

УДК 81'25

DOI 10.47388/2072-3490/lunn2025-72-4-59-74

РЕАЛИЗАЦИЯ ТЕКСТООБРАЗУЮЩИХ КАТЕГОРИЙ В МАШИННОМ ПЕРЕВОДЕ

О. В. Петрова, М. Ю. Родионова

Нижегородский государственный лингвистический университет
им. Н. А. Добролюбова, Нижний Новгород, Россия

В статье рассматривается проблема воспроизведения в машинном переводе таких свойств текста, как цельность и связность. Анализируются причины их нарушения и проявления этих нарушений на примере конкретных переводов, выполненных разными программами автоматического перевода. Показано, что в основе этого лежит неспособность машины воспринимать текст как единое структурно-семантическое и коммуникативное целое, все компоненты которого объединены многочисленными внутритекстовыми связями, а также отсутствие понимания машиной всех видов контекста. Контекст для программы зачастую ограничивается предметной областью текста, а когерентность и когезия в некоторых случаях не выходят за рамки абзаца. Показано, что эти категории нарушаются даже в пределах одного предложения. Результатом неспособности машины выявлять внутритекстовые связи становится их некорректное воспроизведение в переводе. Неверно установленные и неверно воспроизведенные внутритекстовые связи в ряде случаев приводят к содержательным и смысловым ошибкам, которые могут скрываться за внешне правильной формой. Кроме того, распространение машинного перевода, часто используемого без постредактирования, приводит к постепенному разрушению самого понятия «текст». Занятия по анализу текста выявляют у студентов-переводчиков те же проблемы восприятия, что и у компьютерных программ: многие студенты не соотносят значения слов и логических коннекторов с контекстом, не видят контекста за пределами абзаца, не воспринимают текст в контексте внеязыковой ситуации, затрудняются в выявлении внутритекстовых связей. Необходимо целенаправленно учить студентов восприятию текста, показывая на примере переведенных машиной текстов, к чему может приводить нарушение основных текстообразующих категорий.

Ключевые слова: текст; контекст; текстообразующие категории; когерентность; когезия; внутритекстовые связи.

Цитирование: Петрова О. В., Родионова М. Ю. Реализация текстообразующих категорий в машинном переводе // Вестник Нижегородского государственного лингвистического университета им. Н. А. Добролюбова. 2025. Вып. 4 (72). С. 59–74. DOI 10.47388/2072-3490/lunn2025-72-4-59-74.

Coherence and Cohesion in Machine Translation**Olga V. Petrova, Maria Yu. Rodionova**

Linguistics University of Nizhny Novgorod, Nizhny Novgorod, Russia

This paper addresses the problem of reproducing such fundamental textual properties as coherence and cohesion in machine translation. The research material consists of texts translated by three different computer translation programs, which are analyzed in order to identify the causes and consequences of a machine's inability to adequately reproduce these text-forming categories. The study demonstrates that the principal reason for violations of textuality lies in the machine's inability to perceive a text as a unified structural, semantic, and communicative whole whose components are interconnected by numerous intra-textual links. An additional factor is the machine's inability to comprehend context in its full complexity. For most computer programs, context is often reduced to the subject domain of the text, while coherence and cohesion may, in some cases, be maintained only at the sentence level. In other instances, these categories are violated even within a single paragraph. As a result of the machine's failure to identify intra-textual links, such links are frequently reproduced incorrectly in translation. Erroneously established or inadequately rendered intra-textual connections may lead to distortions of content and meaning, which are sometimes concealed behind an apparently correct grammatical or lexical form. Moreover, the widespread use of automatic translation, often without post-editing, contributes to the gradual erosion of the very notion of the "text" as a coherent and cohesive communicative unit. The paper also draws attention to pedagogical implications of these findings. Classroom experience in text analysis reveals that translation students often demonstrate perceptual difficulties similar to those of computer programs: they fail to relate word meanings and logical connectors to the broader context, do not perceive coherence beyond the paragraph level, overlook the extra-linguistic situation of the text, and experience difficulties in identifying intra-textual links. The study emphasizes the importance of deliberately teaching students to perceive texts as integrated wholes, using machine-translated texts as illustrative examples of the consequences that arise from violations of basic text-forming categories.

Key words: text; context; text-forming categories; coherence; cohesion; intra-textual links.

Citation: Petrova, Olga V., & Rodionova, Maria Yu. (2025) Coherence and Cohesion in Machine Translation. *LUNN Bulletin*, 4 (72), 59–74. DOI 10.47388/2072-3490/lunn2025-72-4-59-74.

1. Введение

Качество машинного перевода анализируется в многочисленных работах, авторы которых расходятся в оценках его сегодняшнего состояния и перспектив его развития. Но так или иначе все, от энтузиастов до скептиков, признают, что при несомненных успехах нейросетевого перевода результаты пока далеки от совершенства. Все признают, что в подавляющем большинстве случаев использовать компьютерный перевод без постредактирования нельзя. Постредактирование постепенно превращается в самостоятельную профессию, и, говоря о требованиях, которые сегодня предъявляются к постредакторам, А. Н. Малявина, В. В. Сдобников и А. А. Шеронов отмечают, что, «помимо знаний и компетентности, которые ожидаются от каждого переводчика и редактора, специалист в области ПРМП должен разбираться в типах ошибок, которые присущи

именно машинному переводу» (Малявина, Сдобников, Шеронов 2025: 91). Поэтому вполне естественно, что многие исследователи обращаются к изучению, классификации и типологизации ошибок, характерных для переводов, выполненных компьютерными программами.

Авторы статей отмечают, что проблемы часто возникают с семантикой, «так как при переводе семантических конструкций необходимы базы данных, которые в настоящее время не разработаны. Помимо этого, немалые затруднения вызывают переводы сложных грамматических, синтаксических и лексических конструкций» (Переходько, Мячин 2017: 96). Исследователи выделяют лексические (непереведенность отдельных лексических единиц или их неверный перевод), морфологические (неверные формы числа, падежа, времени, наклонения, степени сравнения, неправильный выбор предлогов), синтаксические ошибки (неправильное расположение слов в предложении) (Кочеткова, Ревина 2017: 109). Близкую классификацию приводит и П. С. Митин, анализирующий переводы с арабского языка: синтаксические, семантические, лексические и морфологические ошибки (Митин 2021).

В качестве недостатков отмечают неспособность программ обеспечить точную передачу идиом, грамматики и разговорных выражений (Качалова, Лишацкая 2022: 233). И даже признавая, что в ряде случаев программа способна создать полный перевод исходного текста без фактических и содержательных ошибок, исследователи констатируют, что пока еще не удается достичь адекватной степени идиоматичности и аутентичности переводного текста (Павлова, Рябова 2022: 258).

Оценивая степень приемлемости машинных переводов, авторы статей отмечают, что она оказывается разной при переводе разных типов текстов. По мнению ряда авторов, наилучших результатов при использовании машинного перевода можно достичь в случае текстов, написанных в техническом и официально-деловом стиле (Митин 2021; Ртищева 2017). И если в отношении технических текстов с этим мнением согласны не все, то неспособность компьютерных программ переводить художественные тексты практически ни у кого не вызывает сомнений. Объясняют это разными причинами — как общими для проблем машинного перевода, так и спецификой художественного текста. Так, Е. С. Фидельская отмечает значительное количество ошибок грамматического, лексического и стилистического характера, затрудняющих выполнение художественной задачи (Фидельская 2019: 86). И. И. Абдулганеева отмечает, что «особые трудности для перевода представляют национально-маркированные и эмоционально-экспрессивные единицы, дериваты со словообразовательными формантами интенсификации, а также предикативные конструкции сослагательного наклонения, эллиптические конструкции» (Абдулганеева 2022: 5). В наличии «образных выражений, слов в переносных значениях, метафор, метонимий, инвертированных синтаксических конструкций, неполных предложений и других стилистических приемов

и изобразительных средств» видит основную трудность при машинном переводе художественных текстов Е. М. Литвинова (Литвинова 2012: 63).

При всем разнообразии задач, которые ставят перед собой исследователи, типов текстов и языковых пар все приведенные мнения и оценки объединяет одно: авторы большинства работ практически все проблемы видят в трудностях передачи тех или иных единиц — лексических, грамматических или стилистических. И даже в случае перевода художественных текстов речь все равно идет о проблемах передачи каких-то единиц, а не текста как такового. Показательно, что в тех немногочисленных случаях, когда проводятся достаточно масштабные эксперименты с переводом художественных текстов (как, например, в Университете Массачусетса в Амхерсте [Thai, Karpinska, Krishna, Ray, Inghilleri, Wieting, et al. 2022]) и полученные в ходе этих экспериментов результаты высоко оцениваются (Губайловский 2022), речь идет лишь о переводе фрагментов текста, но никак не целостных текстов.

Судя по всему, о практическом использовании машинного перевода для художественных текстов сегодня речь всерьез не идет. Но в любом случае и научные, и технические, и публицистические тексты являются именно текстами, с присущими им текстовыми характеристиками, а поэтому естественно ожидать, что в результате их перевода должен также получиться именно текст, а не набор отдельных более или менее удачно переведенных предложений. Поэтому представляется интересным посмотреть, как справляются компьютерные программы с сохранением текстуральности, т. е. с воссозданием в переводе основных текстообразующих категорий.

2. Характеристика материала и методов исследования

Материалом для исследования послужили переводы целостных публицистических текстов, выполненные программами *Google Translate*, *DeepL* и «Яндекс Переводчик». Полученные переводы анализировались с точки зрения сохранения в них основных текстообразующих категорий.

3. Результаты исследования и их обсуждение

Став объектом специального лингвистического изучения в середине XX века, текст с тех пор получил множество научных определений. По свидетельству О. Н. Копытова, уже к 1990-м годам насчитывалось около 250 таких определений (Копытов 2011: 150). Эту же цифру, со ссылкой на В. А. Пищальникову и Ю. А. Сорокина (Пищальникова, Сорокин 1993), приводит М. В. Воронец, объясняя это как сложностью и разнообразием текстов, так и разнообразием подходов к изучению самого этого понятия (Воронец 2021).

Самое общее определение текста дает «Лингвистический энциклопедический словарь»: текст — это «объединённая смысловой связью последовательность знаковых единиц, основными свойствами которой являются связность

и цельность» (ЛЭС 1990: 507). И. Р. Гальперин в своем ставшем классическим определении текста пишет: «Текст — это произведение речетворческого процесса, обладающее завершенностью, объективированное в виде письменного документа, литературно обработанное в соответствии с типом этого документа, произведение, состоящее из названия (заголовка) и ряда особых единиц (сверхфразовых единств), объединенных разными типами лексической, грамматической, логической, стилистической связи, имеющее определенную целенаправленность и прагматическую установку» (Гальперин 2006: 18).

Характеризуя разные подходы к изучению текста, Г. Я. Солганик отмечает, что для лингвистики текст — «это результат целенаправленного речевого творчества, целостное речевое произведение», в котором «особенно важно подчеркнуть целостность как одну из главных особенностей текста» (Солганик 2005: 8).

Обобщая многочисленные исследования, посвященные выявлению сущности текста и его основных характеристик, А. А. Кожанов и Г. Н. Россихина пишут: «Многие отечественные лингвисты считают основными свойствами текста цельность, связность и отдельность» (Кожанов, Россихина 2013: 84).

Поскольку текст «выступает одновременно явлением лингвистическим, психологическим, семиотическим и даже физическим, имея особое синтаксическое, просодическое, прагматическое и в целом коммуникативное строение», в аспекте перевода необходимо учитывать не только чисто лингвистические его характеристики (Пищальникова, Сонин 2017: 265). Особое значение приобретает рассмотрение текста как инструмента коммуникации, при котором на первый план выдвигается «необходимость введения для его исследования особых единиц и категорий, связанных в современной лингвистике с понятием текстуальности» (там же: 272).

Очевидно, что при таком разнообразии подходов к определению самого понятия «текст» разными оказываются и наборы выделяемых категорий, получивших название текстообразующих. Так, И. Р. Гальперин выделяет следующие категории: информативность, членимость, когезия (понимаемая им как внутри-текстовые связи), континуум, автосемантия отрезков текста, ретроспекция и проспекция, модальность, интеграция и завершенность текста (Гальперин 2006). Л. Н. Мурзин и А. С. Штерн считают основными свойствами текста связность и цельность (Мурзин, Штерн 1991). Н. С. Валгина рассматривает в качестве основных конструктивных признаков текста целостность и связность, объясняя это тем, что текст как объект лингвистического исследования является прежде всего информационным и структурным единством, т. е. функционально завершенным речевым целым (Валгина 2003: 29). В. Дресслер и Р.-А. де Богранд уточняют понятие связности, вводя, наряду с когезией, когерентность, а также называют такие категории, как информативность, интенциональность и восприимчивость (*intentionality and acceptability*), ситуативность и интертекстуальность (De Beaugrande, Dressler 1981).

При этом, как отмечают В. А. Пищальникова и А. Г. Сонин, в центре внимания большинства исследователей оказываются «категории **когерентности / цельности / целостности** (относимой к плану содержания текста) и **когезии / связности** (план выражения)» (Пищальникова, Сонин 2017: 272). Сохранение именно этих характеристик при переводе текста и представляется интересным проверить в первую очередь.

Исходная гипотеза состояла в том, что любая компьютерная программа как при порождении текста, так и при переводе создает тексты, в которых нарушены текстообразующие категории. Эту гипотезу подтвердил ИИ. Нейросети *ChatGPT* были заданы два вопроса: 1. Может ли ИИ при порождении текста нарушать текстообразующие категории? 2. Могут ли нарушать текстообразующие категории программы машинного перевода?

Были получены следующие ответы:

1	2
<p>Искусственный интеллект (ИИ), особенно в контексте генерации текста, может нарушать текстообразующие категории по нескольким причинам:</p> <p>1. Отсутствие понимания контекста: ИИ не обладает истинным пониманием языка или контекста, как это делает человек. Он основывается на шаблонах и статистических вероятностях, что может привести к неуместным или несогласованным высказываниям.</p> <p>2. Генерация без логики: Иногда ИИ может создавать текст, который грамматически правильный, но лишен логической последовательности или смыслового содержания. Это может проявляться в виде нелепых или абсурдных утверждений.</p> <p>3. Ограниченность обучающего материала: Модели обучаются на больших объемах текстов, и если в этих данных присутствуют ошибки, предвзятости или неуместные примеры,</p>	<p>Да, программы машинного перевода могут нарушать текстообразующие категории по нескольким причинам:</p> <p>1. Проблемы с контекстом: Машинный перевод может не учитывать контекст, в котором используется фраза или слово. Это может привести к неправильному переводу, который не соответствует смыслу оригинала.</p> <p>2. Лексические и грамматические ошибки: Перевод может содержать ошибки в выборе слов или грамматических конструкциях, что делает текст трудным для понимания или даже абсурдным.</p> <p>3. Культурные различия: Машинный перевод может не учитывать культурные особенности и идиоматические выражения, что может привести к переводу, который звучит неестественно или неправильно в целевом языке.</p> <p>4. Стиль и тон: Программы перевода могут не передавать стиль и тон оригинального текста, что может изменить восприятие сообщения.</p>

<p>это может отразиться на качестве генерируемого текста.</p> <p>4. Креативность и оригинальность: ИИ может генерировать оригинальные идеи, но зачастую они могут не соответствовать установленным жанровым или стилистическим нормам, что также может восприниматься как нарушение текстообразующих категорий.</p> <p>5. Этические и культурные аспекты: ИИ может не учитывать культурные и этические нормы, что может привести к созданию текста, который воспринимается как оскорбительный или неуместный.</p>	<p>Например, формальный текст может быть переведен в разговорный стиль, что нарушает его текстообразующую категорию.</p> <p>5. Сложные структуры: Некоторые языки имеют сложные синтаксические структуры, которые трудно адекватно перевести. Это может привести к потере смысла или изменению структуры предложения.</p> <p>6. Неоднозначность: Слова и фразы могут иметь несколько значений, и машинный перевод может выбрать неправильное значение, что также нарушает текстообразующие категории.</p>
<p>Эти факторы подчеркивают сложность и многообразие языка, а также ограничения современных технологий ИИ.</p>	<p>Таким образом, хотя программы машинного перевода значительно улучшились за последние годы, они все еще могут сталкиваться с проблемами, которые влияют на качество и точность перевода.</p>

Обращение к ИИ объясняется не верой в его способность постичь некую абсолютную истину, а тем, что он отбирает и обобщает то, что по запрашиваемой теме пишут и публикуют в интернете. Таким образом, ответы ИИ отражают основные мнения и выводы, содержащиеся в научных статьях и докладах. Сформулированные в такой краткой форме, они позволяют понять, каковы преобладающие в научном сообществе взгляды на интересующую нас проблему.

Чтобы проверить обоснованность ответа насчет нарушения текстообразующих категорий при генерировании текстов, достаточно обратиться к такому сообщению, сгенерированному ИИ (*AL Overview Google*):



Книги Хайнлайна были переведены на английский язык, потому что они были написаны на английском языке в первую очередь. [Роберт А. Хайнлайн](#) был американским писателем-фантастом и переводил свои книги на английский язык самостоятельно.

ИИ нашел информацию о том, что Роберт Хайнлайн был американским писателем и писал на английском языке, но при этом генерирует текст, в котором одновременно сообщает, что его книги на этот язык переводились, причем переводил их он сам.

Первая из названных причин нарушения текстуальности явно не срабатывает: дело не в отсутствии понимания контекста — ИИ в данном случае сам его себе создал. Что касается логики, то здесь два, казалось бы, взаимоисключающих случая: ИИ устанавливает причинно-следственную связь между тем, что книги были написаны на английском языке, и тем, что они на английский язык переводились, и одновременно на основе того, что Хайнлайн был американским писателем, инферирует — в терминологии М. Е. Коровкиной и А. Л. Семёнова (Коровкина, Семенов 2022) — то, что он переводил книги на английский язык самостоятельно. Причины этого в процитированном ответе, который дал *ChatGPT*, не называются: просто констатируется вероятность логических нарушений.

Не оправдывается и третье объяснение — ограниченность обучающего материала. В данном случае материала было явно достаточно, но связность текста при этом нарушена. Нельзя генерирование такого текста объяснить и остальными двумя причинами — креативностью и отсутствием учета этических факторов.

Итак, из всего, что на вопрос о нарушении текстуальности при генерировании текста ответил ИИ, остается только сам факт такой возможности. Попытка сформулировать причины не удалась: дело, очевидно, в чем-то другом.

Посмотрим на ответы, которые *ChatGPT* дал по поводу нарушений текстуальности при переводе. Три из них — первый, второй и шестой — фактически сводятся к употреблению неверных слов, в результате ли ошибочного выбора лексико-семантического варианта (ЛСВ), являющегося лексическим соответствием определенного слова оригинала, или просто в результате неверно найденного соответствия. Не очень понятен смысл пятого ответа, поскольку предполагается, что все синтаксические структуры исходного языка (особенно «сложные», т. е. специфичные для данного языка) в базе у машинного переводчика должны быть. Но, так или иначе, речь опять идет об ошибках в передаче отдельных единиц — на этот раз грамматических.

Что касается шестого ответа, объясняющего возможные ошибки неоднозначностью слов и фраз, то здесь дело обстоит несколько сложнее. Сказать, что, если в базе будут все значения слова, машина обязательно выберет правильное, нельзя по меньшей мере по двум причинам. Первая — непонимание машиной контекста, а вторая — неспособность машины учитывать цель перевода. Сегодня, когда перевод все чаще становится потребностноориентированным (Шамилов 2023), получатель ждет не просто понятного текста, а текста, соответствующего его конкретным запросам. А узкопрофессиональных словоупотреблений в базе чаще всего нет, на что указывает О. И. Кузьмин, который пишет, что сегодня в машинном переводе необходим индивидуальный подход, т. е. «фокусировка

внимания на частных случаях употребления» (Кузьмин 2021: 51). Получатель же воспринимает отклонения от узкопрофессионального узуса как лексическую ошибку.

К тексту как таковому в какой-то степени относятся третий и четвертый ответы, в которых говорится, что машина может не учитывать культурные особенности текста и его стилистические характеристики. Действительно, без учета культурных особенностей текста невозможна адекватная адаптация текста к инокультурному реципиенту. Как отмечают А. Ю. Ивлева и Е. А. Кузнецова, «переводчик играет роль медиатора культур, благодаря которому достигается оказание такого воздействия на получателя, какое оригинал оказывает на адресата. Личность переводчика при этом проявляется через отбор языковых средств и переводческих приемов, делающих возможным перевыражение текста для иноязычного реципиента» (Ивлева, Кузнецова 2024: 60). Неправильный выбор таких средств действительно может привести к нарушению связности текста.

Понятно, что автоматический переводчик не обладает личностью. Однако *ChatGPT* в своем ответе ссылается на наличие идиоматических выражений, что опять возвращает нас к ошибкам в передаче отдельных единиц. К тому же идиоматические выражения для конкретной пары языков, так же как и синтаксические структуры, предположительно, должны быть в базе.

В целом ответы ИИ подтверждают обоснованность впечатления, что при анализе недостатков машинного перевода исследователи концентрируются на ошибках в передаче отдельных единиц.

Посмотрим на ситуацию с точки зрения воспроизведения машиной текстуальности, и в первую очередь тех текстообразующих категорий, в наличии и важности которых сходится большинство исследователей: когерентности / цельности / целостности и когезии / связности.

Отмеченная ИИ неспособность воспринимать контекст действительно приводит к появлению ошибок. И речь здесь идет о любых видах контекста — широком, узком, лингвистическом, ситуативном, имплицитном. Показательны в этом отношении ошибки, допущенные программой «Яндекс Переводчик» при переводе статьи о фильме Чаплина «Новые времена».

В описываемой сцене дело происходит на заводе:

*There are many delicious, yet profoundly pertinent moments, like his natural desire to snatch a moment's respite from **the belt** and steal a smoke in the men's room. The **televised image of the scowling man in the front office flashes on the wall** with a command to drop that cigarette and get back to work.*

«Яндекс Переводчик»:

*Здесь много восхитительных, но в то же время очень уместных моментов, например, его естественное желание на минутку отвлечься от **ремня** и украдкой покурить в мужском туалете. **Телевизионное изображение хмурого мужчины***

в приемной стряхивает пепел на стену, приказывая бросить сигарету и вернуться к работе.

Слово *belt* до этого в тексте уже встречалось, причем в контексте, не оставляющем сомнений в том, что это конвейер:

After screwing the bolts in the ever-faster belt, which attains a lunatic speed at the bidding of the polished, well-clad gentleman in the executive's office, Charlie goes in as crazy the tempo.

Машина, не воспринимая контекста, в обоих случаях выбрала вариант *ремень*. Но если получатель перевода на основе контекста в первом случае понял, что речь идет о конвейере, то он и повторное употребление слова *ремень* воспримет как обозначающее тоже конвейер. Нарушения когерентности не произошло.

Хуже обстоит дело с переводом второго предложения. Здесь уже происходит искажение смысла, действительно приводящее к нарушению когерентности. Теряется связь между героем, находящимся в уборной, и телевизионным изображением, якобы находящимся в чьей-то приемной. Непонятно, какой пепел на какую стену это изображение стряхивает и каким образом что-то приказывает. Действительно, среди множества значений глагола *to flash* есть такое, которое с натяжкой может означать стряхивание чего-то. Но оно никак не вписывается в контекст ситуации, описание которой в предыдущем предложении сама машина уже воспроизвела. Можно лишь предположить, что на выбор этого значения повлияло наличие в том же предложении слова *cigarette*, в результате чего было добавлено слово *пепел*. В итоге сцена описана как разворачивающаяся в приемной, где чье-то телевизионное изображение стряхивает пепел на стену. Таким образом, выбор варианта перевода основывался не на самом тексте, а на не имеющих никакого отношения к этому тексту содержащихся в базе ассоциативных связей между словами. Связи же телевизионного изображения с тем, где в это время находится и что делает герой фильма, машина не видит и, следовательно, не отражает в переводе.

Примеров того, что машина не воспринимает контекст, много. И в большинстве случаев это действительно приводит к нарушению когерентности. Но нарушение когерентности в переводе может возникать не из-за непонимания машиной контекста, а из-за того, что она неверно устанавливает внутритекстовые связи в тексте оригинала. Показателен в этом отношении перевод следующего фрагмента из статьи о необходимости пересмотреть правила присуждения Нобелевской премии:

*But talk to scientists in private, and many will grumble. The Nobels are a great way to get people interested in science, they'll say, and it's good that we have them. But there have been strange omissions, with people who should have won a prize denied. The grumpiest complain that despite the glitz of the prizes, and the ensuing media attention in articles like this one, **the rules that govern them** no longer reflect the way modern science works.*

Google Translate:

*Но поговорите с учеными наедине, и многие будут ворчать. Нобели — отличный способ заинтересовать людей наукой, скажут они, и хорошо, что они у нас есть. Но были странные упущения, когда людям, которые должны были выиграть приз, было отказано. Самые сварливые жалуются, что, несмотря на блеск призов и последующее внимание средств массовой информации к статьям, подобным этой, **правила, которыми они управляют**, больше не отражают то, как работает современная наука.*

При всем несоответствии *Нобелей* и *выигрывания призов* узусу русского языка общий смысл был бы понятен, если бы не ошибка в переводе *the rules that govern them*. Можно, конечно, говорить о грамматических проблемах — об отсутствии падежных форм у местоимений в английском языке и их наличии в русском. Но этой ошибки не было бы, если бы машина воспринимала смысловые связи. В оригинале местоимение *them* может относиться либо к *The Nobels*, либо к *prizes*, которые ничем управлять не могут. Компьютер же соотнес это местоимение с *scientists* (либо с *the grumpiest*), в результате чего они превратились в деятелей, управляющих с помощью каких-то правил. Случай, когда нарушение внутритекстовых связей привело к нарушению смысла.

Конечно, все ошибки перевода так или иначе проявляются в использовании неверных единиц. Но, как уже было показано, это не всегда приводит к нарушению текстообразующих категорий. Поэтому значительно важнее понять, чем эти ошибки бывают вызваны. Так, например, одно и то же предложение исходного текста (*Instead, its uniforms will be promoted as gender neutral, **even removing packaging** if the colour signifies a gender divide*) Google Translate и DeepL перевели по-разному, но в обоих случаях неверно:

Google Translate:

*Вместо этого его униформа будет продвигаться как гендерно-нейтральная, **даже если будет удалена упаковка**, если цвет означает гендерное разделение.*

DeepL:

*Вместо этого школьная форма будет рекламироваться как гендерно нейтральная, **даже с упаковкой**, если ее цвет указывает на разделение полов.*

Смысл этих двух переводов получился противоположный, но в обоих случаях противоречащий оригиналу. В переводе Google Translate получается, что школьная форма будет рекламироваться и продаваться как гендерно нейтральная, **несмотря на возможное удаление упаковки**, а в переводе DeepL — что это будет происходить, **несмотря на возможное наличие упаковки**. Ни та, ни другая программа не соотнесла удаление упаковки со способом продвижения товара как гендерно нейтрального. И дело здесь не в неспособности машины воспринять контекст. Дело опять в неспособности выявить внутритекстовые связи.

Во многих случаях ошибки в восприятии машиной средств когезии бывают не столь очевидными и действительно создают возможность неверного понимания смысла.

Оригинал:

For years, people searched for traces of the frogs. Scientists ran extensive surveys, and pastors offered rewards to their congregants for anyone that could find one.

Google Translate:

В течение многих лет люди искали следы лягушек. Ученые проводили обширные опросы, а пасторы предлагали своим прихожанам вознаграждение за каждого, кто мог их найти.

Ошибки в машинном переводе многообразны. Помимо уже обсуждавшихся, можно назвать нестабильность в передаче одних и тех же имен собственных, использование разных лексических соответствий для обозначения одного и того же предмета, варьирование при переводе на русский язык местоимений мужского и женского рода по отношению к одному и тому же человеку и многое другое. Все это нарушает когерентность текста. Но главное не перечень конкретных лексических или грамматических ошибок, а понимание основных причин, лежащих в их основе: неспособность программ-переводчиков воспринимать текст как единое структурное, смысловое и коммуникативное целое и неспособность воспринимать сам текст как существующий в определенном ситуативном контексте.

Рассматривая роль контекста для полноценного понимания и воспроизведения текста при переводе, О. Б. Лукманова пишет, что переводчик должен извлекать из текста «дополнительные смыслы, возникающие в результате многообразных связей между различными внутритекстовыми и внетекстовыми элементами», которые образуют внутритекстовые и межтекстовые ассоциативно-смысловые цепочки (Лукманова 2023: 284). Совершенно очевидно, что компьютерный переводчик этого делать не может. Не случайно, говоря о задачах в развитии машинного перевода, О. И. Кузьмин выделяет смысловую связность текста как один из главных критериев оценки качества переведенного текста (Кузьмин 2021: 41).

4. Заключение

В основе большинства ошибок, допускаемых программами компьютерного перевода, лежит неспособность машины воспринимать текст как единое структурно-семантическое и коммуникативное целое, все компоненты которого объединены многочисленными и многообразными внутритекстовыми связями. При несомненном улучшении качества машинного перевода в текстах продолжают встречаться ошибки, многие из которых вызваны тем, что машина воспринимает и переводит не текст, а последовательность существующих как бы вне контекста слов, словосочетаний и грамматических форм. Понятие контекста для

программы зачастую ограничивается предметной областью текста, а когерентность и когезия в некоторых случаях не выходят за рамки абзаца и могут нарушаться даже в пределах одного предложения. В результате переведенный текст не всегда является текстом в собственном, терминологическом смысле этого слова. И у этого есть два последствия. Во-первых, неверно установленные и / или неверно воспроизведенные внутритекстовые связи в ряде случаев приводят к содержательным и смысловым ошибкам, которые могут скрываться за внешне правильной формой. Во-вторых, все большее распространение машинного перевода, часто используемого без постредактирования, приводит к постепенному разрушению самого понятия «текст». Занятия по анализу текста со студентами выявляют у них те же проблемы, что и у компьютерных программ: многие не соотносят значения слов и логических коннекторов с контекстом, не видят контекста за пределами абзаца, не воспринимают текст в контексте внеязыковой ситуации, затрудняются в выявлении внутритекстовых связей.

Поскольку нет оснований ожидать, что в ближайшее время проблемы с текстуральностью машинного перевода будут разрешены, необходимо целенаправленно учить студентов восприятию текста, в том числе используя машинный перевод для иллюстрации того, к чему может приводить нарушение основных текстообразующих категорий.

Список литературы / References

- Абдулганеева И. И.* К вопросу о специфике перевода текстов художественного дискурса: возможен ли машинный перевод? // Переводчик 2030: обучение профессионально ориентированному переводу в меняющемся мире. Сборник научных статей международной научно-практической конференции, приуроченной к празднованию Дня преподавателя перевода. Одинцово: МГИМО, 2022. С. 5–10. [Abdulganeeva, Irina I. (2022) K voprosu o specifike perevoda tekstov hudozhestvennogo diskursa: vozmozhen li mashinnyj perevod? (On the Issue of the Specifics of Translation of Literary Discourse Texts: is Machine Translation Possible?). In *Perevodchik 2030: obuchenie professional'no orientirovannomu perevodu v menyayushchemsya mire. Sbornik nauchnyh statej mezhdunarodnoj nauchno-prakticheskoy konferencii, priurochennoj k prazdnovaniyu Dnya prepodavatelya perevoda* (Translator 2030: Training in Professionally Oriented Translation in a Changing World: A collection of scientific articles from the international scientific and practical conference dedicated to the celebration of Translation Teacher's Day). Odintsovo: MGIMO, 5–10. (In Russian)].
- Валгина Н. С.* Теория текста. Москва: Логос, 2003. [Valgina, Nina S. (2003) *Teoriya teksta* (The Theory of the Text). Moscow: Logos. (In Russian)].
- Воронец М. В.* Теория текста. [Электронный ресурс]. Барнаул: АлтГПУ, 2021. URL <https://library.altspu.ru/dc/pdf/voronec1.pdf> (дата обращения: 24.05.2025). [Voronec, Mariya V. (2021) *Teoriya teksta* (The Theory of the Text). Retrieved from: <https://library.altspu.ru/dc/pdf/voronec1.pdf> (2025, May 24). Barnaul: AltGPU. (In Russian)].

- Гальперин И. Р. Текст как объект лингвистического исследования. Москва: КомКнига, 2006. [Galperin, Ilya R. (2006) *Tekst kak ob'ekt lingvisticheskogo issledovaniya* (Text as an Object of Linguistic Research). Moscow: KomKniga. (In Russian)].
- Губайловский В. А. Машина учится делать переводы художественных текстов. Получается совсем неплохо [Электронный ресурс] // Сетевое издание TechInsider. 9.11.2022. URL <https://www.techinsider.ru/technologies/news-1565203-ii-perevodit-literaturnye-proizvedeniya-i-pochti-ne-ustupaet-cheloveku-perevodchiku/> (дата обращения: 22.05.2025). [Gubajlovskij, Vladimir A. (2022) *Mashina uchitsya delat' perevody hudozhestvennyh tekstov. Poluchaetsya sovsem neploho* (The Machine is Learning to Translate Literary Texts. It's Doing Quite Well). *Setevoe izdanie TechInsider*. Retrieved from: <https://www.techinsider.ru/technologies/news-1565203-ii-perevodit-literaturnye-proizvedeniya-i-pochti-ne-ustupaet-cheloveku-perevodchiku/> (2025, May 22). (In Russian)].
- Ивлева А. Ю., Кузнецова Е. А. Личность переводчика как фактор межкультурной коммуникации при переводе публицистических текстов // Вестник Нижегородского государственного лингвистического университета им. Н. А. Добролюбова. 2024. Вып. 4 (68). С. 57–75. [Ivleva, Alina Yu., & Kuznecova, Ekaterina A. (2024) *Lichnost' perevodchika kak faktor mezhkul'turnoj kommunikacii pri perevode publicisticheskikh tekstov* (The Translator's Personality as a Factor of Intercultural Communication in Translating Journalistic Texts). *LUNN Bulletin*, 4 (68), 57–75. (In Russian)]. DOI 10.47388/2072-3490/lunn2024-68-4-57-75.
- Качалова М. К., Лихацкая Н. В. Машинный перевод художественной литературы // Science and Technology Research — 2022. Сборник статей V Международной научно-практической конференции. Петрозаводск: Международный центр научного партнерства «Новая Наука», 2022. С. 230–233. [Kachalova, Marina K., & Lihackaya, Natal'ya V. (2022) *Mashinnyj perevod hudozhestvennoj literatury* (Machine Translation of Fiction). In *Science and Technology Research — 2022. Sbornik statej V Mezhdunarodnoj nauchno-prakticheskoj konferencii* (Science and Technology Research — 2022. Collection of articles of the V International scientific and practical conference). Petrozavodsk: Mezhdunarodnyj centr nauchnogo partnerstva «Novaya Nauka», 230–233. (In Russian)].
- Кожанов А. А., Россихина Г. Н. Лингвистическое понятие текста // Филологические науки. Вопросы теории и практики. 2013. № 9 (27). В 2-х ч. Ч. 1. С. 83–87. [Kozhanov, Aleksandr A., & Rossihina, Galina N. (2013) *Lingvisticheskoe ponyatie teksta* (Linguistic Concept of Text). *Philology. Theory and Practice*, 9 (27), 83–87. (In Russian)].
- Копытов О. Н. О фундаментальных категориях текста // Вестник ИГУ. 2011. № 3 (15) С. 149–157. [Kopytov, Oleg N. (2011) *O fundamental'nyh kategoriyah teksta* (On the Fundamental Categories of the Text). *The Bulletin of Irkutsk State University*, 3 (15), 149–157. (In Russian)].
- Коровкина М. Е., Семёнов А. Л. Инферирование и функциональный подход к тексту: на материале синхронного перевода // Вестник Российского университета дружбы народов. Серия: Теория языка. Семиотика. Семантика. 2022. № 13 (2). С. 337–352. [Korovkina, Marina E., & Semyonov, Arkadij L. (2022) *Inferirovanie i funkcional'nyj podhod k tekstu: na materiale sinhronnogo perevoda* (Inferencing and Functional Approach to Text: Based on Simultaneous Interpreting). *RUDN of Language Studies, Semiotics and Semantics*, 13 (2), 337–352. (In Russian)]. DOI: 10.22363/2313-2299-2022-13-2-337-352.
- Кочеткова Н. С., Ревина Е. В. Особенности машинного перевода // Филологические науки. Вопросы теории и практики. 2017. № 6 (72). В 3-х ч. Ч. 2. С. 106–109. [Kochetkova, Natal'ya S., & Revina, Elena V. (2017) *Osobennosti mashinnogo perevoda* (The Features of Machine Translation). *Philology. Theory and Practice*, 6 (72), 106–109. (In Russian)].
- Кузьмин О. И. Перспективы современных систем машинного перевода // Вестник Нижегородского государственного лингвистического университета им. Н. А. Добролюбова. 2021. Вып. 1 (53). С. 41–52. [Kuz'min, Oleg I. (2021) *Perspektivy sovremennyh sistem*

- mashinnogo perevoda (Modern Machine Translation Systems: Trends and Prospects). *LUNN Bulletin*, 1 (53), 41–52. (In Russian)]. DOI 10.47388/2072-3490/lunn2021-53-1-41-52.
- Литвинова Е. М. Машинный перевод. Москва: НОУ ВПО «МГТА», 2012. [Litvinova, Ekaterina M. (2012) *Mashinnyj perevod* (Machine Translation). Moscow: NOU VPO «MGTA». (In Russian)].
- Лукманова О. Б. Вертикальный контекст в русских переводах детективных романов Дороти Л. Сэйерс // Вестник Нижегородского государственного лингвистического университета им. Н. А. Добролюбова. 2023. Вып. 4 (64). С. 282–303. [Lukmanova, Ol'ga B. (2023) Vertikal'nyj kontekst v russkih perevodah detektivnyh romanov Doroti L. Sejers (The Vertical Context in Russian Translations of Detective Novels by Dorothy L. Sayers). *LUNN Bulletin*, 4 (64), 282–303. (In Russian)]. DOI 10.47388/2072-3490/lunn2023-64-4-282-303.
- Малявина А. Н., Сдобников В. В., Шеронов А. А. Постредактирование машинного перевода: теоретический и дидактический аспекты // Вестник Нижегородского государственного лингвистического университета им. Н. А. Добролюбова. 2025. Вып. 1 (69). С. 84–100 [Malyavina, Anna N., Sdobnikov, Vadim V., & Sheronov, Anton A. (2025) Postredaktirovanie mashinnogo perevoda: teoreticheskij i didakticheskij aspekty (Machine Translation Post-Editing: Theoretical and Didactic Issues). *LUNN Bulletin*, 1 (69), 84–100. (In Russian)]. DOI 10.47388/2072-3490/lunn2025-69-1-84-100.
- Митин П. С. Типичные ошибки машинного перевода общеполитической лексики (на примере перевода прессы арабской республики Египет) // Вестник Московского университета. Серия 22. Теория перевода. 2021. № 1. С. 17–34. [Mitin, Pavel S. (2021) Tipichnye oshibki mashinnogo perevoda obshchepoliticheskoy leksiki (na primere perevoda pressy arabskoj respubliky Egipet) (Typical Errors in Machine Translation of Sociopolitical Vocabulary: a Case Study of Translations of Egypt's Mass Media). *Moscow University Translation Studies Bulletin*, 1, 17–34. (In Russian)].
- Мурзин Л. Н., Штерн А. С. Текст и его восприятие. Свердловск: изд-во Урал. ун-та, 1991. [Murzin, Leonid N., & Shtern, Alla S. (1991) *Tekst i ego vospriyatie* (The Text and its Perception). Sverdlovsk: izd-vo Ural. un-ta. (In Russian)].
- Павлова О. Ю., Рябова М. Ю. Типология ошибок в машинном переводе // Полилингвальность современной культуры: Сборник статей Международной научной конференции в рамках I Международного научно-образовательного форума «Филологическая наука и образование в Кузбассе». Кемерово: КемГУ, 2022. С. 247–258. [Pavlova, Oksana Yu., & Ryabova, Marina Yu. (2022) Tipologiya oshibok v mashinnom perevode (Typology of Errors in Machine Translation). In *Polilingval'nost' sovremennoj kul'tury: Sbornik statej Mezhdunarodnoj nauchnoj konferencii v ramkah I Mezhdunarodnogo nauchno-obrazovatel'nogo foruma «Filologicheskaya nauka i obrazovanie v Kuzbasse»* (Polylinguality of Modern Culture. Collection of Articles of the International Scientific Conference within the Framework of the 1st International Scientific and Educational Forum “Philological Science and Education in Kuzbass”). Kemerovo: KemSU, 247–258. (In Russian)].
- Переходько И. В., Мячин Д. А. Оценка качества компьютерного перевода // Вестник Оренбургского государственного университета. 2017. № 2. С. 92–96. [Perekhod'ko, Irina V., & Myachin, Dmitriy A. (2017) Ocenka kachestva komp'yuternogo perevoda (Evaluation of the Quality of Computer Translation). *OSU Bulletin*, 2, 92–96. (In Russian)].
- Пищальникова В. А., Сорокин Ю. А. Введение в психопэтику. Барнаул: Изд-во АГУ, 1993. [Pishchal'nikova, Vera A., & Sorokin, Yuriy A. (1993) *Vvedenie v psihopoetiku* (Introduction to Psychopoetics). Barnaul: iz-vo AGU. (In Russian)].
- Пищальникова В. А., Сонин А. Г. Общее языкознание. Москва: Р.Валент, 2017. [Pishchal'nikova, Vera A., & Sonin, Aleksandr G. (2017) *Obshchee yazykoznanie* (General Linguistics). Moscow: R.Valent. (In Russian)].

- Ртищева Н. Г. Оценка качества работы систем машинного перевода // Решетневские чтения: Материалы Международной науч.-практ. конф., посвящ. памяти генерального конструктора ракетно-космических систем академика Михаила Федоровича Решетнева. Красноярск: СибГУ им. М. Ф. Решетнева, 2017. Вып. 21 (2). С. 704–705. [Rtishcheva, Natal'ya. G. (2017) *Ocenka kachestva raboty sistem mashinnogo perevoda* (Evaluation of the Quality of Machine Translation Systems). *Reshetnevskie chteniya: Materialy` Mezhdunarodnoj nauch.-prakt. konf., posvyashh. pamyati general'nogo konst-ruktora raketno-kosmicheskix sistem akademika Mixaila Fedorovicha Reshetneva* (Reshetnev readings: Proceedings of the International Scientific and Practical Conference, dedicated to in memory of Academician Mikhail Fedorovich Reshetnev, General Designer of Rocket and Space Systems). Krasnoyarsk: SibGU im. M. F. Reshetneva, 21 (2), 704–705. (In Russian)].
- Солганик Г. Я. К определению понятий «текст» и «медiateкст» // Вестник Московского университета. Серия 10. Журналистика. 2005. № 2. С. 7–15. [Solganik, Grigorij Ya. (2005) *K opredeleniyu ponyatij «tekst» i «mediatekst»* (On the Definition of the Concepts of “text” and “media text”). *Lomonosov Journalism Journal*, 2, 7–15. (In Russian)].
- Фидельская Е. С. Оценка эффективности систем машинного перевода в работе с текстом художественного стиля // Актуальные проблемы лингвистики и лингводидактики. Материалы IV Всероссийской научно-практической конференции. Ульяновск: УлГУ, 2019. С. 83–89. [Fidel'skaya, Ekaterina S. (2019) *Ocenka effektivnosti sistem mashinnogo perevoda v rabote s tekstem hudozhestvennogo stilya* (Evaluation of the Effectiveness of Machine Translation Systems in Working with Artistic Text). In *Aktual'nye problemy lingvistiki i lingvodidaktiki. Materialy IV Vserossijskoj nauchno-prakticheskoy konferencii* (Actual Problems of Linguistics and Linguodidactics. Proceedings of the IV All-Russian Scientific and Practical Conference). Ulyanovsk: UIGU, 83–89. (In Russian)].
- Шамилов Р. М. Что такое потребностноориентированный перевод? // Вестник Нижегородского государственного лингвистического университета им. Н. А. Добролюбова. 2023. Вып. 1 (61). С. 116–135. [Shamilov, Raviddin M. (2023) *Chto takoe potrebnostno-orientirovannyj perevod?* (What Is a Needs-Centered Translation?). *LUNN Bulletin*, 1 (61), 116–135. (In Russian)]. DOI: 10.47388/2072-3490/lunn2023-61-1-116-135.
- De Beaugrande, Robert, & Dressler, Wolfgang U. (1981) *Introduction to text linguistics*. London & New York: Longman. Retrieved from: https://www.academia.edu/31596533/Introduction_to_Text_Linguistics (2025, May 28).
- Thai, Katherine, Karpinska, Marzena, Krishna, Kalpesh, Ray, Bill, Inghilleri, Moira, Wieting, John & Iyyer, Mohit. (2022) *Exploring Document-Level Literary Machine Translation with Parallel Paragraphs from World Literature*. Retrieved from: <https://doi.org/10.48550/arXiv.2210> (2025, May 25).

Словари / Dictionaries

- ЛЭС — Лингвистический энциклопедический словарь / Гл. ред. В. Н. Ярцева. М.: Советская энциклопедия, 1990. 685 с. [Yartseva, Viktoriya N. (ed.) (1990) *Linguisticheskij Entsiklopedicheskij Slovar'* (Linguistic Encyclopaedic Dictionary), Moscow: Sovetskaya entsiklopediya. (In Russian)].