

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ
федеральное государственное бюджетное образовательное учреждение высшего
образования
«Тольяттинский государственный университет»

Гуманитарно-педагогический институт

(наименование института полностью)

Кафедра «Теория и практика перевода»

(наименование)

45.03.02 Лингвистика

(код и наименование направления подготовки / специальности)

Перевод и межкультурная коммуникация

(направленность (профиль) / специализация)

ВЫПУСКНАЯ КВАЛИФИКАЦИОННАЯ РАБОТА (БАКАЛАВРСКАЯ РАБОТА)

на тему Машинный перевод текста технической документации с английского на
русский язык

Обучающийся

Н.Ю. Усачев

(Инициалы Фамилия)

(личная подпись)

Руководитель

к. филол. н., доцент С. М. Вопияшина

(ученая степень (при наличии), ученое звание (при наличии), Инициалы Фамилия)

Тольятти 2025

Аннотация

Актуальность исследования обусловлена возрастанием значения технической документации в условиях глобализированной инженерной и производственной деятельности. С увеличением объёмов информации и необходимостью её оперативного перевода машинный перевод становится востребованным инструментом, однако не всегда обеспечивает требуемую точность, что требует вмешательства переводчика.

Объектом исследования является машинный перевод текстов технической документации.

Предметом исследования выступают особенности применения систем машинного перевода к таким видам технических текстов, как чек-листы и спецификации, а также методы оценки качества полученных переводов.

Целью работы является изучение и описание применения машинного перевода к текстам чек-листов и спецификаций с позиции качества перевода, а также выработка практических рекомендаций по улучшению результатов машинной обработки подобных текстов.

Задачи исследования включают: 1) анализ понятия технической документации и её жанровых особенностей; 2) изучение основных принципов функционирования машинного перевода, его разновидностей и технологий; 3) рассмотрение этапов предредактирования и постредактирования как способов повышения качества перевода; 4) проведение предпереводческого анализа и машинного перевода выбранных текстов технической документации; 5) оценка качества перевода с помощью специализированных метрик и выявление типичных ошибок.

Структура. Работа включает введение, две главы, заключение и список используемую литературы.

В первой главе раскрываются теоретические аспекты: даётся определение технической документации, её классификация, особенности лексики, синтаксиса и грамматики, а также рассматриваются принципы

машинного перевода. Вторая глава содержит практическую часть исследования: проводится анализ и предредактирование технических текстов, выявляются типичные ошибки машинного перевода и оценивается его качество на основе метрики TER и статистики ошибок. В заключении подводятся итоги исследования.

Список используемой литературы включает 38 источника научной литературы.

Общий объем работы составляет 62 страницы.

Оглавление

Введение.....	5
1 Теоретические основы исследования.....	8
1.1 Определение понятия технической документации	8
1.2 Классификация текстов технических документов	13
1.3 Практика использования машинного перевода в текстах технической документации.....	15
2 Стратегия машинного перевода технической документации (на примере чек-листа и спецификации).....	26
2.1 Использование машинного перевода при работе с текстами технической документации.....	26
2.1.1 Предпереводческий анализ чек-листа	26
2.1.2 Предпереводческий анализ спецификаций.....	30
2.2 Предредактирование текста.....	35
2.2.1 Предредактирование текста чек-листа	35
2.2.2 Предредактирование текста спецификации.....	39
2.3 Анализ результата машинного перевода.....	41
2.3.1 Анализ результата машинного перевода чек-листов.....	41
2.3.2 Анализ результата машинного перевода спецификаций.....	44
2.4 Оценка качества машинного перевода	477
Заключение	50
Приложение А Ручной перевод текста чек-листа.....	56
Приложение Б Ручной перевод текста чек-листа	59
Приложение В Ручной перевод текста спецификации.....	61
Приложение Г Статистика ошибок в текстах чек-листов.....	62

Введение

Техническая документация – это важный компонент научно-технического и производственного взаимодействия, обеспечивающий правильную эксплуатацию, сборку, настройку, тестирование и сопровождение продукции. В таких текстах, как чек-листы и спецификации, критически важна точность формулировок и корректность передачи информации. В условиях глобализации и высокой скорости информационного обмена применение машинного перевода становится не только востребованным, но и необходимым инструментом для ускорения рабочих процессов. Однако точность и адекватность перевода машиной не всегда соответствуют требованиям, предъявляемым к техническим текстам.

В связи с этим особенно актуальным становится вопрос использования машинного перевода при переводе текстов технической документации.

Объектом исследования являются машинный перевод текстов технической документации.

Предметом исследования являются особенности применения систем машинного перевода к текстам чек-листов и спецификаций и оценка перевода.

Целью исследования изучение и описание применения машинного перевода к текстам чек-листов и спецификаций с позиций качества перевода.

Задачами данной работы являются:

- 1) проанализировать понятие и жанровые особенности технической документации.
- 2) изучить принципы функционирования машинного перевода, его основные типы и технологии.
- 3) рассмотреть этапы предредактирования и постредактирования как способы повышения качества машинного перевода.
- 4) выполнить предпереводческий анализ и машинный перевод текстов технической документации.

5) оценить качество переведённых текстов с использованием соответствующих метрик и проанализировать типичные ошибки.

Материалом исследования послужили англоязычные тексты чек-листов объемом в 16 122 знаков и спецификаций объёмом в 34 109 знаков.

В работе использовались следующие методы исследования: общенаучные методы анализа и синтеза, описательный метод, метод предпереводческого анализа текста, статистический метод, метод сплошной выборки, метод оценки качества перевода.

Теоретической базой исследования послужили работы по стилистике М. Н. Кожиной, И. Р. Гальперина, И. В. Арнольд А. Л. Пумпянского, Г. Д. Орловой, а также работа А. Т. Булатовой по документоведению и В. М. Лейчика по терминоведению. Также были задействованы исследования на тему машинного перевода Л. Н. Беляевой, В. А. Нуриева и Kenny D.

Практическая значимость исследования состоит в том, что работу можно использовать как рекомендации при работе с переводом технических текстов при помощи машинного перевода.

Основные положения данной работы были апробированы на научно-теоретической конференции «Студенческие дни науки ТГУ» 2025 года в конкурсе докладов секции «Лингвистика и межкультурная коммуникация».

Структура работы. Работа состоит из введения, двух глав, заключения и списка используемой литературы.

В первой главе настоящего исследования подробно раскрывается понятие технической документации. Особое внимание уделяется её структурным особенностям и функциям в различных профессиональных и производственных контекстах. Также в этой части работы представлена классификация технических документов по различным критериям, включая назначение, структуру и область применения. Кроме того, анализируется современный процесс перевода технической документации с использованием средств машинного перевода.

Вторая глава посвящена проведению предпереводческого анализа

конкретных видов технических текстов, таких как чек-листы и технические спецификации. В рамках этого анализа исследуются лексико-синтаксические особенности исходных текстов, а также выполняется предредактирование с целью улучшения качества итогового перевода. Далее подробно рассматриваются типичные ошибки, возникающие в процессе машинного перевода, с их классификацией и объяснением причин возникновения. Отдельным разделом этой главы является оценка качества перевода, где представлены количественные и качественные показатели, а также статистический анализ распределения различных типов ошибок.

В заключительной части работы подводятся итоги проведённого исследования. Делается обобщение полученных результатов, формулируются основные выводы и рекомендации, направленные на повышение эффективности перевода технической документации.

Приведён список используемой литературы, на основании которой строилось теоретическое и практическое содержание работы. В общей сложности в список вошло 38 источников, из которых 28 публикаций на русском языке и 10 – на английском.

1 Теоретические основы исследования

1.1 Определение понятия технической документации

Техническая документация принадлежит к научно-техническому подстилю научного стиля. Рассмотрим данный стиль в целом [14, с. 73].

По определению М. Н. Кожинной: «научный стиль – это научная сфера общения и речевой деятельности, связанная с реализацией науки как формы общественного сознания; отражает теоретическое мышление, выступающее в понятийно-логической форме» [16 с. 289].

Иными словами, стиль научного мышления представляет собой форму внешнего выражения определенной познавательной позиции [23 с. 160].

Научный стиль имеет свои подстили. Всего в научном стиле выделяют следующие пять подстилей [26 с. 10]:

- собственно-научный (монография, статья);
- учебно-научный (учебники);
- научно-популярный (очерк, книга);
- научно-технический (техническая документация);
- научно-информативный (реферат, конспект).

Техническая документация относится к научно-техническому подстилю.

В повседневной жизни научную и техническую документацию могут приравнять друг к другу, однако, помимо сходных черт, существует некоторое различие между ними.

Прежде всего, в документации превалирует использование языка для специальных целей (Language for Specific Purpose). Однако, каждый из текстов ставит разные коммуникативные цели. Научная документация затрагивает в основном статьи по той или иной области знания, учебники, диссертации и так далее. В то же время, техническая документация больше отвечает за разного рода инструкции, мануалы к технике, спецификации, чек-листы, таблицы с данными и тому подобные [29].

Из этого следует, что техническая документация отражает техническую идею, вложенную при создании того или иного текста. Подобное суждение сходится с определением технической документации А. Т. Булатовой: «технические документы – это обобщающее название графических и текстовых документов, отражающих техническую мысль» [8 с. 27].

Таким образом, под техническими документами следует понимать совокупность материалов, предназначенных для фиксации результатов инженерной и проектной деятельности, в том числе в сферах строительства, технологии и конструирования. Эти документы служат средством отображения и передачи технической мысли, сформированной в процессе разработки соответствующих проектов.

Обозначим черты технической документации.

Обратим внимание на то, как реализуются стилеобразующие факторы технической документации, т.е задающие параметры текста (лексические, синтаксические, грамматические): сфера общения, цель общения, которые задают всё остальное (вид речи, форма речи, способ изложения, объективность, точность, ясность, экспрессивность (выразительность), тип информации, стандартизированность, полнота изложения) [11 с. 115].

Цель общения технической документации – передача информации между специалистами и предписание.

Функцией технической документации является выражение мыслей и понятий, появившихся при создании технических (и не только) устройств, изложение определенной информации в формате предписаний, необходимых для правильного функционирования устройств. Или же при их установке или во время проведения ремонтных работ [9 с. 550].

Также отметим, что в текстах научного стиля преобладает когнитивная информация [2 с. 4].

Обозначим стилистические черты данного стиля:

Одной из основных лексических черт технической документации является обильное использование в ней терминологической лексики или же,

как обозначил данную черту И. В. Гальперин – терминологичность [12 с. 424]. Каждая научная дисциплина имеет упорядоченную систему терминов, которая отражает её предмет и методы исследования [19 с. 79].

А. Л. Пумпянский выделил у научно-технической литературы следующий состав, который распадается на три главные группы слов:

– слова общего языка, которые составляют основу любого научного и технического текста;

– общенаучные слова, т.е. слова, характерные для описания научных исследований независимо от области знания;

– термины, на основе которых строится метаязык любой науки [24].

Понятие термина имеет множество определений. В данной работе за основу взято понятие В. М. Лейчика: «лексическая единица определенного языка для специальных целей, обозначающие общее – конкретное или абстрактное – понятие теории определённой специальной области знания или деятельности» [17 с. 32].

Дополнительно, выделим, что тексты технической документации не подразумевают под собой двусмысленности [33]. Все слова, представленные в научной документации являются однозначными, для той или иной отрасли. Данная черта важна, так как двусмысленность трактовки может привести к ошибкам на производстве.

Ещё одной лексической особенностью технической документации является использование сокращений. Если тот или иной термин используется множество раз во всём документе, то может быть принято решение сократить его до аббревиатуры, например: «*internal combustion engine (ICE), quality assurance (QE)*» и так далее. Данный механизм позволяет сжать текст и упростить его восприятие для целевой аудитории.

Отдельно можно указать на «сухость» технической документации. Техническая документация, как и научный стиль в целом, редко допускает использование эмоционально окрашенной лексики [13 с. 102].

Ещё одной лексической чертой конкретно технической документации

можно назвать отсылку на другие документы, который могут быть частью производственного процесса, например: «*For reference, see document RPXXXXX3P00110000*».

Рассмотрим синтаксические особенности данного жанра.

К синтаксическим особенностям относят использование сложносочинённых предложений [32]. При этом сама по себе техническая документация стремится к лаконичности, в связи с чем, в ней также очень часто прослеживается использование простых предложений [6 с. 9-10]. Это способствует лучшему восприятию текста, упрощая его прочтение.

Из-за своей предписывающей функции, очень часто в текстах технической документации можно найти повелительное наклонение. Пример: «*Press the red button. Unplug the device*».

Нередко в текстах технической документации можно встретить использование клише, главным образом именной характер морфологических элементов предложения, полную систему связующих компонентов [15 с. 66]. Данная черта предназначена для стандартизации и упрощения информации.

Перейдем к грамматическим особенностям.

Одной из важных черт научно-технического жанра и научного стиля в общем является нейтральность. Данная черта проявляется, в основном, в неопределенно-личных конструкциях [18 с. 424]. Также, признаком данной черты является использование пассивных конструкций [1, с. 335]. Это позволяет читающему сконцентрироваться на объекте действия, а не на субъекте. Например: «*The equipment must be calibrated before use* или *The file has been successfully uploaded*».

Важно отметить, что активный залог также широко используется. Это упрощает текст и делает его более прямолинейным. Можно выделить следующие примеры: «*The system performs diagnostics automatically*».

Ещё одним пунктом можно выделить использование инфинитивных конструкций. Пример: «*Ensure to check the power supply*».

Далее, подчеркнём обширное употребление модальных глаголов,

которые выражают обязательность или возможность тех или иных действий. Наиболее часты в употреблении глаголы: *must*, *should* и *can*. Примеры: «*The user must verify the settings*»; «*The cable should be connected securely*»; и «*The device can operate in both modes*».

Дополнительно можно выделить преобладание в тексте существительных. Они играют ключевую роль в описании процессов, объектов и характеристик. Пример: «*Power failure may lead to data loss*» и «*The installation of the device requires precision*».

Для описания действий предпочтение отдается отглагольным существительным вместо глаголов, что делает текст формальным и стандартизированным. Для примера: «*Connection establishment is required before data transfer*» и «*Regular maintenance ensures reliability*».

Обратимся к особенностям структуры подобного вида текстов. Необходимо отметить неоднородность данного пункта. В зависимости от документа его структура будет значительно видоизменяться. Тексты инструкций будут иметь структуру, присущую подобного рода текстов. В то же время, чек-листы имеют совершенно другую структуру и оформление. Однако, существуют общие структурные особенности.

Например, очень часто в научно-технической документации представляются таблицы и списки, для упрощения восприятия информации. Помимо этого, в текстах технической документации часто прибегают к структурированию на разделы и подразделы.

Таким образом, можно сделать вывод, что техническая документация, являясь разновидностью научного стиля, обладает рядом характеристик, схожих с другими жанрами внутри этого стиля. Вместе с тем, как особый вид текстов, фиксирующих и передающих техническую мысль, она отличается рядом уникальных стилевых признаков, которые формируют её как самостоятельный жанр. Эти различия проявляются на уровне лексики, синтаксиса, грамматических конструкций, а также в особенностях структурной организации текста.

1.2 Классификация текстов технических документов

Технические документы занимают особое место в профессиональной коммуникации, поскольку служат основой для разработки, производства, эксплуатации и сопровождения технических изделий и процессов. После определения понятия технического документа, рассмотрим классификацию документации.

Г. Д. Орлова выделяет следующие виды технических текстов [22 с. 12]:

- научно-техническая литература, т.е. монографии, сборники и статьи по различным проблемам технических наук;
- учебная литература по техническим наукам (учебники, руководства, справочники);
- техническая и товаросопроводительная документация (паспорта, технические описания, инструкции по эксплуатации и ремонту, основные технические данные и др.; накладные, упаковочные талоны, комплектовка и др.);
- техническая реклама: рекламные объявления, фирменные каталоги, проспекты;
- проектная документация: проекты, расчеты, чертежи;
- патенты.

Существует несколько основных критериев классификации подобного рода технической документации. Всего их четыре:

- назначение документа (например: *разработка, производство, эксплуатация*);
- содержательную структуру (например: *описательные, регламентирующие*);
- целевую аудиторию (например: *для разработчиков, конечных пользователей, специалистов по контролю качества*);
- жанровые особенности (например: *спецификации, инструкции, чек-листы*).

Рассмотрим данные критерии подробнее.

Назначение документа. Этот критерий определяет, какую функцию выполняет документ в рамках производственного или эксплуатационного процесса. Например, документы для разработки служат для передачи информации о проектировании изделия, тогда как эксплуатационные документы обеспечивают пользователей практическими инструкциями.

Если речь заходит о классификации документации по её назначению, то их различают на следующие основные виды [25]:

– конструкторская документация. Описывает конструкцию изделия, его состав, устройство и функциональные особенности. Примеры: *чертежи, спецификации, технические условия*;

– технологическая документация. Содержит информацию о процессах изготовления и обработки изделий. Например: *технологические карты, инструкции по сборке*;

– эксплуатационная документация. Обеспечивает пользователей информацией для безопасного и эффективного использования изделий. Это могут быть руководства по эксплуатации, инструкции пользователя, паспорта изделий;

– программная документация. Применяется для сопровождения программного обеспечения, включающая описания алгоритмов, спецификаций и руководство пользователя;

– отчётная документация. Включает документы, фиксирующие результаты испытаний, проверок, приёмки или сертификации изделий.

Содержательная структура. Данный тип критериев обозначает функцию документа. Например, существуют описательные и регулирующие документы.

Описательные тексты технических документов передают читателю понимание о характеристиках и структуре объекта, его предназначении и свойствах. Они дают читателю понимание, что из себя представляет технический объект и как он функционирует. Одним из самых ярких примеров

подобного типа текстов будут технические паспорта.

В свою очередь регламентирующие документы содержат инструкции, предписания или перечни действий, которые необходимо выполнить для корректного использования того или иного устройства.

Целевая аудитория определяет то, каким образом будет составляться текст. В зависимости от уровня подготовки пользователей изменяются стиль, структура и уровень детализации текста. Для людей, знакомых с темой (разработчиков, различных профессионалов) могут использоваться более технические термины, схемы. Конечный же пользователь получит более простую версию текста, которая будет менее сложна в понимании.

Жанровые особенности, в свою очередь, определяет сам жанр конкретного документа. Существует большое множество жанров, например [10, с. 380]:

- инструкции и мануалы;
- чек-листы;
- спецификации;
- технические описания;
- проектная документация;
- технологические карты;
- отчетные документы.

После рассмотрения классификации всех документов, перейдём к практике использования машинного перевода в текстах технической документации.

1.3 Практика использования машинного перевода в текстах технической документации

Машинный перевод (далее МП) широко используются переводчиками при работе со специальными текстами (научных, научно-технических, информационных). Какое-то время, до создания моделей нейронного

машинного перевода, полагалось, что только подобные специальные тексты могут быть стабильно качественно переведены машиной. Однако сейчас это не является правдой. С приходом нейронных сетей некоторые исследования показывают увеличение производительности перевода на 36%, если измерять количество переводимых слов на единицу времени, если сравнивать с работой без помощи машины [5 с. 23]. Помимо этого, машинный перевод технической документации позволяет справляться с большими объемами информации при переводе, что ещё больше ускоряет процесс.

Под машинным переводом подразумевается процесс перевода текста с исходного языка (ИЯ) на переводящий язык (ПЯ), посредством специального программного обеспечения [20 с. 4].

Машинный перевод перетерпел множество различных итераций, прежде чем достичь своих нынешних высот. Некоторые из данных вариантов переводчика точно используются и по сей день [35]. Рассмотрим версии машинного перевода:

- машинный перевод «по правилам» (1950 – 1980-е годы): использование правил лингвистики, чтобы перевести текст из ИЯ в ПЯ. Показал свою неэффективность из-за различной структуры каждого из языков;
- машинный перевод на основе трансформаций (1980 – 1990-е годы): подход использовал двуязычный словарь и набор правил для переноса смысла из ия в пя. являлся заметным улучшением по сравнению со своим предшественником, но не был лишён ограничений;
- межъязыковой машинный перевод (1990-е годы): этот подход использовал промежуточный язык (интерлингва), который представлял значение исходного языка. он был эффективнее своих предшественников, но требовал значительных человеческих усилий для разработки интерлингвы;
- машинный перевод на основе корпусов (1990-е годы): включал в себя перевод на основе примеров и статистический машинный перевод, а также использование параллельного корпуса для получения новых данных для будущих переводов;

- машинный перевод на основе примеров (1990–2000-е годы): в этом подходе перевод строился на основе сходства между параллельными предложениями на исходном и целевом языках. этот метод был эффективнее предыдущих, но требовал большого объёма данных для обучения программы;
- статистический машинный перевод (2000–2010-е годы): этот подход основывался на статистических моделях, использующих большие объёмы двуязычных данных для перевода текста. это было значительным улучшением по сравнению с более ранними подходами, и он широко применялся в промышленных отраслях;
- нейронный машинный перевод (2010 – по настоящее время): данный подход использует глубинное обучение нейронных сетей, чтобы перевести текст. на данный момент является ведущим методом машинного перевода из-за качественного улучшения по сравнению с остальными методами, указанными ранее;
- гибридный машинный перевод (настоящее время): включает в себя два или более метода машинного перевода, такие, например, как перевод «по правилам», статистический и нейронный машинный перевод. в данном случае каждый тип перевода отвечает за тот или иной аспект перевода, таких как грамматика, лексика и контекст, после чего результаты складываются, получая перевод.

Стратегия машинного перевода немного отличается от того, что используется при ручном переводе.

ISO 11669 делит процесс перевода на 4 стадии, которые проиллюстрированы на рисунке 1[36].

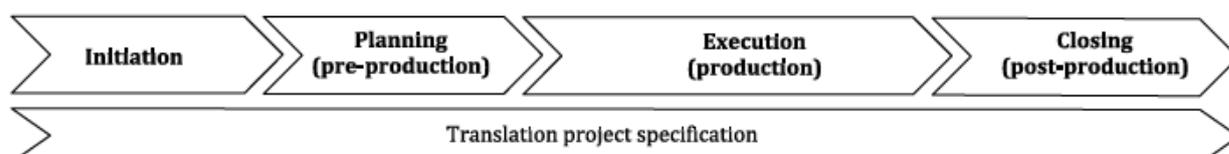


Рисунок 1 – Этапы перевода

Initiation (начальный этап): является подготовительным этапом, на котором заказчик высказывает требования к работе переводчика, а переводчик проводит предпереводческий анализ.

Planning (планирование, предварительная подготовка): на данном этапе текст на ИЯ подготавливается к работе, а также финализируется техническое задание.

Execution (исполнение): на данном этапе осуществляется перевод в соответствии с обговоренным техническим заданием.

Closing (завершающий этап или постпроизводство): переводчик получает обратную связь по результату проделанной работы. Производится оценка перевода в соответствии с выполнением технического задания.

В контексте машинного перевода, меняются исполняемые в каждом пункте действия. Рассмотрим данные отличия.

Начальный этап перевода остается неизменным. Однако, на этапе предварительной подготовки переводчик должен провести предредактуру текста.

Предварительная редактура – это собрание терминологических и стилистических правил, которые подготовят текст для последующей обработки машиной. Это позволяет улучшить результат перевода ещё до того, как программа переведет текст, из-за чего является важным аспектом работы с машинным переводом [30].

Л. Н. Беляева в своей статье «Машинный перевод в работе переводчика: практический аспект», выделяет следующий алгоритм действий предредактуры:

– введение в иноязычный текст артиклей там, где это необходимо или грамматически оправданно. это обусловлено тем, что вставка артиклей вручную помогает машине правильно интерпретировать количество, известность и тип объекта;

– повторение элементов при сочинительной связи словосочетаний в предложении. это делает симметричную структуру более ясной и облегчает машинное определение связи между частями;

– введение союзов при использовании бессоюзной связи между предложениями. это способствует указанию логической связи между частями, помогая системе построить правильную структуру предложения;

– устранение конструкций в скобках в середине именной группы или в середине предложения. данное исправление делает структуру более линейной и однозначной;

– замена окказиональных аббревиатур на полные наименования либо введение специальных символов, предотвращающее их перевод как обычных слов. это способствует переводу машинной тех аббревиатур, которые не распространены или ей неизвестны;

– устранение эллипсисов, неформальных конструкций и метафор, восстановление недостающих элементов и переформулировка неформальных/поэтических конструкций в нейтральный стиль делает текст понятным для алгоритма.

– приведение к единому виду конструкций, которые могут иметь разное написание [4 с. 15]. Это помогает системе распознавать повторяющиеся элементы и использовать правильные словари перевода.

После предредактуры текста, его обрабатывает программа МП. На данный момент самым часто-используемым вариантом машинного перевода является нейронный машинный перевод (НМП).

На данный момент, большинство НМП следует модели, выстроенной на работе Аишиша Васвани и др., в их исследовании «*Attention is all you need*» [31]. Данная модель переводит текст, следуя обозначенным ниже шагам:

– токенизация. Текст разбивается на мелкие элементы, или же токены, преобразовывая символы во что-то, с чем может работать программа. Это могут быть как слова, так и слоги или символы. Например, предложение: «*Я люблю перевод*», превратиться, в восприятии программы «Я», «люблю» и

«перевод». После чего, в зависимости от переводческой модели, может быть дополнительно раздроблена.

– векторизация или создание эмбеддингов. Машина не может работать с текстом напрямую, поэтому она преобразует текст в числа, каждое из которых обозначают значение, грамматику и контекст токена. Пример: «Я» имеет условный вектор $[0.2, -0.1, 0.7]$, «люблю» имеет вектор $[0.8, 0.3, -0.5]$ и так далее. Именно с ними будет работать машина в дальнейшем.

– позиционная кодировка. В зависимости от модели переводчика, к уже указанным векторам могут быть добавлены векторы позиций. Они необходимы, чтобы указать позицию слова в предложении. Однако модели, который задействуют Рекуррентные нейронные сети могут обойти этот шаг.

– прохождение через Encoder (кодировщик). Кодировщик принимает последовательность векторов и строит внутреннее представление предложения. Чтобы учесть контекст предложения задействуется Self-attention (само-внимание) и Multi-Head Attention. С помощью данных инструментов программа понимает, что слово «люблю» относится к слову «перевод», а не наоборот.

– прохождение через Decoder (декодировщик). На данном этапе программа прогнозирует, с какого слова лучше начать предложение, после чего, с помощью функции Softmax прогнозирует каждое последующее слово. Для машины это выглядело бы примерно так:

- I;
- I love;
- I love translation.

Также важно отметить, что в данном процессе также задействованы такие алгоритмы как *Beam Search* и *Sampling*.

– детокенизация. После генерации последовательности токенов происходит обратный процесс – сборка финального текста. Мы получаем результат: I love translation.

По получении финального текста, переводчик может провести пост-

редактуру. Пост-редактура необходима для проверки правильности итогового текста с грамматической, орфографической, пунктуационной и стилистической стороны. В контексте текстов технической документации это также важно, чтобы проверить терминологическое единообразие в тексте.

Постредактура делится на два вида: легкое и полное постредактирование. Легкое постредактирование включает минимальные правки текста и выполняется сравнительно быстро. Полное постредактирование подразумевает исправление всех недочетов работы МП и требует больше времени. Однако в среднем, оба этих метода предполагают куда более быстрое выполнение работы, чем если бы компьютерные средства не были бы задействованы [27].

Международная организация по стандартизации (*International Standards Organization, ISO*) разработала стандарт по постредактированию ISO 18857:2017 (*Translation services – Post-editing of machine translation output – Requirements*). В нем дается следующая характеристика этим двум видам постредактирования: «Легкое постредактирование это процесс, направленный на достижение простого понимания текста, без стремления создать продукт, сравнимый с переводом, выполненным человеком» [37].

В то же время полное постредактирование характеризуется как «процесс направленный на получение продукта, сопоставимого с переводом, выполненным человеком» [37].

Углубимся в особенности обоих уровней постредактирования. Легкое постредактирование предполагает следующее [38]:

- стремление к семантически корректному переводу;
- сохранение языковой нормы языка перевода;
- проверка наличия (или опущения) той или иной информации в процессе перевода машиной;
- редактирование фрагментов текста, которые имеют сниженную лексику (если это не предполагалось текстом изначально, например в художественных целях);

- отсутствие необходимости вносить стилистические правки и исправлять структуру перевода, для улучшения стилистики текста.

Полное постредактирование предполагает следующее:

- стремление к грамматически, синтаксически и семантически правильному переводу;

- адекватный перевод терминологии, в случае если он невозможен, непередаваемые термины включаются в список «непереводимых»;

- проверка наличия (или опущения) той или иной информации в процессе перевода машиной;

- редактирование фрагментов текста, которые имеют сниженную лексику (если это не предполагалось текстом изначально, например, в художественных целях);

- правильность форматирования текста.

В рамках машинного перевода технических документов, из-за важности правильного понимания текста, чаще всего используется именно полная постредактура.

После выполнения всех шагов, указанных выше, необходимо оценить качество работы машинного перевода. В настоящий момент для подобной цели существует ряд метрик, каждая из которых имеет свои преимущества и недостатки. Рассмотрим эти метрики [7 с. 224]:

- BLEU (*Bilingual Evaluation Understudy*) – является одной из самых распространённых метрик оценки перевода. BLEU измеряет совпадение n-грамм между сгенерированным переводом и эталонным переводом, что дает представление о том, насколько хорошо работает модель. Основная идея данной метрики заключается в том, что чем больше совпадений между сгенерированным переводом и эталонным, тем выше соответственно значение. Однако BLEU имеет свои ограничения, такие как недостаточный учет контекста и семантики текста;

- METEOR (*Metric for Evaluation of Translation with Explicit ORdering*) – другая популярная метрика. Её особенностью является то, что в ней

учитывается не только точность и полнота перевода, но также оцениваются и порядок слов и фраз в переводе. Вычисление метрики зависит от precision и recall для слов и фраз в переводе. Затем, вычисляется среднее гармоническое от precision и recall, чтобы дать оценку качества перевода;

– TER (*Translation Error Rate*) – это метрика, оценивающая качество машинного перевода посредством подсчета количества ошибок в переводе. Данная метрика, подсчитывает число редактирований в переводе в сравнении с эталонным текстом [21 с. 105]. Под ошибками в данном случае понимаются: вставки, замены, переупорядочивание слов в переводе, грамматические, синтаксические и лексические ошибки. С помощью данной метрики можно оценить частоту ошибок в машинном переводе, а также является индикатором качества перевода, относительно эталонного текста. Оценка зависит от коэффициента варьируется от общего количества ошибок. Метрика выглядит следующим образом (таблица 1).

Таблица 1 – Интерпретирование результата TER

TER (%)	Качество перевода	Комментарий
0-20%	Очень хорошее	Почти не требует правок
20-35%	Хорошее	Требует мелких корректировок:
35-50%	Среднее	Нуждается в доработке
50-70%	Плохое	Серьезные ошибки
70% и выше	Очень плохое	Почти не соответствует смыслу

Отметим, что для ни одна из выше приведенных метрик не является идеальным показателем качества перевода. Рекомендуется использовать соответствующие метрики, в зависимости от контекста и специфики задачи, для которой и проводится оценка качества.

По результатам анализа теоретического материала первой главы мы можем прийти к следующим выводам:

- техническая документация как жанр принадлежит к научно-техническому подстилю научного стиля и отличается высокой стандартизованностью, точностью и однозначностью выражения.

- коммуникативная цель технической документации заключается в передаче специализированной информации и предписаний между профессионалами, обеспечивая корректное использование и обслуживание технических устройств;

- лексические особенности включают терминологичность, отсутствие экспрессивности, использование сокращений и отсылок на сопутствующие документы, а также преобладание существительных и отглагольных форм;

- синтаксические черты выражаются в использовании простых и сложносочинённых предложений, клише, повелительного наклонения и нейтральных конструкций, включая пассивный залог и инфинитивные формы;

- грамматические особенности технической документации ориентированы на объективность и безличность, что достигается через пассивные конструкции, модальные глаголы и нейтральный стиль изложения;

- структура текстов варьируется в зависимости от жанра (инструкции, спецификации, чек-листы и т.д.), однако общими чертами являются наличие таблиц, списков и разделов для повышения читаемости и удобства восприятия;

- классификация технической документации осуществляется по четырем основным критериям: назначение, содержательная структура, целевая аудитория и жанровые особенности;

- машинный перевод технической документации прошёл эволюцию от правил и статистики к нейросетевым моделям, обеспечивая высокую скорость и эффективность перевода при условии предварительной и последующей обработки текста;

– предредактура текста необходима для повышения качества машинного перевода и включает структурные и лексико-грамматические правки, направленные на устранение неоднозначностей и упрощение восприятия машиной;

– постредактура делится на лёгкую и полную, причём для технических текстов предпочтительно использование полной постредактуры из-за высокой значимости точности терминов и структурных элементов;

– оценка качества перевода осуществляется с помощью метрик BLEU, METEOR и TER, каждая из которых имеет свои особенности и ограничения; оптимальный выбор метрики зависит от целей и контекста перевода.

2 Стратегия машинного перевода технической документации (на примере чек-листа и спецификации)

2.1 Использование машинного перевода при работе с текстами технической документации

Практическая часть данной работы включает рассмотрение использования машинного перевода в текстах технической документации. Рассмотрим процесс перевода с использованием машинного перевода в соответствии со стратегией, представленной в пункте 1.3.

2.1.1 Предпереводческий анализ чек-листа

Чек-лист – это перечень пошаговых последовательных действий, которые необходимо воспроизвести, чтобы получить желаемый результат при работе с тем или иным прибором. Отметка уже выполненной работы помогает увидеть прогресс уже произведенной работы. Именно в постепенном выполнении всех необходимых шагов в правильном порядке и заключается смысл чек-листа как документа [28 с. 83].

Обозначим экстралингвистические черты чек-листа.

Автором чек-листа является профессионал в определенной области (либо группа экспертов), действующие от имени организации (например: *Технический писатель, разработчик техдокументации*). Автор обладает компетентностью в сфере, к которой относится чек-лист, и составляет документ в рамках регламентов или стандартов.

Цель текста: обеспечить контроль над выполнением заданий и последовательностью процедур или этапов, минимизировать ошибки, упростить проверку и повысить эффективность работы. За счёт этого чек-лист относится к эксплуатационной документации.

Реципиентами текста могут являться как специалисты в данной области или технический персонал, которые используют чек-лист в своей работе. От

реципиента ожидается соблюдение пунктов чек-листа.

Способом передачи сообщения является тот или иной носитель информации: бумажный, электронный и другие.

Чаще всего чек-листы создаются и распространяются внутри организаций. Реже, они распространяются специализированными агентствами и нормативными организациями (ISO, DIN, ГОСТ).

Поводом для создания текста может послужить:

- необходимость формализовать повторяющиеся процессы;
- требования стандартизации и регламентирования;
- обеспечение качества, безопасности или соответствия нормам.

Коммуникативной целью текста является передача структурированной инструкции для выполнения и/или контроля деятельности исполнителя. Чек-лист не предполагает дискуссию – он призван контролировать соответствие и обеспечивать чёткое исполнение.

К лингвистическим факторам относятся:

Темой текстов чек-листов является контроль, выполнение и проверка тех или иных операций. Она всегда узкоспециализирована и отражает область применения. Область применения пишется как внутри документа, так и в его заголовке (рисунок 2)

**« Ensure the Electrical Electronic Product Process
Feature convergence »**

PROCESS IDENTITY SHEET

Рисунок 2 – Название и область применения чек-листа

Текст содержит перечень конкретных шагов, требований, задач или критериев, которые необходимо проверить или выполнить для корректной работы устройства (рисунок 3)

Process activities			
APPLICABILITY	RTF	RTR	RTK
	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
2 - Ensure the convergence of electronic feature integration RRRRRRRP0011010			
	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
3 - Realize Architecture Form & validate the ISR RPRRR001100011010			
	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
4 - Realize Architecture File EE RRRRRRRP0011010			
	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
5 - Apply the communication network standards RRRRRRRP0011010			
	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
6 - Realize the SFA Architecture, the SDE/SDM RRRRRRRP0011010			

Рисунок 3 – Лист задач

Структура текста может изменяться в зависимости от организации, создавшей чек-лист. Пример оформления на рисунке 4. Зачастую характеризуется высокой степенью формализации и стандартизации.

Périmètre du Processus : (cocher les cases adéquates) ⁽¹⁾	Logique de rattachement : (cocher les cases adéquates) ⁽²⁾	
Technical means	Personnel & competences ⁽³⁾	

Рисунок 4 – Пример оформления чек-листа

К невербальным элементам текста можно отнести:

- таблицы,
- цветовую кодировку (рисунок 5),

- иконки и символы,
- графические элементы (логотипы, рамки, линии).

Vehicle :						
Checked by :						
Date :						
#D1	#D3	#D5	CF platform	OPF	VPC all vehicles	

Рисунок 5 – Пример цветовой кодировки

Синтаксические особенности рассматриваемого текста характеризуются преимущественно упрощённой структурой предложений, что способствует высокой степени однозначности и функциональной направленности изложения. Примером может служить следующее утверждение: «*EE architecture shall respect the state of the art of both companies*» или «*producing the same Network parameter by different ECU on the same vehicle must be prohibited*», демонстрирующая использование нормативно-обязательных формулировок. Характерным является также активное применение императивных конструкций, направленных на регламентацию действий. Это достигается за счёт использования инфинитивных конструкций глагола: «*Ensure the convergence of electronic feature integration*», «*Realize Architecture Form*», «*Apply the communication network standards*». При этом отмечается сознательное избегание вводных слов и конструкций, что обусловлено стремлением к предельной ясности и лаконичности.

Лексический состав текста отличается высокой степенью специализированности. В нем широко представлены термины, характерные для предметной области, такие как *Busload*, *Electronic Architecture*, *Inter System Request*, *Ground Distribution Diagram* и другие. Стилизовыми доминантами являются сжатость и функциональность формулировок, например: «*Apply the*

communication network standards», «*Realize the SFA Architecture*», «*Prepare and animate EiPF*».

Эмоционально окрашенная лексика полностью отсутствует, что соответствует научно-техническому стилю и обеспечивает нейтральный тон изложения.

Особое место занимает использование стандартизированных формулировок, например: «*Companies' Safety team shall define common list*» или «*under agreement at companies' archi/Network level, this rule can have exception*», что подчёркивает регламентирующий характер документа.

Таким образом, тексту присущ нейтральный, объективный тон, исключающий выражение субъективных оценок. Допустимо включение императивных конструкций, служащих целям нормативного предписания и операционной конкретизации требований.

С точки зрения прагматического воздействия, текст оказывает операционное и регламентирующее воздействие, направляя, побуждая и фиксируя действия пользователя.

2.1.2 Предпереводческий анализ спецификаций

Рассмотрим понятие спецификации. Спецификация – это детальное описание или план того, как что-то должно быть сделано или произведено [34].

Спецификации относятся к конструкторской документации и описывают конструкцию, состав и устройство продукта. Прежде, чем переходить к предпереводческому анализу, необходимо уточнить, что сами по себе спецификации бывают шести различных типов. Данные типы представляют из себя [3, с. 5-7]:

- спецификации требований (к продукту);
- функциональная спецификация;
- спецификация дизайна продукта;
- рабочая спецификация;

- спецификации, которые представляют собой документы с приказами;
- спецификации, как технические стандарты.

Это уточнение является важным, так как в зависимости от типа спецификации, может изменяться и её лексическое, грамматическое, структурное наполнение.

Охарактеризуем экстралингвистические особенности спецификаций:

Автором спецификации, как правило, выступает инженер, разработчик, технический писатель или междисциплинарная команда специалистов, обладающая профессиональными знаниями в соответствующей области. Часто спецификация создается в рамках проектной документации или требований к продукции.

Интенция автора спецификации – зафиксировать технические характеристики, параметры, требования, условия или свойства объекта (товара, системы, программного обеспечения, строительного элемента и т. д.). Автор стремится обеспечить однозначную интерпретацию параметров и предотвратить ошибки в производстве, разработке или эксплуатации.

Реципиентами текста являются: инженеры и разработчики, заказчики и подрядчики, инспекторы, переводчики.

Передача сообщения осуществляется через бумажные или электронные носители информации.

Текст создаётся преимущественно для использования компании или организации-разработчика.

Поводом для создания документа обычно является:

- разработка нового продукта или системы;
- документация уже существующего продукта;
- соблюдение стандартов (ISO, DIN, ГОСТ).

Коммуникационная цель текста: зафиксировать параметры, ограничения, стандарты и условия проекта. Спецификация служит официальным источником информации для всех участников проекта.

Перейдем к лингвистическим факторам:

Темой текста спецификации является описание характеристик, требований, функций стандартов, условий работы и допустимых отклонений проекта. Тематика спецификации зависит от отрасли. В данной работе проводится анализ дизайнерской спецификации электронной архитектуры (рисунок 6).

<u>DESIGN SPECIFICATION</u> (RNDS) Vehicle generic specifications Design specification RNDS-B-00000 v0.0	
<u>STANDARD</u>	

Рисунок 6 – Спецификация дизайна

Спецификация представляет собой нормативно-технический документ, содержащий систематизированное изложение требований к изделию, системе или процессу. В ее содержании формулируются технические, функциональные и эксплуатационные требования, а также указываются ключевые параметры проекта, включая габаритные размеры, массу и иные количественные характеристики. Особое внимание уделяется условиям эксплуатации, в которых предполагается функционирование объекта.

Кроме того, в спецификации приводятся графики (рисунок 7), формулы, таблицы (рисунок 8) и иные аналитические материалы, обеспечивающие точность представления данных.

Разработка и оценка документа осуществляется на основании определённых стандартов, которые также фиксируются в тексте.

Структурная организация спецификации варьируется в зависимости от

применяемого нормативного подхода, однако в большинстве случаев включает титульный лист, введение, основную часть (разделённую на логически обоснованные разделы и подразделы), приложения, глоссарий терминов и перечень нормативных документов, регламентирующих или дополняющих рассматриваемую область (рисунок 9).

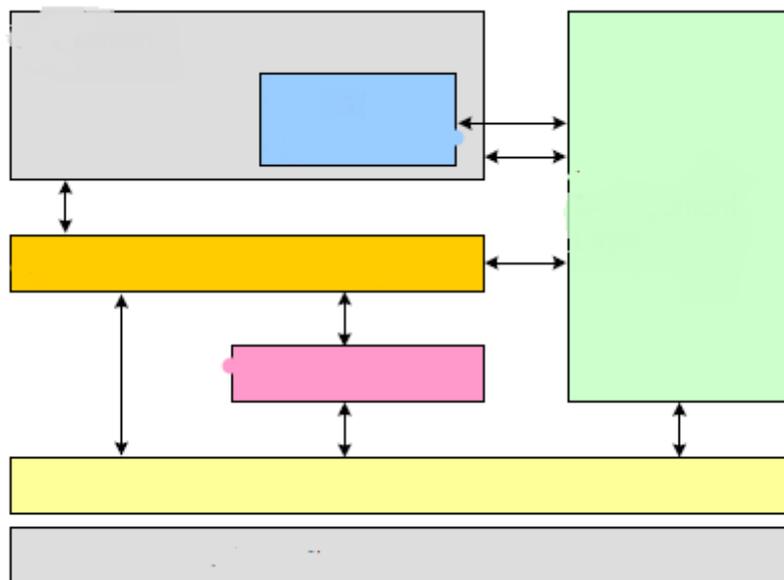


Рисунок 7 - Пример схемы в спецификации

Byte1	Byte2	...	Byte14	Byte15	Byte16
Data	Data	...	Data	Data	Data

Рисунок 8 - Пример таблицы в спецификации

CONTENTS

	Page
FOREWORD	4
INTRODUCTION	6
1. SCOPE	7
2. NORMATIVE REFERNCES	7
3. TERMS AND DEFINITION	8
4. SYMBOLS AND ABBREVIATED TERMS	8

Рисунок 9 – Пример оглавления спецификации

Значимую роль в информационной насыщенности документа играют невербальные компоненты: таблицы служат основным инструментом представления параметров; диаграммы, графики, чертежи и схемы используются как средства визуализации технической информации; формулы и математические обозначения выполняют функцию формализованного описания зависимостей и расчётов.

Синтаксическая организация текста спецификации характеризуется преобладанием простых конструкций, таких как «*For the Bus fault accounting specification, refer to §8.3.1.*» и сложносочинённых конструкций «*If ECU manages diagnostic frame, ECU shall recover within 100ms*», широким использованием страдательного залога, а также модальных и условных выражений, таких как *shall, must, if* и другие, что придаёт тексту нормативно-обязательный характер, например: «*Network parameter by different ECU on the same vehicle must be prohibited*», «*The Remote frame shall NOT be used*».

Лексическая структура текста отличается высокой степенью терминологической насыщенности, включающей специализированную лексику такую как *busload, frame, Network Manager, Identifier, Header, Anti Replay counter, Electronic Control Unit*; аббревиатуры, к примеру: *ID, I-P layer, EEPROM, RAM ECU, EE, ASIL, MAC/AR, CAN_FD, CAN_HS*; числовые значения и коды: *CANSW-FD056a, CANSW-FD060a, CANSW001a*; а также формулы: *Link Layer(V=x1), Link Layer (V=x2), >dt*. Преобладает однозначная лексика, что способствует минимизации интерпретационных рисков и обеспечивает строгость научно-технического изложения.

Преобладает официально-деловая и нейтральная тональность текста. В некоторых случаях задействуются императивно-регламентирующий тон, например: «*The Remote frame shall NOT be used*»; «*Data Length for data frame shall be 64 bytes maximum*». Субъективность, эмоциональность и оценочность в тексте не допускаются.

Прагматическое воздействие текста сводится к тому, чтобы

сформировать основу для проектирования, производства, проверки и сертификации продукта. Она воздействует как нормативный или контрактный документ, фиксирующий обязанности сторон. В юридическом контексте может служить доказательством соответствия или нарушения требований.

2.2 Предредактирование текста

2.2.1 Предредактирование текста чек-листа

Прежде чем обрабатывать текст с помощью программы для перевода, необходимо провести предредактуру. Данный этап позволит увеличить качество конечного результата. Приведенные примеры призваны продемонстрировать, как конечный результат подобной трансформации, так и разницу, между изначальным текстом и текстом предредактированным.

Рассмотрим результаты предредактуры чек-листов и спецификаций, из переведенных документов, общим объёмом в 18 тысяч знаков. Ниже, мы проанализируем ряд примеров, сравнив оригинальный текст с текстом после предпереводческого анализа, выделив отличия между ними, и почему были совершены те или иные изменения.

Оригинал: *Rule: EE architecture shall respect the state of the art of both companies. EE architecture solution shall not degrade it and solution has to be discussed with companies' Safety experts if new safety concern is found. Definition of state of the art: It's de-facto standard in industry as result of benchmark OR it's countermeasure for past experienced problem and no other solution can be proposed.*

Предредактированный текст: *The EE architecture shall respect the state of the art of both companies. The EE architecture solution shall not degrade it. The solution has to be discussed with the companies' Safety Experts if a new safety concern is found.*

Definition of the state of the art:

It is the de-facto standard in the industry as a result of a benchmark, OR it is

a countermeasure for a past experienced problem, and no other solution can be proposed.

Поясним проведенные изменения:

Были добавлены артикли «the», «a», «the» в следующих местах: *EE architecture shall respect the state of the art; if a new safety concern is found; the de-facto standard in the industry.*

В оригинальном тексте артикли были опущены, что характерно для технических текстов на английском языке, особенно в черновых или корпоративных документах. Однако для машинного перевода отсутствие артиклей затрудняет интерпретацию. Устойчивое выражение *the state of the art* требует определённого артикля. *A new safety concern* – неопределённый артикль здесь указывает, что речь идёт о каком-либо ещё не определённом инциденте. *The industry* – артиклем подчеркивается что имеется в виду вся отрасль, а не абстрактное понятие.

Также, в данном примере было добавлено повторное упоминание подлежащего: *The EE architecture solution shall not degrade it.* В оригинале вторая часть предложения «*solution has to be discussed...*» начинается без явного подлежащего. Это допустимо для человека, но не для машины. Предредактура вводит повтор подлежащего *The solution*, чтобы структура стала формально завершённой и понятной алгоритму.

Сложносочиненное предложение было разделено на два: *The EE architecture solution shall not degrade it. The solution has to be discussed....* Это решение было принято из-за того, что машинный перевод лучше справляется с короткими, логически завершёнными фразами. Длинные предложения с сочинёнными частями без повтора элементов приводят к потере логики или неверной расстановке субъектов/объектов.

Был исправлен эллипсис *it's* и добавлены артикли *a de-facto standard, a past experienced problem*. Сокращение *it's* заменено на полную форму *it is* – это повышает формальность и устраняет потенциальную неоднозначность, так как *it's* может быть интерпретировано как *it has* в некоторых случаях.

Артикли были добавлены чтобы сделать текст грамматически завершённым и понятным для анализа.

Рассмотрим следующий пример:

Оригинал: *Rule: Companies' EE architecture needs to consider maximum ASIL ranking to define EE Architecture. Precondition: Companies' Safety team shall define common list of ECU/EE parts with maximum ASIL ranking. Implementation: -In case of ASIL decomposition, solution has to be compliant with maximum ASIL ranking of each component. -EE architecture solution shall minimize the number of ASIL related parts. -If it needs to define different(new) level of ASIL from original part, it needs to have common agreement on ASIL level.*

Предредактура: *Rule: Companies' EE architecture needs to consider the maximum Automotive Safety Integrity Level (ASIL) ranking to define the EE architecture. Precondition: Companies' Safety Team shall define a common list of Electronic Control Unit (ECU) / EE parts with the maximum ASIL ranking. Implementation: -In case of ASIL decomposition, the solution has to be compliant with the maximum ASIL ranking of each component. -The EE architecture solution shall minimize the number of ASIL-related parts. -If it needs to define a different (new) level of ASIL from the original part, it needs to have a common agreement on the ASIL level.*

Были проведены следующие изменения:

Аббревиатуры были заменены на полные наименования при первом упоминании, чтобы исключить ошибочное или произвольное толкование при переводе. Многие системы машинного перевода могут не распознать контекст или неверно интерпретировать аббревиатуру как слово. Указание полной формы даёт системе точную семантическую опору и улучшает терминологическую точность.

Повтор слова *the solution* вместо опущенного подлежащего позволяет обособить логические единицы предложения и устранить двусмысленность. Без такого повторения машинный переводчик может ошибочно "привязать" часть предложения к неверному субъекту.

Был добавлен дефис для формирования корректного сложного прилагательного. Это делает семантику более чёткой: «части, связанные с ASIL», а не «части, ASIL и связанные» (возможное ошибочное толкование при отсутствии дефиса). Это важно для обеспечения грамматической корректности и устранения неоднозначности.

Рассмотрим ещё один пример:

Оригинал: «*Rule: HMI parameter must be functional command (deflate tire indication) and not refers to a specific HMI solution (Blinking, Standstill ON-OFF light). Note: Under agreement at companies' EE archi/Network level, this rule can have exception (example : MILLamp, FlashingIndicatorStatusDisplay already agreed in CIA)*».

Предредактура: «*Rule: An HMI parameter must be a functional command, for example, deflate tire indication, and must not refer to a specific HMI solution for example, blinking, standstill ON-OFF light. Note: Under agreement at the companies' EE architecture / Network level, this rule can have exceptions, for example, MILLamp, FlashingIndicatorStatusDisplay, which are already agreed in CIA*».

Объясним внесенные изменения:

Были устранены скобки внутри предложения («*for example, deflate tire indication*»; «*for example, blinking, standstill ON-OFF light*»). Они были заменены на вводные обороты. Это формализует структуру текста и делает смысл вводимых примеров явным. Данное изменение необходимо, из-за того, что машинный перевод часто неправильно обрабатывает скобки внутри предложения, особенно если в них вложены термины или нестандартные конструкции.

Был исправлен некорректный синтаксис и продублирован модальный глагол. В оригинале нарушено согласование: *must be* сочетается с инфинитивом *refer*, но используется форма *refers*. Мы исправили грамматику и сделали конструкции параллельными. Это помогает системе точно интерпретировать каждое предикативное требование как отдельное,

равнозначное условие.

К более мелким исправлениям можно отнести расшифровку *EE archi* как *EE architecture*. Полное слово делает фразу однозначной и технически точной.

Также было добавлено придаточное предложение. «*Already agreed in CIA*» является неполной конструкцией. В предредактура было восстановлено полное придаточное предложение. Это делает фразу законченной и понятной для перевода.

2.2.2 Предредактирование текста спецификации

Аналогично предыдущему параграфу, ниже будут представлены примеры оригинального и предредактированного текста спецификации.

Оригинал: «*Activation/deactivation of the communication process upon Application request (communication layers initialization, request to the Data Link Layer to configure CAN or CAN-FD objects, ...)*».

Предредактура: «*Activation or deactivation of the communication process upon an Application request, including the initialization of the communication layers and the request to the Data Link Layer to configure the Controller Area Network (CAN) or Controller Area Network with Flexible Data rate (CAN_FD) objects*».

Отметим отличия между двумя вариантами.

Слэш был заменен на связку *or*, так как знак «/» часто порождает семантическую неоднозначность в машинном переводе. Он может быть интерпретирован как «или», «и», «вместо» или как часть технического обозначения. Замена на *or* устраняет неоднозначность.

Как и в предыдущих примерах, аббревиатуры раскрываются при первом упоминании, чтобы обеспечить точную интерпретацию термина.

Были уточнены логические связи при помощи вводного оборота («*including the initialization*» вместо «*communication layers initialization*»). В оригинале элементы даны в скобках, что делает их статус в предложении

грамматически неясным. В предредактуре вводится явное пояснение *including...*, чем интегрирует эти действия в состав предложения как конкретные примеры или составные части процесса.

Следующий пример. Оригинал: «*The Application shall initialize the system resources used by the Interaction-Presentation layer and start the communication process. The node shall be initialized and ready to send and receive CAN frames within 300ms after power-up or Wire WakeUp. If the node performs WakeUp / Sleep by CAN, the node shall comply with the time interval specified in [REF4]. If Application cannot prepare the valid value to Interaction-Presentation layer from the beginning and the node cannot transmit the valid value within above specified time interval, the specific value is sent*».

Предредактура: «*The Application shall initialize the system resources used by the Interaction-Presentation layer and shall start the communication process. The node shall be initialized and shall be ready to send and receive CAN frames within 300 ms after power-up or after Wire WakeUp. If the node performs WakeUp or Sleep by CAN, the node shall comply with the time interval specified in [REF4]. If the Application cannot prepare the valid value to the Interaction-Presentation layer from the beginning, and if the node cannot transmit the valid value within the above-specified time interval, the specific value is sent*».

Разберем внесенные изменения. Вспомогательный глагол *shall* был повторён для параллельности внутри текста. Данное изменение делает структуру предложения симметричной и грамматически однозначной. Машинный перевод точнее интерпретирует параллельные модальные конструкции, когда глагол повторяется – это предотвращает слияние действий и обеспечивает правильную передачу модальности.

Введение предлога *after* перед *Wire Wakeup* обусловлено необходимостью исключить неопределенность. Таким образом программа понимает, что оба события (*power-up* и *Wire WakeUp*) стоят в одинаковом отношении ко времени.

Разделение сложного предложения и повтор союзов в «*...from the*

beginning and the node cannot transmit...» нужно, чтобы разбить текст на две логичные, завершённые части. Это позволяет улучшить структурную читаемость. Также, это подчёркивает двойное условие, при котором наступает действие *«specific value is sent»*.

2.3 Анализ результата машинного перевода

2.3.1 Анализ результата машинного перевода чек-листов

В данном разделе представлен анализ результатов машинного перевода чек-листов с целью выявления типичных ошибок и недостатков автоматического перевода. Для перевода использовался переводчик Google Translate. Анализ перевода позволяет утверждать, что в машинном переводе часто встречаются буквализмы, смысловые ошибки, неточности и искажения, а также неясности. Ниже рассмотрим конкретные примеры, иллюстрирующие эти проблемы и предложим корректные варианты перевода.

Первый пример:

Правило: «Архитектура EE должна соответствовать современному уровню обеих компаний, а решение архитектуры EE не должно его ухудшать. Решение должно обсуждаться с экспертами по безопасности компаний, если обнаружена новая проблема безопасности.»

Определение современного уровня техники:

Это фактический стандарт в отрасли в результате сравнительного анализа или контрмера для проблемы, возникшей в прошлом, и никакое другое решение не может быть предложено».

Анализ перевода позволяет утверждать, что машинный перевод подвержен буквализмам – дословной передаче устойчивых выражений. Так, фраза *«state of the art of both companies»* переведена как «современный уровень обеих компаний», что является дословной калькой. Это снижает точность и ясность текста. Более уместный перевод – «соответствовать текущему уровню развития технологий обеих компаний».

В следующем предложении наблюдается еще один буквализм – *de-facto standard* переведено как «фактический стандарт в отрасли». В русском языке устойчивым является заимствованный термин «де-факто стандарт», и именно он должен использоваться: «Это стандарт де-факто в отрасли...».

Также встречается калька с английского в выражении «*countermeasure for a problem*», переведенном как «контрмера для проблемы». Термин «контрмера» имеет специфическую окраску, связанную с военной лексикой, что в данном контексте неуместно. Лучше использовать: «меры по устранению проблемы».

Фраза «*no other solution can be proposed*» переведена как «никакое другое решение не может быть предложено» и является неточностью. Это создает ощущение физической невозможности предложить решение, тогда как в оригинале имеется в виду отсутствие альтернатив, признанных приемлемыми. Корректный вариант: «...и не может быть предложено другое (приемлемое) решение».

Рассмотрим следующий пример:

«Правило: Архитектура EE компаний должна учитывать максимальный рейтинг ASIL для определения архитектуры EE.»

Предварительное условие:

Команда по безопасности компании должна определить общий список деталей ECU/EE с максимальным рейтингом ASIL.

-В случае разложения ASIL решение должно соответствовать максимальному рейтингу ASIL каждого компонента.

-Решение архитектуры EE должно минимизировать количество деталей, связанных с ASIL.

-Если необходимо определить другой (новый) уровень ASIL от исходной детали, необходимо иметь общее соглашение по уровню ASIL».

Анализ перевода указывает на наличие, буквализмов искажение, неточность и неясность.

Так, выражение «*ECU/EE parts*» переведено как «детали ECU/EE».

Термин «детали» неточен в профессиональной технической среде. Более правильный вариант – «компоненты» или «блоки ECU/EE».

Далее, *ASIL decomposition* переведено как «разложение ASIL», что является дословным переводом. В профессиональной терминологии используется выражение «декомпозиция ASIL», которое точнее отражает суть процесса: «в случае декомпозиции ASIL...».

Фраза «*a different (new) level of ASIL from the original part*» переведена как «новый уровень ASIL от исходной детали». Здесь нарушается синтаксическая связность и теряется точность смысла. Корректный вариант – «уровень ASIL, отличный от уровня исходного компонента».

Также наблюдается неточность в оригинале: «*to have a common agreement on the ASIL level*» в оригинале стало «необходимо иметь общее соглашение по уровню ASIL» в переводе. В оригинале речь идет о достижении консенсуса между сторонами. Перевод должен уточнять субъект: «необходимо общее согласие между компаниями по уровню ASIL».

Фраза «*shall minimize the number of ASIL-related parts*» переведена как «минимизировать количество деталей, связанных с ASIL». Такой перевод является недостаточно ясным. Более корректный вариант – «минимизировать количество компонентов, подлежащих требованиям ASIL».

Рассмотрим следующий пример:

«Правило: Параметр HMI должен быть функциональной командой, например, индикация спуска шины, и не должен ссылаться на конкретное решение HMI, например, мигание, включение-выключение индикатора остановки. Примечание: По соглашению на уровне архитектуры EE/сети компаний это правило может иметь исключения, например, MILLamp, FlashingIndicatorStatusDisplay, которые уже согласованы».

Анализ перевода показывает, что машинный перевод не всегда точно передает семантические нюансы оригинала. Так, фраза «*standstill ON-OFF light*» переведена как «индикатор остановки», что искажает смысл. Речь идет об индикаторе, сигнализирующем состояние стоянки, а не «остановки» как

действия. Правильный перевод – «индикатор включения/выключения в режиме стоянки».

Выражение *«refer to a specific HMI solution»* переведено как «ссылаться на конкретное решение HMI». Здесь «решение» – калька, создающая путаницу. Более точный вариант – «вариант реализации HMI» или «конкретная реализация HMI».

Фраза *«agreement at the companies' EE architecture / Network level»* переведена как «по соглашению на уровне архитектуры EE/сети компаний». Это создает ложное впечатление, что соглашение достигнуто на уровне архитектуры, а не между компаниями. Это искажение. Лучше: «при согласовании между компаниями на уровне архитектуры EE/сети».

Наконец, конструкция *«which are already agreed in CIA»* переведена как «уже согласованы в CIA», что вызывает вопросы: кем именно согласованы? Следует уточнить субъект согласования. Более ясная формулировка – «уже одобрены в рамках CIA».

2.3.2 Анализ результата машинного перевода спецификаций

Рассмотрим особенности машинного перевода технических спецификаций. Анализ перевода позволяет утверждать, что автоматический перевод часто приводит к кальке технических терминов, смысловым и стилистическим искажениям. Приведем конкретные примеры таких ошибок и предложим корректные варианты перевода.

Первый пример:

«Активация или деактивация процесса связи по запросу приложения, включая инициализацию уровней связи и запрос к канальному уровню для настройки объектов сети CAN или сети контроллеров с гибкой скоростью передачи данных (CAN_FD)».

Фрагмент *«controller area network with flexible data rate (CAN_FD)»* переведен как «сеть контроллеров с гибкой скоростью передачи данных

(CAN_FD)». Это дословная калька, нарушающая принятую терминологию. Анализ показывает, что машинный перевод не учитывает общепринятые сокращения и их значение, что приводит к смысловой ошибке. Корректный перевод: «контроллерная сеть с гибкой скоростью передачи данных (CAN_FD)» или «сеть контроллеров CAN с гибкой скоростью передачи данных».

Следующая ошибка наблюдается в отрывке «*request to the data link layer for configuration of CAN network objects*», переведенном как «запрос к канальному уровню для настройки объектов...». Здесь слово *objects* в контексте CAN-сетей обозначает элементы конфигурации, такие как сообщения, буферы и т. д., а не абстрактные «объекты». Корректнее: «настройки CAN-объектов» или «настройки объектов CAN-сети».

Фраза «*initialization of communication layers*» переведена как «инициализация уровней связи». Такой перевод является неточным. Термин «связь» слишком общий и не отражает специфику сетевых протоколов. Лучше использовать: «инициализация уровней коммуникационного стека» или «инициализация коммуникационных уровней».

Также фраза «*activation or deactivation of the communication process*» переведена как «активация или деактивация процесса связи». Анализ показывает, что машинный перевод сохраняет структуру английского оригинала, но нарушает нормы русского языка. В техническом русском тексте употребляется: «включение или отключение коммуникационного процесса».

Выражение «*communication process by request of the application*» переведено как «процесс связи по запросу приложения», что создает семантическую неясность. Более уместный вариант: «коммуникационный процесс, инициируемый приложением».

Следующий пример:

«Приложение должно инициализировать системные ресурсы, используемые уровнем Interaction-Presentation, и начать процесс связи.»

Узел должен быть инициализирован и должен быть готов отправлять

и получать кадры CAN в течение 300 мс после включения питания или после Wire WakeUp. Если узел выполняет WakeUp или Sleep by CAN, узел должен соблюдать временной интервал, указанный в [REF4].

Если приложение не может подготовить допустимое значение для уровня Interaction-Presentation с самого начала, и если узел не может передать допустимое значение в течение указанного выше временного интервала, должно быть отправлено определенное значение».

В выражении «*used by the Interaction-Presentation layer*» перевод «используемые уровнем Interaction-Presentation» представляет собой буквализм. В таких случаях наименования архитектурных уровней рекомендуется транслитерировать и грамматически корректно интегрировать в предложение. Возможный вариант: «используемые уровнем *Interaction-Presentation*» или «используемые уровнем «Interaction-Presentation»». Также допустимо добавить пояснение в скобках: «уровнем *Interaction-Presentation* (уровнем взаимодействия и представления)».

Фраза «*performs WakeUp or Sleep by CAN*» переведена как «выполняет WakeUp или Sleep по CAN». Это дословная калька и искажение. Оно не соответствует принятой терминологии. В технической русскоязычной документации употребляются выражения: «пробуждение по CAN» и «переход в спящий режим по CAN», либо «осуществляет *WakeUp/Sleep* через CAN». Корректный вариант: «выполняет пробуждение или переход в спящий режим по CAN».

В предложении «*the specific value is sent*» перевод «должно быть отправлено определенное значение» является неясностью. Слово «определенное» не передаёт сути – это значение, установленное заранее. Корректнее использовать «заранее определённое значение» или «предустановленное значение».

2.4 Оценка качества машинного перевода

Для того, чтобы дать оценку переводу, мы прибегнем к такой метрике, как TER (или Translation Edit Rate). В дополнение, мы также рассмотрим общую статистику по ошибкам, которые встречались в переводе.

Следуя формуле TER мы можем выявить процент в одном примере:

$$TER = \frac{\text{редактирования}}{\text{общее количество слов эталона}} \quad (1)$$

Эталонные переводы чек-листов представлены представлены в приложении А.1 и А.2. Эталонный перевод спецификации представлен в приложении Б.1. Для того, чтобы определить коэффициент во всех примерах задействуем формулу:

$$TER = \frac{\text{сумма всех } TER}{\text{количество примеров}} \quad (2)$$

Рассмотрим средний коэффициент TER (TERavg) для текстов чек-листов:

$$TER_{avg} = \frac{0.29+0.24+0.18+0.23+0.17+0.20+0.19+0.17+0.18+0.17}{10} = 20.2\% \quad (3)$$

Итого, общий коэффициент TER перевода чек-листов составляет около 20%. Данный результат можно считать хорошим. Правки в таком переводе минимальны. Отдельно можно отметить, что ни один из отдельных примеров не превысил 30%.

Однако, как и упоминалось в первой главе исследования, метрики не могут в полной мере показать качество перевода. В случае с TER, не оценивается «тяжесть» ошибки, только количество редактирований. Данное упущение можно заметить, если посмотреть на статистику.

В приложениях В.1 и Г.1 можно найти категории ошибок, совершённых машиной в текстах чек-листов и спецификаций. Всего было допущено: 9 смысловых ошибок (18% от всех ошибок), 10 буквализмов (29%), 7 искажений (20%), 5 неточностей (15%) и 8 неясностей (18%).

Из этих данных можно сделать вывод, что программа перевода склонна ошибаться, когда переводит специализированную лексику. Также, из 10

примеров в шести были найдены смысловые ошибки, что является весомым недостатком.

Перейдем к анализу текстов спецификаций. Эталонные предложения спецификаций представлены в приложении А. Рассмотрим средний коэффициент TER для спецификаций.

$$TER_{avg} = \frac{0.20+0.13+0.14+0.13+0.30+0.16}{6} = 14\% \quad (4)$$

В данном случае можно говорить о меньшем общем количестве изменений в тексте, что означает перевод более приближенный к эталону.

Если обратить внимание на статистику в данном типе текста, можно отметить следующие допущенные ошибки: всего была допущена 1 смысловая ошибка (4%), 8 буквализмов (29%) 7 искажений (26%), 7 неточностей (26%), 4 неясностей (15%).

Из данных можно сделать вывод, что: модель машинного перевода демонстрирует слабости как в точности (буквализм, смысловые ошибки), так и в выразительности (неясность, неточность).

С учетом предоставленных выше фактов, можно заключить, что общий коэффициент ошибок при машинном переводе сравнительно мал. Несмотря на это, полная постредактура всё ещё является необходимым этапом перевода, в связи с возможным наличием в тексте смысловых и лексических ошибок, которые могут нарушить семантическую целостность текста.

По результатам исследования, проведенного во второй главе, можно сделать следующие выводы:

- чек-листы и спецификации как виды технической документации обладают чётко выраженными лингвистическими и экстралингвистическими особенностями, включая нормативность, терминологичность, формализованную структуру и нейтральный стиль;
- предредактирование текста существенно повышает качество перевода за счёт устранения неоднозначностей, раскрытия аббревиатур, нормализации синтаксиса и грамматики;

– типичными ошибками машинного перевода являются буквализмы, лексические и синтаксические искажения, смысловые неточности и ошибки в передаче технических терминов;

– машинный перевод чек-листов показал *ter* около 20%, что говорит о хорошем качестве перевода, требующем лишь умеренной постредактуры;

– перевод спецификаций дал ещё более низкий *ter*, что свидетельствует о лучшей структурной предсказуемости и удобстве текста для машинной обработки;

– статистический анализ ошибок подтвердил необходимость постредактуры, особенно в отношении терминов, логических связей и технических выражений;

– комплексный подход – предпереводческий анализ, предредактирование и последующая оценка качества – обеспечивает наиболее эффективное использование машинного перевода в технической документации;

– постредактирование остаётся обязательным этапом, особенно в критически важных документах, где точность и однозначность формулировок имеют решающее значение.

Таким образом, проведённый в главе анализ продемонстрировал, что эффективное использование машинного перевода в сфере технической документации возможно только при комплексном подходе, включающем предпереводческий анализ, предредактирование и последующую оценку перевода.

Заключение

Техническая документация неотрывно связана с нынешним технологичным веком. Точный и оперативный перевод подобного рода текстов трудоёмок, от того машинный перевод становится важным инструментом в арсенале переводчика, что позволяет ускорить обработку больших объёмов информации при этом, обеспечивая приемлемое качество перевода.

Проведенное исследование позволило всесторонне рассмотреть процесс машинного перевода технической документации, выявить его ключевые особенности, преимущества и ограничения.

В теоретической части работы было определено, что техническая документация относится к научно-техническому подстилю научного стиля и отличается высокой степенью стандартизованности, терминологичности и однозначности выражения. Были проанализированы жанровые и структурные особенности таких видов документации, как чек-листы и спецификации. Особое внимание уделено их синтаксическим, лексическим и грамматическим характеристикам, которые непосредственно влияют на результат машинного перевода.

Было проведено сравнение между научной и технической документацией, что позволило уточнить лингвостилистические характеристики технических текстов как подгруппы научного стиля и на основе этого дано более уточнённое определение термина данной главы.

После были рассмотрены стилеобразующие черты этого жанра, а также его стилистические черты: лексические, грамматические, синтаксические и структурные.

Во втором параграфе данной главы, были рассмотрены виды классификации технической документации Г. Д. Орловой, а также изучены критерии, по которым определяется жанровая принадлежность того или иного технического текста.

В третьем параграфе было дано понятие машинного перевода. Далее, были рассмотрены этапы эволюции систем машинного перевода: от правил и статистических моделей до современных нейронных и гибридных подходов. Было подчеркнуто значение этапов предредактирования и постредактирования для достижения качественного результата, особенно в работе с техническими текстами, где ошибки могут привести к критическим последствиям. В конце параграфа были приведены системы оценки машинного перевода.

Вторая глава включает в себя предпереводческий анализ текстов чек-листов и спецификаций. Далее была проведена предредактура вышеуказанных текстов, с последующей обработкой их при помощи программы машинного перевода. После был проведен анализ полученных данных. Показанные результаты показали, что машинный перевод способен обеспечить высокий уровень точности при условии корректной подготовки исходного текста; основные ошибки перевода касаются терминологии, калькирования, смысловых искажений и буквального переноса англоязычных конструкций; качество перевода спецификаций оказалось выше по сравнению с чек-листами, что подтверждается более низким показателем TER.

Таким образом, было подтверждено, что при соблюдении определённой стратегии – включающей анализ текста, предредактирование, машинный перевод и последующую редактуру – машинный перевод может быть эффективным инструментом в работе с технической документацией. Однако он не может полностью заменить профессионального переводчика, особенно при необходимости точной передачи отраслевой терминологии и сохранения стилистической нейтральности текста.

Список используемой литературы

1. Арнольд И. В. Стилистика. Современный английский язык : учебник для вузов. 5-е изд., испр. и доп. М. : Флинта : Наука, 2002. 384 с.
2. Айларова С.О. Способы передачи когнитивной информации при переводе научно-популярных текстов // Язык и перевод в меняющемся мире. 2021. С. 4-9
3. Баталин С. В., Кохташвили Н. И., Фоменко О. С., Янкина Е. В. Reading and translating specifications / под ред. С. В. Баталина. Волгоград : ВолгГТУ, 2016. 128 с.
4. Беляева Л. Н. Машинный перевод в работе переводчика: практический аспект // Вестник ПНИПУ. Проблемы языкознания и педагогики. 2019. № 2. С. 22-30.
5. Беляева Л. Н. Машинный перевод в современной технологии процесса перевода // Известия Российского государственного педагогического университета им. А.И. Герцена. 2022, № 203. С. 22-30 [Электронный ресурс]. URL: <https://rep.herzen.spb.ru/publication/1549> (дата обращения: 13.05.2025).
6. Богданова Е. Н. Современный научно-технический текст: стиль и грамматика // Инновационные исследования и разработки в области гуманитарных и социально-экономических наук. 2019. С. 9-13
7. Борунов И. А. Метрики оценки перевода NMT-перевод – основные модели, оценка качества // Вестник Самарского университета. История, педагогика, филология. 2024. Т. 30, № 3. 2024. С. 226.
8. Булатова А. Т. Документоведение : учеб. Пособие. М. : Инфра-М, 2005. 183 с.
9. Булгакова И. В., Максимович А. Н. Стилистические и языковые особенности технической документации // Актуальные проблемы деятельности подразделений УИС. 2018. С. 628.

10. Вопияшина С. М. Жанры инженерной коммуникации и их роль в подготовке инженера в международном контексте // Современное инженерное образование: вызовы и перспективы 2023. 432 с.
11. Гаджаева С. Х. Стилеобразующие факторы и языковые особенности научного стиля // Динамика взаимоотношений различных областей науки в современных условиях. 2017. С. 114-117.
12. Гальперин И. Р. Очерки по стилистике английского языка. М. : Изд-во литературы на иностранных языках, 1958. 458 с.
13. Зайцева Н. Н., Нерушева Т. В. Общая характеристика научного стиля речи как разновидности русского литературного языка // Актуальные вопросы лингвистики и лингводидактики в контексте межкультурной коммуникации 2024. С. 961.
14. Занина А. Л., Краснопёрова Ю. В. Лингвистические особенности технической документации как жанра научного стиля // Вестник научного общества студентов, аспирантов и молодых ученых. 2015. №3. С. 72-75
15. Калинина С. С. Особенности научно-технической разновидности научного стиля в русском языке // Вестник науки и о образования, Олимп, 2017. №2. С. 65-69.
16. Кожина М. Н. Стилистика русского языка. М. : Стилистика русского языка, 2008. 289 с.
17. Лейчик В. М. Терминоведение: предмет, методы, структура. 4-е изд. М. : Книжный дом «ЛИБРОКОМ», 2009. 256 с.
18. Мизюкаева Т. Н. Особенности научно-технического стиля // Евразийское научное объединение. 2021. С. 423-425.
19. Мурзабекова А. Р. Дискурсивные характеристики научного стиля // Инновационная наука, 2020. №1-6. С. 78-81.
20. Мурзабекова А. Р., Алиев З. Г. Машинный перевод: системы машинного перевода, их недостатки и преимущества // Инновационное развитие науки и техники, 2020, 81 с.

21. Нуриев В. А., Егорова А. Ю. Методы оценки качества машинного перевода: современное состояние // Информатика и её применения. Том 15. №2. 2021. С. 104-111 с.
22. Орлова Г. Д. Пособие по переводу английской научно-технической литературы : учеб. Пособие. Тула : Изд-во ТулГУ, 2006. 175 с.
23. Пискорская С. Ю. Стиль научного мышления и стиль научного познания // Вестник сибирского государственного аэрокосмического университета им. Академика М.Ф. Решетнева. 2007. №2. С. 160-163.
24. Пумпянский А. Л. Чтение и перевод английской научно-технической литературы: лексика, грамматика, фонетика, упражнения. Мн. : Попурри, 1997. 608 с.
25. Российский сертификационный центр [Электронный ресурс]. URL: <https://rossertcentr.ru/dokumentatsiya> (дата обращения: 13.05.2025).
26. Серебренникова Н. Г. Научный стиль речи: учебное пособие для студентов-бакалавров технических направлений / Авт.-сост. Н.Г. Серебренникова. –Новосибирск: Издательство НГТУ, 2015. – 88 с
27. Хромова А. А., Лукманова Р. Р. Постредактирование англо-русского машинного перевода: проблемы, методы и оптимизация // Сравнительно-сопоставительные исследования. Том 17. №3. 2024. [Электронный ресурс]. URL: <https://philology-journal.ru/article/phil20240138/fulltext> (дата обращения: 13.05.2025).
28. Ярыгина З. А., Ярыгин А. Н. Чек-лист – новая технология в школьном образовании? // Педагогическое мастерство и современные педагогические технологии. 2019. 232 с.
29. Akthem A. A. Style of technical language: oil and gas discourse [Электронный ресурс]. URL: <https://www.researchgate.net/publication/385094355> (дата обращения: 13.05.2025).

30. Ana G. A. Pre-editing and post-editing [Электронный ресурс]. URL: https://www.researchgate.net/publication/338310852_Pre-editing_and_post-editing (дата обращения: 13.05.2025).
31. Ashish V., Noam S., Niki P., Jakob U., Llion J., Aidan N. G., Lukasz K., Illia P. Attention Is All You Need [Электронный ресурс]. URL: <https://arxiv.org/abs/1706.03762> (дата обращения: 13.05.2025).
32. Ailian Y. An analysis of stylistic features of English for science and technology based on functional linguistics [Электронный ресурс]. URL: <http://www.forestchemicalsreview.com/index.php/JFCR/article/view/1116> (дата обращения: 13.05.2025).
33. Blanka F. K. Scientific prose style and its specifics [Электронный ресурс]. URL: <https://www.researchgate.net/publication/275537459> (дата обращения: 13.05.2025).
34. Cambridge free English dictionary [Электронный ресурс]. URL: <https://dictionary.cambridge.org/dictionary/english/specification> (дата обращения: 13.05.2025).
35. Hanımnur M., Yaşar A., Mehmet C. O. The Evolution of Machine Translation: A Review Study [Электронный ресурс]. URL: https://www.researchgate.net/publication/381916795_The_Evolution_of_Machine_Translation_A_Review_Study (дата обращения: 13.05.2025).
36. ISO 11669:2024 Translation projects – General guidance, Part 4: General – May, 2024
37. ISO 18857:2017 Translation services – Post-editing of machine translation output – Requirements, Part 6: Requirements of full post-editing, April 2017.
38. Kenny D. Machine Translation for Everyone: Empowering Users in the Age of Artificial Intelligence. Berlin : Language Science Press, 2022.

Приложение А

Ручной перевод текста чек-листа

Таблица А.1 – Ручной перевод текста чек-листа

Оригинал	Эталонный перевод
<p>Rule: EE architecture shall respect the state of the art of both companies. EE architecture solution shall not degrade it and solution has to be discussed with companies' Safety experts if new safety concern is found. Definition of state of the art: It's de-facto standard in industry as result of benchmark OR it's countermeasure for past experienced problem and no other solution can be proposed.</p>	<p>Правило: Архитектура EE должна соответствовать передовому уровню технологий обеих компаний. Решение по архитектуре EE не должно снижать этот уровень. В случае выявления новой проблемы безопасности решение должно быть обсуждено с экспертами по безопасности компаний. Определение передового уровня технологий: Это либо фактический отраслевой стандарт, сформированный в результате бенчмаркинга, ЛИБО контрмера, разработанная для ранее возникшей проблемы, при этом проблема не может быть решена другими методами.</p>
<p>Rule: Companies' EE architecture needs to consider maximum ASIL ranking to define EE Architecture. Precondition: Companies' Safety team shall define common list of ECU/EE parts with maximum ASIL ranking. Implementation: -In case of ASIL decomposition, solution has to be compliant with maximum ASIL ranking of each component. -EE architecture solution shall minimize the number of ASIL related parts. -If it needs to define different(new) level of ASIL from original part, it needs to have common agreement on ASIL level.</p>	<p>Правило: Архитектура EE компании должна учитывать максимальный рейтинг ASIL для определения архитектуры EE. Предварительное условие: Команда по безопасности компаний должны определить общий список деталей ECU/EE с максимальным рейтингом ASIL. -В случае разложения ASIL решение должно соответствовать максимальному рейтингу ASIL каждого компонента. -Если необходимо определить другой (новый) уровень ASIL от исходной детали, необходимо иметь общее соглашение по уровню ASIL.</p>
<p>Rule: HMI parameter must be functional command (deflate tire indication) and not refers to a specific HMI solution (Blinking, Standstill ON-OFF light). Note: Under agreement at companies' EE archi/Network level, this rule can have exception (example : MILLamp, FlashingIndicatorStatusDisplay already agreed in C1A).</p>	<p>Параметр HMI должен быть функциональной командой, например, индикация спуска шины, и не должен ссылаться на конкретное решение HMI, например: мигание, включение-выключение индикатора остановки. По соглашению на уровне архитектуры EE / сетевой архитектуры компаний, это правило может иметь исключения (например, MILLamp, FlashingIndicatorStatusDisplay, которые уже согласованы в C1A).</p>

Продолжение приложения А

Продолжение приложения А.1

Оригинал	Эталонный перевод
<p>Rule: Separate HMI frame(PDU) from general purpose Frame for HMI parameters When possible, create a HMI frame (PDU) in ECU and do not mix HMI and functional parameters in the same frame.</p> <p>Implementation: Bus load impact has to be checked when it applies new frame. In case of critical bus load impact, it could be exempt from this rule.</p>	<p>Правило: Отделите фрейм HMI (PDU) от фрейма общего назначения для параметров HMI. При возможности, создайте фрейм HMI (PDU) в ECU и не совмещайте HMI и функциональные параметры в одном фрейме. Параметр HMI должен быть функциональной командой, например, индикация спуска шины, и не должен ссылаться на конкретное решение HMI (например, мигание, включение-выключение индикатора остановки). Влияние на загрузку шины должно быть проверено при применении нового фрейма. В случае критического влияния на загрузку шины может быть сделано исключение из этого правила.</p>

Таблица А.2 – Ручной перевод текста чек-листа

Оригинал	Эталонный перевод
<p>2 - Ensure the convergence of electronic feature integration RPP000000000 Write the EFLG (Electronic Feature Generic List) from the PD (Product Definition), deploy and verify the application of features in systems.</p>	<p>2 - Обеспечить согласованность интеграции электронных функций RPP000000000 Составьте общий список электронных функций (EFLG) на основе определения продукта (PD), разверните функции в системах и проверьте их применение.</p>
<p>3 - Realize Architecture Form & validate the ISR RPP000000000 Write the Architecture Forms for each EE feature declared in series or option in the EFLG from the System Design Document (SDD) and Inter System Request (ISR).</p>	<p>3 – Реализовать архитектурную форму и подтвердить ISR RPP000000000 Составьте AF для каждой функции EE, указанной как серийная или опциональная в EFLG, на основе SDD и ISR.</p>
<p>6 - Realize the SFA Architecture, the SDE/SDM RPP000000000 Collection Electrical Interface Sheet (EIS), SFA Function (SFA-f) and SF-IS requirement to realize Architecture SFA (SFA-a), Electric Distribution Diagram (SDE), Ground Distribution Diagram (SDM).</p>	<p>6 – Реализовать архитектуру SFA, SDE/SDM RPP000000000 Соберите требования к электрическому интерфейсу (EIS), функциональной прикладной схеме (SFA-f) и функциональной безопасности межсистемных требований (SF-IS) для создания функциональной прикладной схемы архитектуры (SFA-a), схемы электрического распределения (SDE) и схемы распределения заземления (SDM).</p>

Продолжение приложения А

Продолжение приложения А.2

Оригинал	Эталонный перевод
7 - Prepare and animate EiPF RPP000000000 Ensuring the convergence and progress of the EiPF tests, depending on the availability and maturity of electrical equipment is Operating Folder (ECU).	7 – Подготовить и организовать EiPF RPP000000000 Обеспечьте согласованность и прогресс испытаний EiPF в зависимости от доступности и зрелости электрического оборудования, как указано в эксплуатационной папке (ECU).

Приложение Б

Ручной перевод текста спецификации

Таблица Б.1 – Ручной перевод текста спецификации

Оригинал	Эталонный перевод
<p>Activation/deactivation of the communication process upon Application request (communication layers initialization, request to the Data Link Layer to configure CAN or CAN-FD objects, ...).</p>	<p>Активация или деактивация процесса передачи данных по запросу приложения, включая инициализацию уровней коммуникации и запрос к канальному уровню на настройку объектов сети CAN или CAN-FD.</p>
<p>Most of the CAN or CAN-FD data link layer is in the CAN or CAN-FD controller hardware. Only the bit timing settings and frame acceptance filtering (frame type and identifier) are configurable. If the filtering capability embedded in the CAN or CAN-FD hardware is limited, a software filter add-on may have to be developed.</p>	<p>Большая часть канального уровня сети CAN или CAN-FD реализована на аппаратном уровне контроллера. Настройке подлежат только параметры синхронизации битов и фильтрация принимаемых кадров (включая тип кадра и идентификатор). Если встроенные возможности фильтрации аппаратного обеспечения CAN или CAN-FD ограничены, может потребоваться реализация дополнительной программной фильтрации.</p>
<p>The Data Link layer shall update all received data and consumed messages when receiving data frames corresponding to the CAN/CAN-FD acceptance filter. Note: For frame acceptance filter configuration, refer to the ECU specification provided separately with each system. Note: All received data may have the DLC value which is different from the DLC value of the received frame defined in [REF8].</p>	<p>Канальный уровень должен обновлять все полученные и потреблённые сообщения при получении кадров, соответствующих фильтру приема CAN или CAN-FD. Примечание: информацию по конфигурации фильтра приема кадров см. в спецификации электронного блока управления (ECU), предоставляемой отдельно для каждой системы. Примечание: полученные данные могут иметь значение DLC, отличающееся от значения DLC принятого кадра, определенного в [REF8].</p>
<p>The Interaction-Presentation layer shall update all received and consumed messages of the Contained I-PDU when receiving contained I-PDU Header corresponding to the software filter. Note: For contained I-PDU Header software filter configuration, refer to the ECU specification provided separately with each system.</p>	<p>Уровень взаимодействия и представления должен обновлять все полученные и использованные сообщения вложенного I-PDU при получении его заголовка, соответствующего программному фильтру. Примечание: для настройки программного фильтра заголовка вложенного I-PDU см. спецификацию ECU, предоставляемую отдельно с каждой системой.</p>

Продолжение приложения Б.1

Продолжение таблицы Б.1

Оригинал	Эталонный перевод
<p>The Application shall initialize the system resources used by the Interaction-Presentation layer and start the communication process.</p> <p>The node shall be initialized and ready to send and receive CAN frames within 300ms after power-up or Wire WakeUp. If the node performs WakeUp / Sleep by CAN, the node shall comply with the time interval specified in [REF4]. If Application cannot prepare the valid value to Interaction-Presentation layer from the beginning and the node cannot transmit the valid value within above specified time interval, the specific value is sent.</p>	<p>Приложение должно инициализировать системные ресурсы, используемые уровнем взаимодействия и представления, и запустить процесс обмена данными. Узел должен быть инициализирован и готов к передаче и приёму кадров CAN в течение 300 мс после включения питания или активации по проводу Wire WakeUp. Если узел выполняет пробуждение или переход в сон по CAN, он должен соблюдать временной интервал, указанный в [REF4]. Если приложение не может сразу подготовить корректное значение для уровня взаимодействия и представления, и если узел не может передать его в установленные сроки, должно быть отправлено специальное значение.</p>
<p>In case of container type frame PDU to be transmitted, ECU shall calculate Overall size of all contained I-PDUs as shown in Figure 7-4 and Figure 7-5 corresponding to the total of all payloads of the contained I-PDUs plus the total length of the corresponding I-PDU headers.</p>	<p>При передаче PDU контейнерного типа ECU должен рассчитать общий размер всех вложенных I-PDU, как показано на рисунках 7-4 и 7-5. Этот размер включает суммарный объём всех полезных данных вложенных I-PDU, а также общую длину соответствующих заголовков I-PDU.</p>

Приложение В

Статистика ошибок в текстах чек-листов

Рисунок В.1 – статистика ошибок в текстах чек-листов



Приложение Г

Статистика ошибок в текстах спецификации

Рисунок Г.1 – статистика ошибок в текстах спецификаций

