL: <https://botkinmoscow.ru/about/o-bolnitse/botkinskaya-bolnitsa-segodnya/> (дата обращения: 03.07.2025).
8. НМИЦ здоровья детей – Национальный медицинский исследовательский центр здоровья детей [Электронный ресурс]. URL: <https://nczd.ru/about/> (дата обращения: 03.07.2025).
9. ВОКПЦ № 2 – Волгоградский областной клинический перинатальный центр № 2 [Электронный ресурс]. URL: <https://vokpc2.ru/about-vokpc2/>. (дата обращения: 03.07.2025).