

Санкт-Петербургский государственный университет

**ЛИНЬ Ин**

**Выпускная квалификационная работа**

**Нейросети в машинном переводе: сопоставительный анализ переводов  
иноязычных экономических текстов на русский язык**

Уровень образования: магистратура

Направление 45.04.02 «Лингвистика»

Основная образовательная программа ВМ. 5763. «Иностранные языки и  
межкультурная коммуникация в сфере бизнеса и менеджмента»

Научный руководитель:  
доцент, Кафедра иностранных языков  
в сфере экономики и права,  
Солнцева Елена Сергеевна

Рецензент:  
доцент, Негосударственное  
образовательное учреждение высшего  
профессионального образования  
«Санкт-Петербургский Гуманитарны  
университет профсоюзов»,  
Девель Людмила Александровна

Санкт-Петербург  
2022

## Оглавление

Введение.....	3
Глава 1 Теоретическая основа машинного перевода.....	7
1.1 История развития машинного перевода.....	7
1.2 Определение и классификация машинного перевода.....	12
1.3 Специфика и трудности при переводе экономического текста.....	16
1.4 Оценка качества машинного перевода.....	17
1.5 Классификация ошибок при машинном переводе.....	20
Выводы по Главе 1.....	26
Глава 2 Сопоставительный анализ переводов английских экономических текстов на русский язык.....	28
2.1 Горизонтальное сравнение.....	32
2.1.1 Количество ошибок в переведенной версии 202104.....	34
2.1.2 Количество ошибок в переведенной версии 202201.....	38
2.1.3 Согласованность перевода и количество изменений на трех платформах.....	42
2.2 Вертикальное сравнение.....	45
2.2.1 Анализ переводов <i>Google</i> .....	46
2.2.2 Анализ переводов <i>Prompt</i> .....	51
2.2.3 Анализ переводов <i>Yandex</i> .....	53
2.3 Анализ результатов перевода с точки зрения экономических текстов.....	55
2.4 Обсуждение взаимосвязи между машинным переводом и человеческим переводом.....	58
Вывод по Главе 2.....	64
Заключение.....	67
Список литературы.....	70
Список источников.....	74

## Введение

С широким распространением Интернета, ускорением глобализации и все более частым обменом информацией между странами растет потребность людей в переводе, и традиционный ручной перевод далек от того, чтобы удовлетворить быстро растущий спрос на перевод. Способность машин обрабатывать большие объемы данных делает возможным эффективным и быстрый перевод. Внимание людей все чаще обращается к машинному переводу. Машинный перевод открыл новый этап развития, и в соответствии с требованиями времени появилось различное программное обеспечение для машинного перевода.

В последние годы, с быстрым развитием глубокого обучения, был достигнут большой прогресс в исследованиях машинного перевода нейронных сетей (neural machine translation, NMT). На многих языках производительность машинного перевода с использованием нейронных сетей была значительно улучшена, намного превосходя производительность традиционного статистического машинного перевода (statistical machine translation, SMT). В настоящее время исследования NMT стали центром внимания ученых и экспертов, и теперь это самое передовое направление развития исследований в области обработки естественного языка, и оно имеет огромное влияние в академических кругах. Этим объясняется **актуальность** данного исследования.

**Целью** исследования стал сопоставительный анализ качества перевода экономических текстов с английского языка на русский разными системами

машинного перевода с использованием нейронных сетей. В **задачи** исследования входит следующее:

- 1) рассмотреть особенности машинного перевода с использованием нейронных сетей;
- 2) определить критерии оценки качества машинного перевода экономических текстов с использованием нейронных сетей;
- 3) по выделенным критериям провести горизонтальное и вертикальное сравнение результатов перевода экономических текстов с английского языка на русский в системах машинного перевода с использованием нейронных сетей *Google Translate*, *Prompt Translator* и *Yandex Translator* в два разных периода времени;
- 4) проанализировать качество перевода и эффективность обучения исследуемых систем машинного перевода с использованием нейронных сетей.

**Объектом** исследования в данной работе являются результаты перевода экономических текстов различными системами машинного перевода, а **предметом** исследования являются ошибки и изменения в результатах перевода.

**Материалом** исследования послужили десять текстов экономической тематики на английском языке, опубликованных на сайте информационного агентства ТАСС, и результаты их перевода, выполненного в трех системах машинного перевода с использованием нейронных сетей, а именно *Google Translate*, *Prompt Translator* и *Yandex Translator* в апреле 2021 года и январе 2022 года, то есть в общей сложности было проанализировано 60 русских переводов

с английского языка, выдаваемых тремя переводческими платформами за два разных периода времени.

**Методологической** основой исследования послужили сплошная выборка, метод классификации, описательный и сравнительно-сопоставительный методы. Материал для исследования отбирался методом сплошной выборки. Далее для анализа качество переводов, полученных с помощью разных систем машинного перевода использовались метод классификации, описательный и сравнительно-сопоставительный методы.

**Научная новизна** исследования состоит в том, что в работе впервые анализируются и сопоставляются ошибки перевода текстов экономической тематики в трех разных переводческих системах на основе нейросетей за два разных периода времени.

**Теоретическая значимость** работы состоит в том, что в работе впервые критерии оценки качества перевода применены к переводу экономических текстов с английского языка на русский за два разных периода времени.

Полученные данные могут использоваться в дальнейшем изучении принципов работы систем машинного перевода с использованием нейронных сетей.

**Практическая значимость** обусловлена тем, что результаты исследования могут применяться при оценке эффективности машинного обучения нейросети, в разработке улучшений переводческих систем с использованием нейронных сетей, а также в курсе обучения письменному переводу с английского языка на русский.

**На защиту выносятся следующие положения:**

После десятилетий развития появились различные виды машинного перевода, и благодаря быстрому росту спроса на перевод машинный перевод широко используется в обществе. Качество перевода нейросетевой системы машинного перевода все еще очень низкое, поэтому машинный перевод без участия человека никогда не заменит человека.

**Структура:** данная работа включает введение, две главы, заключение и список литературы. Во введении раскрываются цель, задачи, значимость и актуальность. В первой главе представляются обзор истории машинного перевода, определение и классификация машинного перевода, специфика и трудности при переводе экономического текста, оценка качества машинного перевода и классификация ошибок при машинном переводе. Во второй главе исследуются результаты машинного перевода экономических текстов и изменения с течением времени.

## **Глава 1 Теоретическая основа машинного перевода**

Слово «перевод» происходит от латинского слова «translatio», что означает «перенос, перемещение, перенесение или передача». Как видно уже из этого, потребность в переводе существовала давно. Первоначальная цель перевода состояла в том, чтобы дать людям возможность преодолеть трудности, вызванные географическими различиями и барьерами для общения в торговле.

По мере углубленного развития глобализации, в современном мире политические, экономические, научные и культурные связи разных стран становятся все более тесными, и язык играет все более важную роль в обмене информацией, что широко признается. Поэтому страны придают большое значение изучению иностранных языков и развитию услуг перевода. В этой связи большое значение в последние годы приобрел именно машинный перевод. Но в мире так много языков, у людей не хватает энергии и времени, чтобы овладеть всеми, спрос людей на услуги перевода продолжительно и значительно растет, так что человеческий перевод больше не может удовлетворить потребности людей, и появился машинный перевод.

### **1.1 История развития машинного перевода**

История машинного перевода восходит к 1930-1940-м годам. После 60 лет взлетов и падений машинный перевод прошел извилистый и долгий путь развития, который научные круги обычно делят на следующие четыре этапа: [Дроздова, 2015]

**Начало (1947-1964):** с начала 1950-х до первой половины 1960-х годов исследования в области машинного перевода показали тенденцию к росту. Соединенные Штаты и бывший Советский Союз, две сверхдержавы в военных, политических и экономических целях, оказали большую финансовую поддержку проектам машинного перевода, а европейские страны в связи с потребностями геополитики и экономики также уделили значительное внимание исследованиям в области машинного перевода. Исследование машинного перевода переживает бум.

**Спад (1964-1975):** в ноябре 1966 года комиссия ALPAC опубликовала отчет *“Language and Machines: Computers in Translation and Linguistics”*, в котором всесторонне отвергалась целесообразность машинного перевода и рекомендовалось прекратить финансовую поддержку проектов машинного перевода. Публикация этого доклада нанесла серьезный удар по бурно развивающемуся машинному переводу, и исследование машинного перевода почти зашло в тупик.

**Восстановление (1975-1989):** после вступления в 1970-е годы, с развитием науки и техники и все более частым обменом научно-технической информацией между странами, языковой барьер между странами стал более серьезным, традиционный человеческий перевод был далек от удовлетворения потребностей в компьютерах для участия в переводческой работе. В то же время, развитие компьютерных наук и лингвистических исследований, особенно значительное совершенствование компьютерных аппаратных технологий и применение искусственного интеллекта в обработке

естественного языка, способствовало восстановлению исследований машинного перевода с технического уровня. Проекты машинного перевода снова начали развиваться, и последовательно были запущены различные практические и экспериментальные системы

**Новый этап (1990-по н.вр.):** с широким применением Интернета, ускорением процесса мировой экономической интеграции и все более частым международным общением традиционный человеческий перевод более неспособен удовлетворить беспрецедентно растущий спрос на перевод. Это открыло новые возможности для развития машинного перевода. [Карасев И. В., Артюшина Е. А., 2011]

Ниже приведены некоторые важные события в развитии машинного перевода:

В 1933 году советский ученый Петр Смирнов Троянский получил патент на свой доклад «машину для подбора и печатания слов при переводе с одного языка на другой». Его изобретение очень простое, включая карты на 4 языках, пишущую машинку и старый пленочный фотоаппарат.

В 1946 году появился первый в мире электронный компьютер.

В 1947 году Американский математик Уоррен Уивер (Warren Weaver) предложил то, что использовать компьютеры для перевода естественного языка. Поэтому датой рождения машинного перевода как исследовательской области считается этот год. Два года спустя концепция машинного перевода была сформулирована Уорреном Уивером. Он написал ряд статей о машинном переводе и опубликовал «Меморандум Уивера» (Weaver-Memorandum).

7 января 1954 года в штаб-квартире IBM в Нью-Йорке был запущен Джорджтаунский эксперимент - компьютер IBM 701 автоматически перевел 60 русских предложений на английский, что стало первой демонстрацией машинного перевода в истории.

В 1964 году, чтобы оценить прогресс исследований в области машинного перевода, Американская академия наук учредила комитет по проблемам автоматической обработки речи (Automatic Language Processing Advisory Committee, именуемый Комитетом ALPAC), который начал двухлетний всесторонний анализ и тестирование.

В 1966 году Американский комитет по проблемам автоматической обработки речи (ALPAC) опубликовал отчет “Language and Machines: Computers in Translation and Linguistics”, который приводит к прекращению государственного финансирования в этой сфере.

В начало 1980-х исследование машинного перевода оживилось благодаря развитию компьютерной техники. В конце 1980-х и начале 1990-х годов системы машинного перевода, включая голосовой перевод, статистический машинный перевод и другие, были выпущены на рынок. В конце 1990-х годов появился Интернет, и автоматизированный перевод и машинный перевод на основе примеров быстро развивались.

Машинный перевод стал незаменимым помощником для профессиональных и непрофессиональных переводчиков и бюро переводов в 21 веке .

В сентябре 2016 года Google выпустила систему перевода на основе нейронных сетей (GNMT). За последние несколько лет мы были свидетелями

серьезных изменений в переводческих компаниях, которые используют технологии искусственного интеллекта, такие как машинное обучение и технологии нейронных сетей, для обеспечения лучшего машинного перевода.

[Машинный перевод от холодной войны до диплёрнинга]

В последнее десятилетие исследования и индустриализация машинного перевода быстро продвигались вперед при сильной поддержке со стороны правительства и промышленности. Правительство Китая включило исследования в области понимания естественного языка, в том числе технологии машинного перевода, в наброски национального средне- и долгосрочного плана развития науки и техники. В 2009 году правительство США включило автоматический перевод в список 10 наиболее влиятельных технологий будущего. План "2049" и проект BOLT, возглавляемые Министерством обороны США, делают акцент на машинном переводе как на высшей точке современных информационных технологий. Интернет-гиганты, такие как Google и Baidu, также используют его в качестве ключевой точки входа в онлайн-сервисы. По сообщениям, Google Translate ежедневно предоставляет услуги более чем 200 миллионам пользователей по всему миру, переводя один миллиард раз в день и обрабатывая эквивалент одного миллиона книг в переводе ежедневно, что превышает объем текстов, которые профессиональные переводчики всего мира могут перевести за год. Машинный перевод имеет важное теоретическое исследование и промышленное применение.

## 1.2 Определение и классификация машинного перевода

Машинный перевод может пониматься в двух смыслах. В узком смысле, машинный перевод является процессом перевода текстов (письменных или устных) с одного естественного языка на другой язык с помощью компьютерной программы полностью или почти полностью. В широком смысле, машинным переводом называется та область научных исследований, которая находится на стыке лингвистики, математики, кибернетики и имеет цель построение систем, реализующих машинный перевод в узком смысле [Лядащева, 2018].

История машинного перевода начинается в 1933 году. С развитием исследования этой сферы в основном существуют разные виды системы машинного перевода. Классификацию систем машинного перевода можно произвести по-разному [Карасев И. В., Артюшина Е. А., 2011].

По степени участия человека в процессе машинного перевода можно разделить на три вида:

- Полностью автоматический перевод;
- Автоматизированный машинный перевод при участии человека;
- Перевод, осуществляемый человеком с использованием компьютера.

В соответствии с различным принципом работы можно выделить следующие основные виды машинного перевода:

- Машинный перевод на основе правил — Rule-based Machine Translation (RBMT);

- Машинный перевод на основе примеров — Example-based Machine Translation (EBMT);
- Машинный перевод на основе статистики — Statistical-based Machine Translation (SMT). [Дроздова, 2015]

Ниже мы кратко представим принципы работы этих трех видов машинного перевода.

### **Машинный перевод на основе правил:**

Идеи машинного перевода на основе правил начали появляться еще в 1970-х годах [Машинный перевод от холодной войны до диплёрнинга]. Система машинного перевода на основе правил состоит из двуязычного словаря и набора лингвистических правил под каждый язык. Эту систему машинного перевода можно разделить на три категории: системы дословного перевода (Direct Machine Translation), трансферные системы (Transfer-based Machine Translation) и интерлингвистические системы (Interlingual Machine Translation). [Машинный перевод от холодной войны до диплёрнинга]

### **Машинный перевод на основе примеров:**

Основной принцип работы машинного перевода на основе примеров - найти фрагмент текста, наиболее похожий на фрагмент исходного текста из корпуса, извлечь их результаты перевода на целевой язык, выполнить соответствующее преобразование и наконец, получить результат перевода полного предложения. Основная идея машинного перевода на основе примеров была впервые предложена Макото Нагао, который предложил, чтобы люди не проводили глубокий анализ языка при переводе простых предложений, а просто

переводили. Сначала исходное предложение разбивается на несколько фрагментов, а затем эти фрагменты переводятся на целевой язык, каждый фрагмент переводится путем сопоставления примера предложения из корпуса по аналогии, и, наконец, эти переведенные фрагменты объединяются в длинное полное предложение. [Машинный перевод от холодной войны до диплёрнинга]

### **Машинный перевод на основе статистики:**

В системах машинного перевода на основе статистики применяется принцип статистического анализа: в программу загружаются большие объемы текстов на исходном языке и их соответственные переводы, которые выполняются человеком. Программа делает анализ статистики межъязыковых соответствий, синтаксических конструкций и другие, затем опирается на эту статистику при выборе вариантов перевода — это именно самообучение.

[Карасев И. В., Артюшина Е. А., 2011]

С 2013 года, когда исследования глубокого обучения достигли больших успехов, постепенно появился **нейронный машинный перевод** - Neural Machine Translation (NMT)[Дьяченко И.Н., Матыченко Ю.В., 2020]. Это самый новый вид машинного перевода, в котором используются искусственные нейронные сети[Павлов, 2017].

Термин «нейронная сеть» (англ. neural network) обычно имеет два значения:

- 1) Биологическая нейронная сеть (Biological Neural Network) — сеть, состоящая из биологических нейронов, которые связаны или функционально объединены в нервной системе. В нейронауках зачастую определяется как группа нейронов, которые выполняют специфические физиологические функции.
- 2)

Искусственная нейронная сеть (Artificial Neural Network) — сеть, состоящая из искусственных нейронов (программируемая конструкция, имитирующая свойства биологических нейронов). Искусственные нейронные сети используются для изучения свойств биологических нейронных сетей, а также для решения задач в сфере искусственного интеллекта. Следует отметить, искусственные нейронные сети построены по принципу организации и функционирования биологических нейронных сетей [Фаустова, 2017].

В данной работе под термином «нейросеть» понимается искусственная нейронная сеть.

Каждая из вышеперечисленных основных систем машинного перевода имеет свои преимущества [Степанов Д.А., Федорцов П.А., 2019], но также неизбежно имеет некоторые недостатки и ограничения. Например, система машинного перевода на основе правил может точно использовать лингвистические особенности и законы, но разработка полных языковых правил является непростой задачей; система машинного перевода на основе примеров может в полной мере использовать существующие результаты перевода, но обслуживание корпуса требует много труда и затрат; машинный перевод на основе статистики может облегчить трудность в приобретении знаний, но чистые математические методы трудно полностью решить сложные проблемы в языке. Для дальнейшего повышения уровня перевода системы машинного перевода [Новожилова, 2014], специалисты и ученые изобрели новые системы машинного перевода, объединив преимущества вышеперечисленных основных систем машинного перевода. **Гибридный**

**машинный перевод** (Hybrid machine translation — НМТ), который интегрирует несколько подходов к машинному переводу в одной системе машинного перевода, является одной из новых систем машинного перевода. ЯНДЕКС ПЕРЕВОДЧИК представляет собой систему гибридно-машинного перевода. Такая система переводит текст или предложение на основе двух подходов - статистического и нейросетевого.

### **1.3 Специфика и трудности при переводе экономического текста**

Как показывает название, «экономический текст представляет собой особый тип текста, предназначенный для фиксации, хранения и передачи экономических знаний, выступающий как вербальный результат предшествовавшей его созданию экономической дискурсивной деятельности.» [Лядацева, 2018]

На самом деле, экономический текст, являющийся научным текстом по экономике, имеет общие особенности с другими текстами научного стиля. Такой текст характеризуется однозначностью, точностью, сжатостью, логичностью и так далее. В таком тексте существует большое количество определений терминологии [Александрова, 2013], которую трудно понять неспециалистам в области экономики. В связи с этим, перевод экономического текста отличается от общелитературного перевода тем, что не требует от переводчика слишком большой субъективной инициативы для проявления эстетической ценности текста. Перевод экономического текста можно рассматривать как лингвистический процесс превращения одного языкового

продукта, связанного с экономикой, в другой языковой продукт при сохранении неизменного значения[Газизов, 2017]. Это требует от переводчика не только передать максимум информации, но и сделать переводимый текст читабельным.

С непрерывным развитием экономической глобализации экономическая деятельность также развивается и меняется, что приводит к изменению экономической терминологии, создавая тем самым новые трудности для перевода.

#### **1.4 Оценка качества машинного перевода**

После появления машинного перевода один сложный вопрос возникает: как определить, хороша или плоха работа подхода к машинному переводу и системы машинного перевода? Чтобы оценить качество машинного перевода, многие ученые и эксперты исследовали по этому поводу: Нанкевич А.А.[2019], Ртищева Н.Г.[2017], Рычихин А.К.[2019], Легостина М.С.[2019], Переходько И.В., Мячин Д.А.[2017] и так далее.

в настоящее время в основном используются три метода:

- Оценка с точки зрения конкретной задачи
- Экспертная оценка
- Автоматические методы [Шаляпина, 1996]

**Оценка с точки зрения конкретной задачи:**

Оценка с точки зрения конкретной задачи представляют собой технологии оценки качества без эталона. То есть, это автоматическая оценка качества перевода без использования референтных текстов.

Автоматическая оценка качества перевода без использования референтных текстов является новым направлением в области оценки качества перевода.[Улиткин, 2016]

### **Экспертная оценка:**

Это метод, осуществляемый переводчиками-экспертами, которые оценивают качество перевода по двум параметрам: адекватность (adequacy)[Мошкович, 2013] и гладкость текста (fluency). Под словом адекватность подразумевается правильная передача смысла оригинала, а гладкость текста демонстрирует соответствие перевода нормам принимающего языка, правильность с точки зрения грамматики [Koehn, 2010]. По этим двум параметрам переводчики-эксперты оценивают качество переведенного текста по шкале от одного до пяти. Но процесс экспертной оценки занимает много времени и стоит дорого, кроме того, субъективность переводчиков-экспертов является главным фактором, который мешает объективной оценке качества перевода.

### **Автоматические методы:**

Автоматические методы представляют собой процессы, в ходе которых требуется сравнить референтный текст и переведенный текст, обработанный системой машинного перевода, и вычислить, насколько эти тексты похожи. Поэтому наличие референтного текста и переведенного текста является

обязательным и необходимым условием для осуществления оценки. В основе автоматической оценки качества перевода с использованием референтных текстов лежит использование различных метрик, которые используются для упрощения и удешевления оценки качества [Улиткин, 2016]. Для автоматической оценки работы машинных переводчиков зачастую используются показатель **Word Error Rate (WER)**, метрики **BLEU** и **NIST**. Эти инструменты позволяют успешно сравнивать работу разных систем машинного перевода и оценивать улучшения в работе конкретной системы [Vilar, 2006]. Используются также метрики точность (precision), полнота (recall) и F-мера [Koehn, 2010].

Рассмотрим подробнее принципы работы каждого из упомянутых выше инструментов оценивания.

**WER (Word Error Rate - пословный коэффициент ошибок)** - метрика, измеряющая расстояние между машинным и образцовым переводом. По сути WER измеряет минимальное количество изменений, которые необходимо сделать, чтобы из результата работы машинного переводчика получить эталонный перевод [Koehn, 2010]. Это общий показатель системы машинного перевода.

Метрика **BLEU (Bilingual Evaluation Understudy - двуязычная оценка дублера)**, которая была предложена разработчиками компании IBM для определения сходства между машинным переводом и эталонным переводом, представляет собой один из самых популярных и недорогих показателей для оценки качества машинного перевода. Алгоритм BLEU сравнивает

последовательные фразы автоматического перевода с последовательными фразами, которые он находит в эталонном переводе, и взвешенно подсчитывает количество совпадений. Для получения достоверного результата для данного инструмента важен правильный порядок слов [Полякова, 2020, с.86]. Метрика BLEU позволяет учитывать не только точность перевода отдельных слов, но и цепочек слов (n-граммы). Основная идея разработчиков метрики состоит в том, что «чем ближе машинный перевод к профессиональному человеческому переводу, тем лучше» [Papineni, K.; Roukos, S.; Ward, T.; Zhu, W. J., 2002].

Метрика NIST (**National Institute of standards and Technology**) является метрикой, которая была разработана на базе метрики BLEU. Если BLEU просто вычисляет точность n-граммов, добавляя равный вес за каждое точное совпадение, NIST вычисляет также, насколько информативен каждый совпадающий n-грамм [Улиткин, 2016, с.176].

Чтобы глубоко рассмотреть достоинства и недостатки трех разных онлайн-переводчиков Google Translate, Promt Translator и Yandex Translator, ниже используется классификация ошибок для оценки качества перевода выбранных экономических текстов.

## **1.5 Классификация ошибок при машинном переводе**

Как уже упоминалось выше, с развитием машинного перевода также развилась оценка качества перевода, в то же время люди начали глубоко изучать и классифицировать ошибки в машинном переводе. Разные ученые и специалисты предлагали разные варианты классификации ошибок, например:

Крашенинникова Н.А.[2011], Гончаров А.А.[2019], Талызина Е.В.[2016] и Антонова О.А.[2020]. До сих пор не существует широко признанного и единого вида классификации ошибок, то есть, когда мы классифицируем ошибки при машинном переводе, можем использовать вариант, которым пользовались исследователи раньше, и также можем создать собственный вариант классификации в соответствии с собственными исследовательскими потребностями и исследовательскими целями. В этой работе мы позаимствуем вариант классификации ошибок, предложенный в «Error Analysis of Statistical Machine Translation Output», с целью классификации ошибок при переводе выбранных нами экономических текстов.

Данная классификация имеет иерархическую структуру и ошибки при машинном переводе делится на пять классов: пропущенные слова (Missing Words), словопорядок (Word order), неправильные слова (Incorrect Words), неизвестные слова (Unknown Words) и пунктуационные ошибки (Punctuation errors). Эти большие классы делятся на меньшие категории как показано рисунке внизу. [Vilar, 2006]

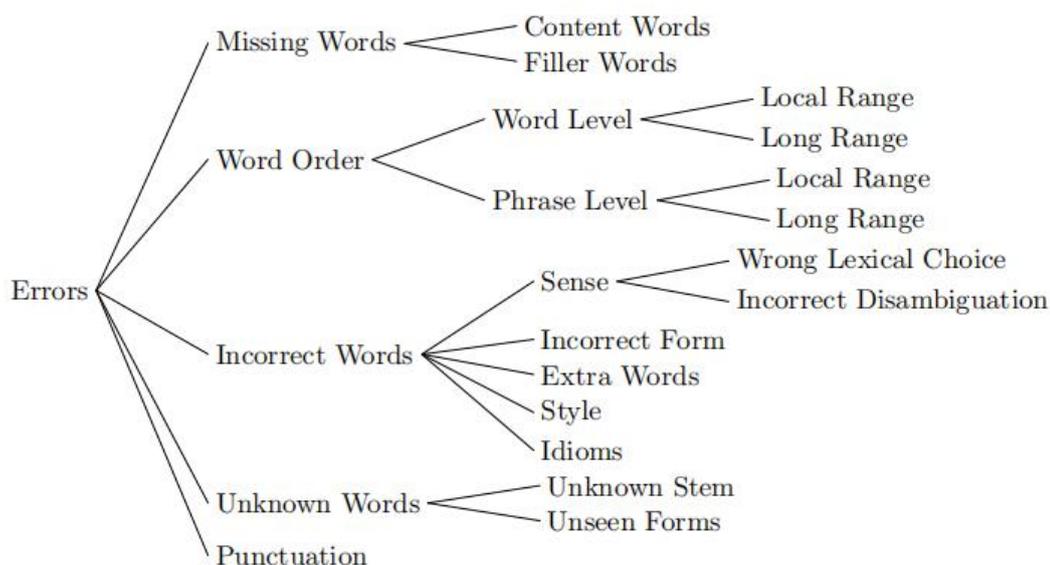


Рис. 1. Классификация ошибок, предложенная Vilar и Llitj'os

### **Пропущенные слова (Missing Words):**

Ошибки «Пропущенные слова» возникают в результате отсутствия слова в переведенном тексте. Эти ошибки делятся на две группы по важности слова:

1. отсутствие главных слов (Content Words) — ошибки, которые возникают когда отсутствие некоторого слова, как существительные и глаголы, приводит к непонятному смыслу предложения. В таком случае пропущенное слово прямо участвует в процессе смыслообразования.

2. отсутствие второстепенных слов (Filler Words) — ошибки, которые возникают когда общий смысл предложения сохраняется при отсутствии слов но перевод оказывается неестественным с точки зрения грамматики.

Второстепенное слово имеют в виду то, что служат для составления грамматически правильного предложения.

### **Словопорядок (Word order):**

Ошибки «Словопорядок» делится на два группы: неправильный порядок на уровне слова и на уровне фразы. И неправильный порядок на каждом уровне выделяется на основе того, насколько необходимо изменить порядок слов или словосочитаний в полученном переводе, чтобы получить правильный и естественный перевод. При этом могут быть два случая: в первом случае необходимость изменения порядка отдельных слов или целых словосочетаний возникает на локальном уровне, то есть, нужно изменить порядок слов или словосочетаний в пределах одного синтаксического отрезка. Во втором случае необходимо перенести слово или словосочетание в другой синтаксический отрезок в том же предложении.

### **Неправильные слова (Incorrect Words):**

«Неправильные слова» является самым обширным классом ошибок при машинном переводе. Данная ошибка допущена когда системе машинного перевода не удастся найти правильный соответственный перевод слова или словосочетания, оказывающегося в оригинальном тексте. Внутри этого класса ошибок выделяются 5 подкатегорий:

1. смысловые ошибки (Sense): смысл предложения оригинала меняется в переведенном тексте из-за неправильного перевод слова или словосочетания. Две причины приводят к неправильному переводу слова или словосочетания — неправильный выбор эквивалента (Wrong Lexical Choice) или неверное снятие неоднозначности (Incorrect Disambiguation ).

2. неправильная грамматическая форма (Incorrect Form). Такая ошибка возникает хотя система машинного перевода выбирает правильное перевод

слова. В флективном языке существует множество изменений форм, посредством которых выражаются отношения между словами. Поэтому правильная грамматическая форма слов в предложении важна для выражения смысла такого предложения. Следует отметить, что русский язык представляет собой флективный язык. В данной работе мы только рассматриваем 5 видов морфологической формы слова: число, падеж, время, наклонение, степени сравнения прилагательных и наречий.

3. лишние слова (Extra Words). Данные ошибки появляются в основном при переводе разговорной речи.

4. стилистические ошибки (Style): используются эквиваленты, которые не соответствуют стилю оригинала. Типичный пример - повторение слова в близком контексте. В этом случае человек-переводчик выбирает синоним и избегает повторения слов.

5. идиоматические выражения, которые система машинного перевода не распознает, переводятся буквально. (Idioms)

### **Неизвестные слова (Unknown Words):**

Ошибки «Неизвестные слова» возникают по причине недостаточного объема корпуса в двух случаях: когда система машинного перевода не идентифицирует слово или основу слова в исходном тексте (Unknown Stem), и не знает определенной формы известных основ слова (Unseen Form). В этих случаях слово не переводится и остается в виде и форме как в оригинале.

### **Пунктуационные ошибки (Punctuation):**

Эти ошибки обычно возникают при работе с языками, не имеющими фиксированных правил пунктуации.

## Выводы по Главе 1

С развитием технологий и необходимостью коммуникации машинный перевод был глубоко интегрирован в нашу жизнь. После десятилетий исследований и разработок появилось множество различных видов машинного перевода. По рабочим принципам машинный перевод можно разделить на разные виды. Среди них в последние годы появилась технология машинного перевода нейронных сетей с использованием глубокого обучения, и качество перевода значительно выросло.

Одновременно с развитием машинного перевода все больше внимания люди уделяют оценке качества перевода, в настоящее время существует примерно три способа оценки качества перевода: оценка с точки зрения конкретной задачи, экспертная оценка, автоматические методы. Автоматическая оценка может быстро определить, является ли качество машинного перевода хорошим или нет. По сравнению с дорогостоящей и трудоемкой экспертной оценкой, автоматическая оценка имеет низкую стоимость и высокую эффективность.

Экономический текст, как некий научный текст, содержит множество терминов и определений. Перевод требует предоставления как можно большего объема оригинальной информации при сохранении объективности и читаемости.

Исходя из представленных выше основных теорий машинного перевода и оценки качества, можно сделать вывод, что исследования в области машинного

перевода продолжают делать прорывные шаги, а теория машинного перевода продолжает развиваться.

## **Глава 2 Сопоставительный анализ переводов английских экономических текстов на русский язык**

На этапе подготовки практической части мы сначала выбрали три популярные платформы онлайн-перевода, а именно Google Translate, Prompt Translator и Yandex Translator. Затем на официальном сайте ТАСС были отобраны десять экономических новостей на английском языке. Потом мы ввели десять текстов на английском языке в три платформы перевода в апреле 2021 года и январе 2022 года соответственно и получили результаты их перевода на русский язык, то есть мы получили в общей сложности 60 русских переводов, выдаваемых тремя переводческими платформами в два разных периода времени.

Мы сравниваем полученные русские переводы по горизонтали и вертикали: **горизонтальное сравнение** означает сравнение результатов перевода, полученных с трех платформ в один и тот же период времени, и выделение результатов перевода, которые несовместимы и которые согласуются друг с другом в трех версиях перевода. Таким образом, мы анализируем степень совпадения результатов машинного перевода экономических текстов с английского языка на русский язык между тремя платформами. В этом процессе мы также обращаем внимание на ошибки в результатах перевода, полученные разными платформами машинного перевода.

Под **вертикальным сравнением** понимается сравнение результатов перевода, полученных на одной и той же платформе в разные периоды времени,

для выяснения изменений в результатах обработки одного и того же английского текста разными платформами перевода за последний год, чтобы проанализировать результаты обучения этих три платформы машинного перевода нейросети. В следующих таблицах 1 и 2 представлена основная информация о трех онлайн-платформах машинного перевода и десяти текстах экономических новостей на английском языке, а также результаты их перевода на русский язык соответственно. Далее мы также используем серию таблиц для наглядного представления результатов анализа.

<b>Platform name</b>	<b>Type</b>	<b>Translation techniques used</b>	<b>Website</b>
PROMT.One	Cross-platform (web application)	Hybrid, rule-based, statistical and neural machine translation	<a href="https://www.online-translator.com/translation">https://www.online-translator.com/translation</a>
Yandex.Translate	Cross-platform (web application)	Statistical and neural machine translation	<a href="https://translate.yandex.com/">https://translate.yandex.com/</a>
Google Translate	Cross-platform (web application)	Statistical and neural machine translation	<a href="https://translate.google.com/">https://translate.google.com/</a>

Таблица 1: Основная информация о трех онлайн-платформах машинного перевода

С развитием машинного перевода сегодняшний машинный перевод больше не ограничивается одной технологией перевода. Согласно Таблице 1 мы можем обнаружить, что выбранные нами переводческие платформы основаны на двух или более технологиях перевода. Promt Translator основан на гибридных, статистических и нейронных технологиях; Яндекс-перевод и Google-перевод основаны на статистических и нейронных технологиях.

Material number	Title	Word count of original	Translation source	version 202104	version 202201
				Word count of translation	
1	IMF improves forecast for US GDP growth rates in 2021 and 2022	189	Google Translate	168	170
			PROMT.One	166	166
			Yandex	171	170
2	IMF praises Russia's measures to fight economic consequences of pandemic	381	Google Translate	326	322
			PROMT.One	326	326
			Yandex	330	329
3	Putin highlights Far East's advanced development as Russia's top priority	239	Google Translate	187	184
			PROMT.One	183	183
			Yandex	185	186
4	Facebook, Twitter, Telegram face \$988,605 fine for prohibited content	183	Google Translate	178	174
			PROMT.One	160	159
			Yandex	177	180
5	Sber, Gazprombank to provide \$2.7 bln loan to Russian Copper Company	181	Google Translate	146	153
			PROMT.One	148	148
			Yandex	150	153
6	Eurasian Economic Commission projects EAEU GDP growth at 3.8% in 2021	200	Google Translate	174	179
			PROMT.One	179	178
			Yandex	176	177
7	Russia ranks second by petroleum exports to US in June 2021 — EIA	168	Google Translate	171	167
			PROMT.One	165	165
			Yandex	172	173
8	Google pays \$47,174 out of \$438,047 of fines in Russia	271	Google Translate	267	245
			PROMT.One	244	244

			Yandex	252	262
9	Shanghai Stock Exchange sets up capital market service base on Hainan	197	Google Translate	160	154
			PROMT.One	161	161
			Yandex	160	159
10	Cross-border money transfers to appear in Faster Payment System in 2022 — Bank of Russia	149	Google Translate	129	128
			PROMT.One	135	135
			Yandex	127	130

Таблица 2: Основная информация о десяти текстах экономических новостей и их переводах

При отборе англоязычных экономических новостей с учетом временных затрат на ручной анализ и качества перевода коротких текстов, обработанных машинным переводом, мы контролировали длину в пределах 300 слов, а десять англоязычных новостных статей насчитывали 2150 слов. Процесс преобразования английского текста в русский текст на различных платформах включает в себя перевод экономических терминов, имен собственных (Например, названий стран, названий организаций, имен людей, названий компаний, названий регионов и т. д.), чисел, процентов и единиц.

Подсчитав количество слов в переводах двух версий каждой платформы и проанализировав изменения количества слов, мы обнаружили, что только количество слов в переводах материалов 1, 2, 3, 5, 7, 8, 9 и 10, которые были переведены Promt Translator, не изменились. Чтобы подсчитать изменения (увеличение или уменьшение) количества слов в десяти переводах, обработанных тремя платформами за два периода времени, мы получаем следующую таблицу. Среди них количество слов в двух версиях, обработанных

Google Translate, составляет 58, в переводе Promt Translator — 2, в переводе Yandex Translator — 25. Хотя количество слов не изменилось, это не означает, что русский перевод не изменился, но мы можем предположить соответственно: Обработка текста Google Translate с течением времени изменилась больше всего, в то время как Promt Translator изменился меньше всего.

<b>Platform name</b>	Google Translate	Prompt Translator	Yandex Translator
<b>The changes in the number of words</b>	58	2	25

Таблица 3: Изменения количества слов на трех платформах

## 2.1 Горизонтальное сравнение

Горизонтальное сравнение относится к сравнению сосуществующих вещей в пространстве. Такое сравнение позволяет людям ясно наблюдать связь между явлениями сходных вещей и понимать их различия и влияния. Прежде всего, мы проводим горизонтальное сравнение результатов перевода, то есть сравниваем результаты перевода, экспортированные десятью английскими исходными текстами на три платформы, и определяем все различия. Все найденные различия вместе отражены в таблицах.

Обнаруженные различия можно разделить на несколько ситуаций. Во-первых, результаты перевода (такие как использование слов, порядок слов, структура предложения и т. д.) между двумя платформами одинаковы, а результаты перевода этих двух платформ отличаются от результатов третьей

платформы перевода. Во-вторых, результаты перевода, экспортируемые тремя платформами, отличаются друг от друга. Путем неполной статистики получается следующая таблица:

	Сходства между Google и Promt	Сходства между Promt и Yandex	Сходства между Google и Yandex	Различия между тремя платформами
Version202104	15	71	45	37
Version202201	61	106	81	42

Таблица 4: Сходства и различия переводов, полученных с трех платформ

Мы обнаружили, что в результатах перевода версии 202104 имеется 71 сходство между переводом *Promt Translator* и переводом *Yandex Translator*, 15 сходств между переводом *Google Translate* и переводом *Promt Translator* и 45 сходства между *переводом Google Translate* и переводом *Yandex Translator*. На трех платформах есть 37 различных переводов. В результатах перевода версии 202201 есть 106 сходств между переводом *Promt Translator* и переводом *Yandex Translator*, 81 сходством между переводом *Google Translate* и переводом *Yandex Translator* и 42 различными переводами между тремя платформами. По этим данным мы обнаружили, что две версии *Promt Translator* имеют больше всего общего с *Yandex Translator*, а *Google Translate* и *Promt Translator* имеют наименьшее. Это может просто показать, что механизм перевода *Promt* и *Yandex* относительно похож, в то время как механизм перевода *Google* и *Promt* меньше всего отличается.

В теоретической части мы упомянули, что существует примерно три метода оценки качества перевода, а именно оценка с точки зрения конкретной задачи, экспертная оценка и автоматические методы. Исходя из характеристик экономических текстов, то есть много имен собственных и выражение точное и строгое, будем оценивать качество перевода по количеству ошибок в переводе и степени согласованности перевода имени собственное.

### 2.1.1 Количество ошибок в переведенной версии 202104

Количество ошибок в переводе оказывает большое влияние на качество перевода, и несомненно, что чем больше ошибок, тем ниже качество перевода. В переведенной версии 202104 мы посчитали ошибки, возникшие на каждой платформе, в разнице, и статистические результаты представлены следующим образом:

Version202104								
Ошибка Google			Ошибка Promt			Ошибка Yandex		
№	Position	Type	№	Position	Type	№	Position	Type
1	M1	Tense	1	M1	Tense	1	M2	Terminology (metaphor) translation error
2	M2	word sense	2	M2	word sense	2	M2	word sense
3	M2	Terminology (metaphor) translation error	3	M2	Terminology (metaphor) translation error	3	M2	Tense
4	M2	word sense	4	M2	word sense	4	M2	Collocation

								problem
5	M2	word sense	5	M2	word sense	5	M5	Grammar
6	M2	Tense	6	M2	word sense	6	M8	word sense
7	M2	Grammar	7	M2	punctuation			
8	M3	Vocabulary repetition	8	M2	Collocation problem			
9	M5	Sentence meaning	9	M3	Sentence meaning			
10	M6	letter case	10	M5	Proper nouns (place names)			
11	M6	Sentence meaning	11	M6	Grammar			
12	M7	Grammar	12	M8	word sense			
13	M8	Punctuation, Translation repeats						
14	M9	Translation repeats						
15	M9	punctuation						
16	M9	preposition						

Таблица 5: Ошибка перевода в версии 202104

Мы обнаружили, что *Google* имеет наиболее очевидные ошибки, достигающие 16, в том числе ошибки перевода экономических терминов, ошибки в значении слов, ошибки в значении предложений, ошибки с заглавными буквами, пунктуационные ошибки, грамматические ошибки, повторный перевод и повторное использование слов, неправильный выбор предлога и неправильное время. В переводе *Prompt* есть двенадцать ошибок, включая ошибки перевода экономических терминов, ошибки значения слов, ошибки перевода географических названий, грамматические ошибки, ошибки

пунктуации, ошибки словосочетания и грамматические ошибки. В переводе Yandex меньше всего ошибок, включая ошибки перевода экономических терминов, ошибки значения слов, ошибки времени, грамматические ошибки и ошибки словосочетания, всего шесть.

Здесь мы проиллюстрируем некоторые из наиболее типичных ошибок в *Google Translate*:

#### 1. Грамматические ошибки

При переводе *“take steps to reduce the regulatory burden”* – примеры выделяются курсивом [*“IMF praises Russia’s measures to fight economic consequences of pandemic”*], результат перевода Google будет *«принятии мер по снизить регуляторную нагрузку»*. То есть использование глагола сразу после «по» — очень очевидная и нелепая ошибка. Чтобы сделать это предложение правильным, мы можем перевести его как *«принятии мер по снижению регуляторной нагрузки»*.

При переводе фразы *“Russia remains second by total petroleum exports to the United States in June 2021”* [ссылка на источник примера снова в квадратной скобке курсивом *“Russia ranks second by petroleum exports to US in June 2021 - EIA”*], результат перевода Google будет *«Россия остается вторым по объему экспорта нефти в США в июне 2021 года»*. Независимо от того, является ли подлежащее слово «страна» или «Россия», слово «второй» должно было быть здесь употреблено в творительном падеже женского рода, и правильный перевод должен быть *«Россия остается второй по объему экспорта нефти в США в июне 2021 года»*.

## 2. Повторение перевода

При переводе “*in the Russian Federation*” результатом перевода Google является «в России. Российская Федерация», «Россия» и «Российская Федерация» используются вместе, чтобы вызвать дублирование, и одно из двух может быть удалено. [*“Google pays \$47,174 out of \$438,047 of fines in Russia”*].

При переводе названия “*Shanghai Stock Exchange sets up capital market service base on Hainan*”, результат перевода Google также имел проблему повторения перевода. Google Translate переводит “*capital market service base*” как «база обслуживания рынков обслуживание рынка стоимости».

Правильным переводом будет «база обслуживания рынков капитала».

## 3. Пунктуационные ошибки

При переводе “*in the Russian Federation*” в шестом абзаце восьмого материала Google выдает результат «в России Российская Федерация», этот результат имеет не только повторение перевода, но и пунктуационные ошибки.

В переводе второго абзаца девятого материала “*Shanghai Stock Exchange sets up capital market service base on Hainan*”, Google переводит “*the provincial branch of Shanghai Stock Exchange*” как «провинциальный филиал Шанхая. Фондовая биржа», между «Шанхая» и «Фондовая биржа» ставится лишняя точка. Правильным результатом должен быть «провинциальный филиал Шанхайской фондовой биржи» или «провинциальный филиал фондовой биржи Шанхая».

#### 4. Ошибка регистра

Также мы обнаружили ошибку регистра в результатах перевода третьего абзаца шестого материала – примеры выделяются курсивом [*“Eurasian Economic Commission projects EAEU GDP growth at 3.8% in 2021”*]. В оригинале говорится, что *“in Armenia and Kyrgyzstan the economic recovery is expected to take more time and complete in 2022”*, а результат Google переводит так: «в Армении и Кыргызстане восстановление экономики ожидается, что это будет больше времени и будет завершено в 2022 году». Игнорируя другие ошибки в переводе этого предложения, мы обнаруживаем, что «Ожидается» не находится в начале предложения, но с заглавной буквы.

#### 2.1.2 Количество ошибок в переведенной версии 202201

Мы также посчитали ошибки на каждой платформе во всех разниц версии 202201. Статистические результаты представлены ниже в таблице 6:

Version202201								
Ошибка Google			Ошибка Promt			Ошибка Yandex		
№	Position	Type	№	Position	Type	№	Position	Type
1	M1	Punctuation, Translation repeats	1	M2	word sense	1	M2	Terminology (metaphor) translation error

2	M2	Terminology (metaphor) translation error	2	M2	Terminology (metaphor) translation error	2	M2	Punctuation
3	M2	Tense	3	M2	Grammar	3	M8	word sense
4	M2	Grammar	4	M2	word sense	4	M9	Grammar
5	M6	letter case	5	M2	Punctuation			
6	M6	letter case	6	M3	Sentence meaning			
7	M6	Translation repeats	7	M5	Grammar			
8	M8	word sense	8	M6	Grammar			
9	M8	Punctuation, Translation repeats	9	M8	word sense			
10	M9	Grammar	10	M9	Grammar			
			11	M10	Proper nouns			

Таблица 6: Ошибка перевода в версии 202201

Из таблицы видно, что по сравнению с версией 202104 количество ошибок трех платформ в версии 202201 уменьшилось в разной степени. Ошибки *Google Translate* сильно уменьшились, всего 10. Сейчас самый неправильный перевод — *Prompt Translator*, всего 11. Ошибки *Yandex* пока меньше всего, всего 4.

Среди них четыре ошибки в переводе *Prompt*, которые заслуживают отдельного исследования:

1. В пятом абзаце материала 2 “*IMF praises Russia’s measures to fight economic consequences of pandemic*”, есть фраза, что “*Directors called on the authorities to remain vigilant*”. Результат перевода *Prompt*: «директоры призвали

власти сохранять бдительность». Здесь подлежащее — «*directors*», сказуемое — «*called in*», и правильная Форма множественного числа слов «*директор*» должно быть «*директора*» вместо «*директоры*».

2. При переводе “*This is a challenge for all and very interesting work*” в четвертом абзаце третьего материала “*Putin highlights Far East’s advanced development as Russia’s top priority*”, результат перевода *Prompt* – «*это задача для всей и очень интересной работы*». На самом деле, изначальный смысл в том, что это вызов для всех, и это тоже очень интересная работа, вызов и работа – это две противопоставленные составляющие. Результат перевода *Prompt* не соответствует смыслу исходного предложения. Правильный перевод должен быть «*это вызов для всех и очень интересная работа*».

3. Название материала 5 – “*Sber, Gazprombank to provide \$2.7 bln loan to Russian Copper Company*”. *Prompt* переводит это название на «*Сбер и Газпромбанк предоставят Русскую медную компанию кредит на \$2,7 млрд*». Первоначальный смысл этого предложения заключается в том, что Сбер и Газпромбанк предоставили Русской медной компании кредит в размере 2,7 млрд долл. Согласно правилам управления глаголов русского языка, «*предоставить*» употребляется с винительным (кого?/что?) и дательным (кому?/чему?) падежами соответственно, то есть «*Русская медная компания*» в этой фразе должна употребляться в дательном падеже вместо винительного, и правильный перевод был бы «*Сбер и Газпромбанк предоставят Русской медной компании кредит на \$2,7 млрд*».

4. В следующей статье «*Eurasian Economic Commission projects EAEU GDP growth at 3.8% in 2021*» есть предложение о том, что “*The economies of Kazakhstan and Russia are supposed to rise by around 4% by the end of this year and may reach the pre-crisis levels as early as in the second half of 2021*”.

Результаты перевода *Prompt* «*Экономика Казахстана и России должна вырасти примерно на 4% к концу текущего года и может выйти на докризисный уровень уже во второй половине 2021 года*». Однако я думаю, что это предложение означает, что и экономика Казахстана, и экономика России растут примерно на 4%, подлежащее стоит во множественном числе, поэтому в этом случае должно быть множественное число подлежащего «экономике» вместо «экономика». И следует использовать «должны» вместо «должна», а «мочь» должна быть в третьем лице множественного числа - «могут».

Ошибка 1 и ошибка 3 являются очень типичными грамматическими ошибками. Перевод *Prompt* не следует грамматическим правилам в процессе перевода, в то время как ошибка 2 и ошибка 4 связаны с тем, что машинный перевод не может правильно понять исходный смысл предложения, что приводит к неправильной интерпретации в процессе перевода. Смысл исходного предложения действительно является существенной разницей между машинным переводом и человеческим переводом: люди-переводчики могут судить о значении предложений на основе содержания всей статьи и в сочетании с тем, что они узнали в повседневной жизни, в то время как машинный перевод может только основываться на оригинальном тексте и языковых правилах. Так что результат машинного перевода будет очень

жестким. При разработке машинного перевода были предприняты усилия для решения этой проблемы, поэтому появились нейронные сети и глубокое обучение, но мы обнаружили, что эта проблема все еще существует в машинном онлайн-переводе, который мы обычно используем до сих пор.

### 2.1.3 Согласованность перевода и количество изменений на трех платформах

Помимо количества ошибок, на качество переводов влияет согласованность перевода, т. е. имеет ли перевод одного и того же существительного один и тот же контекст в одном и том же переводе, чтобы изучить степень согласованности между переводами трех платформ, мы выбрали имя нарицательное “*report*” с высокой частотой в материале 2 и имя собственное “*Faster Payment System*” с высокой частотой в материале 10 – “*Shanghai Stock Exchange sets up capital market service base on Hainan*” и нашли все соответствующие переводы.

Слово “*report*” появляется четыре раза в Материале 2 – “*IMF praises Russia’s measures to fight economic consequences of pandemic*”, и эти четыре “*report*” относятся к “*the report the IMF released on Tuesday*” :

Para1	This is according to the <b>report</b> the IMF released on Tuesday.
Para2	....., according to the <b>report</b> .
Para4	....., according to the <b>report</b> .
Para10	....., according to the <b>report</b> .

Таблица 7: Разные положения слова “*report*” в материале 2

### Google Translate

Para1	Об этом говорится в <b>отчете</b> МВФ, опубликованном во вторник.
Para2	....., - говорится в <b>отчете</b> .
Para4	....., - говорится в <b>отчете</b> .
Para10	....., - говорится в <b>сообщении</b> .

Таблица 8: Разные положения слова “report” в переводе Google

### Prompt Translator

Para1	Об этом говорится в обнародованном во вторник <b>докладе</b> МВФ.
Para2	....., - говорится в <b>докладе</b> .
Para4	....., - говорится в <b>докладе</b> .
Para10	....., - говорится в <b>сообщении</b> .

Таблица 9: Разные положения слова “report” в переводе Prompt

### Yandex Translator

Para1	Об этом говорится в <b>докладе</b> МВФ, опубликованном во вторник.
Para2	....., - говорится в <b>докладе</b> .
Para4	....., - говорится в <b>докладе</b> .
Para10	....., - говорится в <b>докладе</b> .

Таблица 10: Разные положения слова “report” в переводе Yandex

За исключением того, что *Yandex* четыре раза переводил “report” в «доклад», на двух других платформах контекстуальный перевод был непоследовательным.

Далее, посмотрим на имя собственное “Faster Payment System”, которое встречается три раза в материале 10 [*Cross-border money transfers to appear in Faster Payment System in 2022 — Bank of Russia*], и их расположение следующее:

Title	Cross-border money transfers to appear in <b>Faster Payment System</b> in 2022 — Bank of Russia
Para1	Cross-border transfers through the <b>Faster Payment System (FPS)</b> can become available as early as in 2022, the Bank of Russia said on Monday.
Para2	Capabilities of the <b>Faster Payment System (FPS)</b> are anticipated to be expanded on account of implementing cross-border transfers...

Таблица 11: Разные положения “Faster Payment System” в материале 2

### Google Translate

Title	Международные денежные переводы появятся в <b>системе Faster Payment</b> в 2022 году
Para1	Международные переводы через <b>Faster Payment System (FPS)</b> могут стать доступными уже в 2022 году
Para2	Возможности <b>Системы ускоренных платежей (FPS)</b> , как ожидается, будут расширены за счет осуществления трансграничных переводов

Таблица 12: Разные положения “*Faster Payment System*” в переводе *Google*

### Prompt Translator

Title	Трансграничные денежные переводы появятся в <b>системе быстрых платежей</b> в 2022 году
Para1	Трансграничные переводы через <b>Систему более быстрых платежей (ФПС)</b> могут стать доступны уже в 2022 году
Para2	Предполагается, что возможности <b>системы более быстрых платежей (ФПС)</b> будут расширены за счет осуществления трансграничных переводов

Таблица 13: Разные положения “*Faster Payment System*” в переводе *Prompt*

### Yandex Translator

Title	Трансграничные денежные переводы появятся в <b>системе быстрых платежей</b> в 2022 году
Para1	Трансграничные переводы через <b>Систему быстрых платежей (FPS)</b> могут стать доступными уже в 2022 году
Para2	Ожидается, что возможности <b>Более быстрой платежной системы (FPS)</b> будут расширены за счет осуществления трансграничных переводов

Таблица 14: Разные положения “*Faster Payment System*” в переводе *Yandex*

Основываясь на приведенных выше результатах, мы обнаружили, что ни одна из платформ не перевела “*Faster Payment System*” полностью и последовательно. Наиболее признанным переводом “*Faster Payment System*” является «*система быстрых платежей (СБП)*». На самом деле, три платформы имеют разные ошибки при переводе этого имени собственного. *Google*

предпочитает сохранить английские имена собственные, но переводит его в «система ускоренных платежей (FPS)», когда слово появляется в тексте в третий раз; *Prompt* перевел сравнительную степень слова “*faster*” как «более» много раз, а при переводе “*FPS*” также прямо перевел как «ФПС», используя транслитерацию, *Yandex* перевел сравнительную степень слова “*faster*” как «более» один раз.

Объединив переводы этих двух существительных тремя платформами, можно сделать вывод, что будь то *Google*, *Prompt* или *Yandex*, в настоящее время нет возможности перевести часто встречающиеся существительные в одном и том же переводе.

## 2.2 Вертикальное сравнение

Вертикальное сравнение относится к сравнению исторического состояния, текущей ситуации и даже будущего состояния некоего предмета. Путем сравнения тех или иных предметов на разных исторических этапах мы можем понять законы их собственного движения, увидеть этапы и тенденции развития вещей, лучше понять прошлое и понять будущее, предотвратить разделение истории. Сравниваем результаты перевода *Google Translate*, *Prompt Translator* и *Yandex Translator* за два временных периода (версия 202104 и версия 202201) соответственно, и выясняем, где изменились результаты перевода версии 202104 и версии 202201. После статистики выяснилось, что при обработке одного и того же текста *Google*-переводом произошло 185 изменений, в *Prompt*-переводе — 13 изменений, в *Yandex* — 87 изменений.

<b>Platform name</b>	Google Translate	Prompt Translator	Yandex Translator
<b>Total number of changes</b>	185	13	87

Таблица 15: Общее количество изменений на трех платформах

Это подтверждает наше предположение, основанное на изменении количества слов в переводе в начале практической части: обработка текста *Google* со временем меняется больше всего, а *Prompt* меняется меньше всего. Мы делим все происходящие изменения на одиннадцать категорий: морфологические изменения, изменения слов, изменения частей речи, изменения аббревиатур, изменения структуры предложения, изменения порядка слов, изменения заглавных букв, изменения перевода чисел, изменения перевода единиц, изменения пунктуации, и изменения в переводе имен собственных.

Далее мы объясним и проанализируем результаты вертикального сравнения трех платформ одну за другой.

### 2.2.1 Анализ переводов *Google*

Между переводом *Google* версии 202104 и переводом версии 202201 есть 185 изменений. В следующей таблице 16 показаны типы изменений и их количество с апреля 2021 года по январь 2022 года в переводе *Google Translate* десяти статей экономических новостей:

<b>Google</b>	
Type of change	The total
Word change	109

Inflection	21
Word order change	12
Sentence structure change	11
Abbreviation usage	8
Change of part of speech	8
Unit translation change	6
Proper noun translation change	5
Punctuation change	2
Letter case change	2
Digital translation change	1

Таблица 16: Общее количество изменений в переводах *Google*

Как видно из приведенной выше таблицы, наибольшее количество изменений приходится на словоизменения, достигая 109. Среди них 80 изменений в подборе слов, 28 изменений в словарных дополнениях и опущениях и одно изменение, при котором прилагательные становятся прилагательными в сравнительной форме.

В процессе этого изменения некоторые предложения претерпели очень сложные и разнообразные изменения, некоторые изменения являются положительным улучшением, делающим перевод более совершенным, некоторые изменения являются ослаблением, которое делает перевод предложений не таким хорошим, как раньше. Есть также изменения, которые содержат ошибки как до, так и после их возникновения.

Пример сложных изменений в переводе показан в таблице 17:

Original	"The growth of the Belarusian economy in the first half of this year is higher than the commission expected as GDP growth rates totaled 103.3% by the end of the first half of the year," she noted.
Version 202104	«Рост <u>экономики Беларуси</u> в первом полугодии этого года <u>превышает ожидаемый комиссионный показатель</u> , так как <u>по итогам первого полугодия</u> темпы роста ВВП составили 103,3%», - отметила она.

Version 202201	"Рост <u>белорусской экономики</u> в первом полугодии этого года <u>выше, чем ожидала комиссия</u> , так как темпы роста ВВП <u>по итогам первого полугодия</u> составили 103,3%", - отметила она.
-------------------	--

Таблица 17: Пример сложных изменений в перевод

Это предложение взято из четвертого абзаца материала 6 – *“Eurasian Economic Commission projects EAEU GDP growth at 3.8% in 2021”*. В двух версиях этого предложения есть четыре изменения:

1. Изменение части речи: *«Рост экономики Беларуси»* - *«Рост белорусской экономики»*. Первоначально *«Беларусь»* была существительным в родительном падеже для модификации слова *«экономики»*, а в более поздних переводах прилагательное *«белорусская»* использовалось вместо существительного *«Беларусь»* для модификации *«экономики»*.

2. Изменилась структура предложения: *«превышает ожидаемый комиссионный показатель»* становится *«выше, чем ожидала комиссия»*, на самом деле, по смыслу этого предложения, мы можем обнаружить, что перевод последнего больше соответствует смыслу исходного предложения. Это изменение улучшает и оптимизирует перевод этого предложения.

3. Изменен порядок слов: в версии 202104 *«по итогам первого полугодия»* стоит перед *«темпы роста ВВП»*, а в версии 202201 *«по итогам первого полугодия»* стоит после *«темпы роста ВВП»*.

4. Изменения в пунктуации: это предложение на самом деле является прямой цитатой, поэтому в этом предложении используются двойные кавычки. В версии 202104 используется *«...»*, а в версии 202201 — *"..."*.

Пример оптимизированного перевода показан в таблице 18:

Original	This is joint responsibility and work of the government, regions, the authorities of all levels and our major companies, both state-run and private.
Version 202104	Это совместная ответственность и работа правительства, регионов, органов власти всех уровней и наших крупных компаний, как государственных, так и <u>государственных. частный.</u>
Version 202201	Это совместная ответственность и работа правительства, регионов, органов власти всех уровней и наших крупнейших компаний, как государственных, так и <u>частные.</u>

Таблица 18: Пример оптимизированного перевода

Это предложение взято из второго абзаца Материала 3 “*Putin highlights Far East’s advanced development as Russia’s top priority*”. В версии 202104 есть типичная ошибка *Google Translate*, дублирование переводов и пунктуационные ошибки. Повторный перевод «*state-run*» при переводе «*both state-run and private*», с двойным появлением «*государственных*» в предложении и лишней точкой в предложении. В то время как в версии 202201 исправлены дублирующиеся переводы и пунктуационные ошибки, перевод «как государственный, так и частный» является правильным.

Пример ослабленных переводов показан в таблице 19:

Original	...the corresponding court hearing will be held on September 14, the world court No. 422 of the Tagansky district of Moscow....
Version 202104	...соответствующее судебное заседание состоится 14 сентября <u>в мировом суде № 422 Таганского района Москвы...</u>
Version 202201	...соответствующее судебное заседание состоится 14 сентября, <u>мировой суд № 422 Таганского района Москвы...</u>

Таблица 19: Пример ослабленных переводов

Это предложение взято из первого абзаца материала 4 “*Facebook, Twitter, Telegram face \$988,605 fine for prohibited content*”. Первоначальное предложение означает: Заседание состоится в Мировом суде 14 сентября. Версия 202104 использовала наречие местоположения «*в мировом суде № 422 Таганского района Москвы*», а перевод версии 202201 скопировал структуру предложения оригинального английского текста и принял метод дословного перевода. Перевод версии 202104 представляется более приемлемым и понятным.

Примеры ошибок в обоих вариантах перевода показаны в таблице 20 и таблице 21:

Original	(an improvement of 1 pp compared to the January forecast)
Version 202104	(улучшение на 1 п.п. по сравнению с <u>прогнозом в январе</u> ). Январский прогноз)
Version 202201	(улучшение на 1 п.п. по сравнению с <u>прогнозом января</u> ). январский прогноз)

Таблица 20: Первый пример ошибок в обоих вариантах перевода

Первый пример взят из второго абзаца Материала 1 “*IMF improves forecast for US GDP growth rates in 2021 and 2022*”. Единственное изменение, которое происходит здесь, заключается в том, что «*Январский*» становится «*январский*», но дублирующиеся переводы и пунктуационные ошибки сохраняются.

Original	address governance shortcomings and take steps to reduce the regulatory burden
Version 202104	устранении недостатков в управлении и <u>принятии мер по снизить</u> регуляторную нагрузку
Version 202201	устранения недостатков управления и <u>принятия снизить</u>

202201	регулятивную нагрузку
--------	-----------------------

Таблица 21: Второй пример ошибок в обоих вариантах перевода

Второй пример взят из десятого абзаца Материала 2 “*IMF praises Russia’s measures to fight economic consequences of pandemic*”. В версии 202104 “*take steps to reduce*” было переведено как «принятии мер по снизить», а в версии 202201 “*take steps to reduce*” было переведено как «*принятия снизить*». Здесь есть два вида ошибок: первый - это ошибка смысла предложения. В этом предложении «*address governance shortcomings*» и «*take steps to reduce the regulatory burden*» являются противопоставленными компонентами, которые в переводе становятся «управлении» и «принятии мер» — противопоставленными компонентами, что искажает смысл исходного предложения; второй является грамматической ошибкой, ни «*принятии мер по снизить*» в версии 202104, ни «*принятия снизить*» в версии 202201 не соответствуют правилам русской грамматики.

### 2.2.2 Анализ переводов *Prompt*

Между переводом *Prompt* версии 202104 и переводом версии 202201 есть 11 изменений. В следующей таблице показаны типы изменений и их количество с апреля 2021 года по январь 2022 года в переводе *Prompt Translate* десяти статей экономических новостей:

Prompt	
Word change	6
Inflection	2
Punctuation change	2
Change of part of speech	1
Sentence structure change	1

Word order change	1
-------------------	---

Таблица 22: Общее количество изменений в переводах *Prompt*

Как видно из приведенной выше таблицы 22, тип изменения с наибольшим числом — это изменение слова, а всего изменений 6. Среди них 3 изменения в выборе слов и 3 изменения в словарных дополнениях и опущениях.

Большинство изменений в переводе *Prompt* не окажут большого влияния. Но стоит упомянуть два оптимизированных перевода, которые показаны в таблице 23 и таблице 24:

Original	About \$1 trillion should be directed to providing assistance to the population
Version 202104	Около \$1 трлн нужно направить <u>на предусматривать</u> помощи населению
Version 202201	Около \$1 трлн нужно направить <u>на оказание помощи</u> населению

Таблица 23: Первый оптимизированный перевод

Первый пример взят из четвертого абзаца Материала 1 “*IMF improves forecast for US GDP growth rates in 2021 and 2022*”. Согласно правилам русской грамматики, «*направить на*» управляется винительным падежом (кого?/что?). В версии 202104 за «*направить на*» следует глагол «*предусматривать*», что неверно. Но в версии 202201 за «*направить на*» следует существительное «*оказание*». Это изменение исправляет ошибку перевода в версии 202104.

Original	Its member states are Russia, Belarus, Kazakhstan, Armenia and Kyrgyzstan.
Version 202104	<u>Его члены государствами</u> являются Россия, Белоруссия, Казахстан, Армения и Киргизия.
Version	<u>Его государствами-членами</u> являются Россия, Белоруссия,

202201	Казахстан, Армения и Киргизия.
--------	--------------------------------

Таблица 24: Второй оптимизированный перевод

Второй пример взят из шестого абзаца материала 6 “*Eurasian Economic Commission projects EAEU GDP growth at 3.8% in 2021*”. В переводе версии 202104 есть грамматическая ошибка, “*member state*” может быть переведено как «член государства» или «государство-член». Таким образом, правильный перевод должен быть следующим: «Его представители являются Россией, Белоруссией, Казахстаном, Арменией и Киргизией» или «Его участниками являются Россия, Белоруссия, Казахстан, Армения и Киргизия». В этом случае правильным переводом будет версия 202201.

### 2.2.3 Анализ переводов *Yandex*

Между переводом *Yandex* версии 202104 и переводом версии 202201 есть 87 изменений. В следующей таблице 25 показаны типы изменений и их количество с апреля 2021 года по январь 2022 года в переводе *Yandex Translator* десяти статей экономических новостей:

Yandex	
Word change	53
Word order change	11
Unit translation change	8
Change of part of speech	5
Inflection	5
Proper noun translation change	3
Sentence structure change	2

Таблица 25: Общее количество изменений в переводах *Prompt*

Как видно из приведенной выше таблицы, тип изменения с наибольшим числом — это изменение слова, а всего изменения 53. Среди них 47 изменений в выборе слов и 6 изменений в словарных дополнениях и опущениях.

Объединив перевод десяти статей экономических новостей за период с апреля 2021 г. по январь 2022 г. на трех вышеупомянутых платформах, типы и количество изменений, мы видим, что наибольшее количество изменений на трех платформах перевода составляют изменения слов. Помимо изменений слов, на трех платформах произошли изменения порядка слов, изменения частей речи, морфологические изменения и изменения структуры предложения.

Наименьшие изменения в переводах *Prompt* указывают на то, что машинные переводы нейронной сети *Prompt* обучались меньше всего и медленнее всего с течением времени, но два из одиннадцати изменений были оптимизированными переводами, что заслуживает признания. Изменения в *Yandex*-переводе расположены посередине, и большинство изменений мало влияют на качество перевода, такие как замена синонимов, изменение порядка слов и т. д., и не так много аспектов, которые можно использовать для исследования. В *Google Translate* больше всего изменений, и он наиболее достоин нашего исследования и обсуждения. В нем больше всего типов изменений, и результаты, полученные после изменений, различны, что может быть оптимизированным переводом или ослабленным переводом. Ошибка может остаться прежней или может измениться от одного вида ошибки к другому виду ошибки, которая на самом деле может хорошо представлять процесс развития машинного перевода и процесс обучения нейронных сетей, В

целом, машинный перевод постоянно развивается, и нейронные сети постоянно продвигают прогресс машинного перевода, но после тщательного изучения все еще остается много недостатков и дефектов, и в машинном переводе еще много возможностей для улучшения.

### **2.3 Анализ результатов перевода с точки зрения экономических текстов**

Экономические тексты, являющиеся одним из научных текстов, должен максимально сохранять больше информации оригинального текста, сохраняя при этом читаемость и избегая слишком большой субъективной инициативы по выражению эстетической ценности текста. Эти характеристики определяют, что перевод экономического текста отличается от других литературных переводов.

В вышеприведенном исследовании мы проанализировали результаты машинного перевода с точки зрения качества перевода (ошибка перевода и непротиворечивость перевода) и изменений перевода с течением времени. Далее мы проанализируем результаты перевода версии 202104 на основе характеристик экономических текстов.

В нашей подборке экономических текстов на английском языке много цифр и денежных единиц, а также экономических метафор. Все три системы перевода переводили общие номера без ошибок, но были несоответствия между сокращениями (например, “*mln*”) и денежными единицами (например, “\$”).

Например, в материале 8 “*Google pays \$47,174 out of \$438,047 of fines in Russia*” отображается большое количество денежных единиц, как показано в следующей таблице 26:

Original	Google	Prompt	Yandex
\$47,174	47 174 доллара	\$47 174	47 174 доллара
\$438,047	438 047 долларов	\$438 047	438,047 долларов
\$438,047	438 047 долларов США	\$438 047	\$438,047
\$47,174	47 174 доллара США	\$47 174	\$47,174
\$350,437	350 437 долларов США	\$350 437	350 437 долларов США
\$40,436	40 436 долларов США	\$40 436	40 436 долларов США
\$54,040	54 040 долларов	\$54 040	54 040 долларов
\$229,805	229 805 долларов США	\$229 805	229 805 долларов
\$202,769	202 769 долларов США	\$202 769	202 769 долларов
\$54,069	54 069 долларов	\$54 069	54 069 долларов США
3 mln rubles	3 млн рублей	3 млн рублей	3 млн рублей
4 mln rubles	4 миллиона рублей	4 млн рублей	4 млн рублей
17 mln rubles	17 млн рублей	17 млн рублей	17 млн рублей
15 mln rubles	15 млн рублей	15 млн рублей	15 млн рублей
4 mln rubles	4 млн рублей	4 млн рублей	4 миллиона рублей
32.5 mln rubles	32,5 млн рублей	32,5 млн рублей	32,5 млн рублей
3.5 mln rubles	3,5 млн рублей	3,5 млн рублей	3,5 млн рублей
32.5 mln rubles	32,5 млн рублей	32,5 млн рублей	32,5 млн рублей
26 mln rubles	26 млн рублей	26 млн рублей	26 млн рублей

Таблица 26: Денежные единицы в оригинале и их переводы

Согласно приведенной выше таблице, в переводе одного и того же текста только *Prompt Translator* может поддерживать перевод денежных единиц согласованным, в то время как *Google Translate* и *Yandex Translator* имеют несколько версий перевода при переводе денежных единиц. Например, когда *Google Translate* переводит денежную единицу доллара США “\$”, существуют две версии перевода: «доллара» и «долларов США»; когда *Yandex Translator* переводит единицу валюты доллара США «\$», есть три версии перевода: «доллара», «долларов США» и зарезервированная единица «\$». При переводе

“mln”, будь то перевод *Google Translate* или *Yandex Translator*, существуют две версии перевода: «миллион» и «мил», одна - полное название, а другая - аббревиатура.

Конечно, как бы ни переводилась денежная единица, сохраняет ли “mln” аббревиатуру или нет, это не повлияет на понимание статьи читателем при чтении, но наличие разных версий переводов в одном тексте повлияет на строгость научных текстов.

Кроме того, мы часто сталкиваемся с метафорами при чтении экономических текстов. Например, во втором материале “*IMF praises Russia’s measures to fight economic consequences of pandemic*” есть предложение: *Executive Directors commended the authorities for a sizeable response to the crisis which should help put a floor on the downturn and limit scarring*. Согласно словарю [Colin, 2011], “scar” означает шрам или рубцы. Но неуместно переводить “scarring” в «шрамы» или «рубцы» в этом предложении, что сделает это предложение запутанным. Рассмотрим перевод этого предложения тремя системами в версии 202104 (Таблица 27):

Google	Исполнительные директора поблагодарили власти за значительный ответ на кризис, который должен помочь положить конец спаду и ограничить <b>образование рубцов</b> .
Prompt	Директора-исполнители выразили признательность властям за значительные меры реагирования на кризис, которые должны помочь положить конец спаду и ограничить <b>образование рубцов</b> .
Yandex	Исполнительные директора высоко оценили значительный ответ властей на кризис, который должен помочь положить конец спаду и <b>ограничить рубцы</b> .

Таблица 27: Разные переводы метафора в материале 2

Мы обнаружили, что все три системы перевода прямо переводят “*scarring*” в «*рубцы*», что оставляет читателей переводов озадаченными. На самом деле, в сочетании со значением всего предложения, я думаю, что “*limit scarring*” должно означать снижение потерь. Поэтому в этом предложении “*limit scarring*” было бы более уместно перевести на «*уменьшить потери*».

Язык - это продукт человеческого развития, язык постоянно развивается и есть гибкость. Более гибкие и интересные выражения, такие как шутки и метафоры - это языковые барьеры, которые трудно преодолеть человеческим переводчикам. Человеческие переводчики могут обрабатывать перевод в соответствии с фоном исходного текста и намерением первоначального автора, чтобы читатели перевода могли лучше понять смысл, который хочет выразить первоначальный автор. Но пока нет возможности для машинного перевода добиться разумной обработки таких гибких выражений.

Машинный перевод основан на оригинальном предложении и данных, поэтому он имеет определенное преимущество по сравнению с человеческим переводчиком в плане сохранения объективности, но нет способа думать и понимать исходное предложение так же, как человеческий мозг, что также делает машинный перевод неспособным хорошо справляться с метафорой, появляющейся в экономических текстах.

## **2.4 Обсуждение взаимосвязи между машинным переводом и человеческим переводом**

Многие документы, которые переводчики переводят каждый день, придумываются в спешке. Без специального редактирования будут неточные выражения. Человеческие переводчики сталкиваются с некачественными документами, исследуют или просят авторов выяснить смысл оригинального текста, а затем улучшить его в переводе. Но машинный перевод будет только верно переводить на основе оригинальных текстов.

Переход от статистического машинного перевода и машинного перевода на основе фраз к машинному переводу на основе нейронной сети значительно повысил общее качество машинного перевода. Но нейронный машинный перевод опирается на огромные объемы обучающих данных даже больше, чем предыдущие модели. Поскольку самые большие доступные двуязычные наборы данных получены из официальных переводов правительственных документов и религиозных текстов, эти алгоритмы редко подвергаются юмористическим выражениям, словесности и невербальным выражениям.

Самое тревожное, что сторона нейронного машинного перевода часто не признает своих ошибок, а вместо этого действует как плохо подготовленный школьник, пытающийся уйти от ответственности во имя машины. Поскольку некоторые читатели целевого языка могут никогда не знать о проблеме, и могут приписать причудливость перевода самому исходному тексту, эти проблемы менее заметны, чем неуклюжие, явные ошибки.

Можно видеть, что, хотя машинный перевод имеет преимущества высокой скорости и низкой стоимости, машинный перевод также имеет много ограничений и слепых зон.

Во-первых, машинный перевод не чувствителен к культуре. Люди никогда не смогут запрограммировать машины, чтобы понять и испытать определенную культуру. Различные культуры имеют уникальные и разные языковые системы, и машины еще не достаточно мощны, чтобы распознавать некоторые гибкие и сложные выражения через свои собственные возможности понимания, поэтому полученные переводы могут не соответствовать ценностям культуры и конкретным нормам. Это также одна из проблем, которую должно преодолеть машинный перевод.

Во-вторых, машинный перевод не может соединять слова с контекстуальным значением. Во многих языках одно и то же слово может иметь несколько совершенно не связанных значений. В этом случае контекст окажет большое влияние на значение слова, а понимание значения слова в значительной степени зависит от подсказок, прочитанных из контекста. Только люди могут сочетать слова с контекстом, определять их истинное значение и творчески обрабатывать язык для получения полных и точных переводов. Для машинного перевода это, несомненно, очень сложно.

В-третьих, почти все результаты машинного перевода требуют постредактирования человеком. Может быть, компьютер получит систему естественного языка, похожую на систему человека, но в любом случае постредактирование не может быть отброшено. После того, как машина закончит перевод, корректор должен внести окончательные изменения в перевод, чтобы убедиться, что результаты машинного перевода используют правильную грамматику и выражают понятную семантику.

Учитывая упомянутые выше недостатки, споры о том, заменит ли машинный перевод человеческие переводчики, представляются ненужными. В большинстве случаев ответ на этот вопрос - нет. Хотя мы не можем недооценивать потенциал машинного перевода, мы должны понимать, что машинный перевод не может заменить человека до тех пор, пока требуется лингвистическая точность перевода. Вместо этого машинный перевод может помочь людям со всевозможными огромными и утомительными задачами.

Как два различных типа методов перевода, существует очень тесная связь между машинным переводом и переводом человека, что в основном отражается в следующих трех аспектах.

Во-первых, реализация машинного перевода основана на человеческом переводе. С точки зрения эффекта перевода, человеческий перевод может полировать и улучшить результаты машинного перевода. Потому что в процессе машинного перевода функциональным пунктом его системы перевода по-прежнему остается двуязычный корпус выравнивания. Информация в этом корпусе основана на человеческом переводе. Двуязычная ссылка преобразования также основана на этом корпусе. Видно, что на эффект и качество машинного перевода в некоторой степени влияют качество и уровень корпуса, основанного на человеческом переводе. Кроме того, по характеру языка относится к своеобразному гибкому и расходящемуся содержанию. В процессе перевода человеческий переводчик должен пройти через процесс мышления и воссоздания. Из этого измерения, машинный перевод - это процесс гуманизированного мышления и преобразования, перевод текста более помечен

механическими характеристиками, а процесс и метод преобразования более прямые.

Во-вторых, с точки зрения эффективности работы по переводу высокая концентрация и единство машинного перевода определяет, что общая эффективность работы по машинному переводу выше, чем у человеческий перевод в процессе применения. И пока нет ошибок в содержании базовой лексики и параграфов, работа по переводу может вестись с высокой точностью. Этот высокий уровень гарантии с точки зрения эффективности и точности является эффектом и степенью, которые не могут быть достигнуты человеческим переводом.

В-третьих, машинный перевод имеет сильную фиксацию. Это в основном означает, что машинный перевод имеет определенные требования к объему использования. Машинный перевод обычно подходит для переводческих объектов с относительно высокой однородностью с точки зрения общих стандартов перевода и требований к переводу, и может непрерывно выполнять задачи пакетного перевода. Для человеческого перевода его главная особенность заключается в том, что организация и осуществление переводческой работы будет более гибким, и она может удовлетворить потребности различных переводческих потребностей в соответствии с конкретными рабочими потребностями. И при требованиях сложной переводческой работы, хотя эффективность индивидуального человеческого перевода будет ограничена, этот способ перевода может воспользоваться гибкостью кадровой организации для достижения организации и оптимизации

переводческих бригад и повышения эффективности сложного перевода. Эффективность работы и требования к переводу могут соответствовать подробным требованиям к переводу. Это эффект, которого не может достичь машинный перевод.

С точки зрения взаимосвязи между ними машинный перевод и человеческий перевод имеют различия в процессе внедрения и эффекте работы по переводу. Однако в процессе функционирования между ними существует тесная внутренняя связь. Это также показывает, что человеческий перевод и машинный перевод могут обеспечить совместную работу в определенном диапазоне, а с точки зрения конкретных рабочих эффектов эффект перевода также может быть улучшен путем корректировки организационной формы.

Поэтому замена человеческого перевода на машинный не является целью исследования и развития машинного перевода, но ключевым вопросом является то, как лучше использовать отношения между ними для предоставления услуг и руководства для развития работы по переводу.

## Вывод по Главе 2

1. Как в версии 202104, так и в версии 202201, *Promt* и *Yandex* имеют наибольшее сходство в переводе десяти экономических текстов, а *Google* и *Promt* - наименьшее. Это может свидетельствовать о том, что *Promt* и *Yandex* имеют наиболее схожий механизм перевода и поток обработки текста, в то время как *Google* и *Promt* имеют наименее похожий механизм перевода и поток обработки текста.
2. Совмещая результаты перевода версии 202104 и версии 202201, в переводе *Google* и переводе *Promt* много ошибок, что сильно снижает качество перевода. Переводы, переведенные *Yandex*, имеют наименьшее количество ошибок. По количеству ошибок у перевода *Yandex* самое высокое качество перевода, а у *Google* самое низкое.
3. Судя по результатам перевода трех платформ для имен нарицательных и имен собственных в одной и той же статье, ни одна из трех платформ в настоящее время не имеет способа поддерживать согласованность перевода для слов, которые часто встречаются в одном и том же переводе, и существует нет способа сохранить согласованность перевода. Согласованность перевода также снижает качество перевода.
4. Судя по количеству ошибок в переводе и согласованности перевода существительных, качество результатов русского перевода выбранных десяти экономических англоязычных новостных статей переводом *Yandex*

относительно высокое, а перевода *Google* самые явные ошибки. , качество перевода относительно низкое.

5. Путем вертикального сравнения мы обнаруживаем, что обработка одного и того же текста *Google Translate* больше всего менялась с течением времени, *Yandex Translator* занимает второе место по количеству изменений, а *Prompt Translator* — меньше всего изменений. Это частично отражает то, насколько хорошо нейронные сети трех платформ обучались с течением времени.

Нейронные сети трех платформ постоянно обучаются, но нейронная сеть *Google Translate* обучается быстрее всех и, следовательно, больше всего меняется; в то время как нейронная сеть быстрого перевода учится медленнее всего и, следовательно, меньше всего изменяется.

6. Объединяя изменения в обработке текста трех платформ с течением времени, после дальнейшего обучения нейронной сети машинного перевода качество перевода может не улучшаться абсолютно, а может снижаться. И большая часть изменений — это замена слов, особенно замены синонимов, которые мало влияют на качество перевода. В целом машинный перевод постоянно совершенствуется, но все же есть серьезные недостатки. Качество перевода по-прежнему не столь удовлетворительно, а ручное вмешательство по-прежнему требуется для редактирования после перевода.

7. Для анализа качества перевода машинного перевода с точки зрения характеристик экономических текстов можно обнаружить, что при переводе экономических текстов с множеством цифр, символов валют и метафор машинный перевод не может получить скрупулезного и точного результата

перевода. Когда данные существуют в виде арабских цифр, машинный перевод может переводить их более точно, но когда данные существуют в виде относительно большого количества цифр (таких как миллионы, миллиарды), машинный перевод может быть не в состоянии поддерживать согласованность перевода при их переводе. Это приводит к ситуациям, когда большое число в одном и том же переводе имеет и полное название, и аббревиатуру. Кроме того, при переводе символов валют (особенно долларов США) часто встречаются несоответствия перевода, но *Prompt* справляется с этим лучше, чем две другие системы перевода. Как гибкое выражение, метафоры нужно искусственно понимать фон предложения и смысл, который хочет выразить автор.

Машинный перевод обычно использует перевод литералов. Учитывая культурные различия, читатели переведенного текста могут запутаться в буквальном переводе метафоры.

8. Хотя качество перевода машинного перевода постоянно улучшается, оно по-прежнему полно ошибок и упущений при работе с более сложными научными текстами, такими как экономические тексты, а получить безупречный результат перевода тем более невозможно. Если вы хотите получить качественные тексты экономического перевода, вам все равно необходимо освоить переводчиков со знаниями в области экономики и сильным уровнем иностранного языка для выполнения ручного перевода, или постпереводческого редактирования на основе машинного перевода. То есть для получения более совершенного результата перевода по-прежнему требуется участие человека.

## Заключение

В первой главе кратко представлены соответствующие теории машинного перевода вокруг истории развития, определения, классификации машинного перевода, характеристик экономических текстов и трудностей перевода, критериев оценки машинного перевода и ошибок машинного перевода. Мы можем обнаружить, что машинный перевод быстро развивался в последние годы, и появление глубокого обучения значительно улучшило качество перевода машинного перевода. С развитием технологий и потребностей связи исследования машинного перевода продолжают делать прорывы, а теория машинного перевода продолжает развиваться.

Во второй главе мы сравним результаты перевода выбранных английских экономических текстов тремя системами по горизонтали и вертикали. Мы оцениваем качество перевода машинного перевода, подсчитывая ошибки в переводе и проверяя непротиворечивость перевода соответствующих существительных. Мы также изучаем скорость и эффективность глубокого обучения нейронной сети трех систем, анализируя, как каждая система обрабатывает один и тот же текст с течением времени. Наконец, мы используем характеристики экономических текстов в качестве отправной точки для анализа качества результатов машинного перевода экономических текстов, а затем обсуждаем противоречие и взаимосвязь между машинным переводом и человеческим переводом.

Взяв за отправную точку экономические тексты, в этой работе мы протестировали три нейросетевые системы машинного перевода и сделали вывод о том, что экономические тексты обладают определенными характеристиками, включая большое количество чисел, единиц, терминов. Текстовые характеристики экономических текстов также предъявляют высокие требования к точности и совершенству переводческого содержания. В целом, качество перевода нейросетевой системы машинного перевода все еще очень низкое, и системы могут передать основной смысл предложения, но существующие ошибки и несоответствия в переводе говорят нам о том, что результаты перевода нейросетевой системы машинного перевода до сих пор не могут достичь высокого уровня стандарта. Хотя системы машинного перевода на основе нейронных сетей постоянно совершенствуются и учатся, еще есть много возможностей для совершенствования.

Стоит констатировать, что развитие машинного перевода достигло сегодня выдающихся показателей, что в определенной степени облегчило жизнь людей. Программное обеспечение для машинного перевода предоставляет удобные функции поиска слов и перевода в повседневной жизни. Хотя сервис не идеален и значение многих переводов неточное, оно все же расширяет возможности общения людей, снижает стоимость перевода для обычных людей и делает перевод более удобным. В настоящее время все больше и больше программного обеспечения машинного перевода использует для перевода метод нейронной сети (neural machine translation, NMT), который в определенной степени улучшает качество перевода.

С быстрым развитием машинного перевода сегодня многие люди сомневаются, действительно ли машинный перевод заменит человеческий перевод. Но ответ очевиден: машинный перевод без участия человека никогда не заменит человека. Те, кто заинтересован в переводческой деятельности, должны усердно изучать иностранные языки, овладевать навыками перевода, читать много текстов и накапливать устойчивые выражения, чтобы улучшить свои переводческие способности и способствовать развитию машинного перевода.

Перспективы исследования данной проблемы состоят в возможности дальнейшего более глубокого изучения машинного перевода, особенно нейронный машинный перевод, по оптимальному решению и преодолению наличных проблем.

## Список литературы

1. Александрова Е.И. Экономический термин как элемент экономического дискурса при обучении экономическому переводу на материале текстов СМИ // Вестник Московского государственного областного университета. Серия: Педагогика. 2013. № 1. С. 54-60.
2. Антонова О.А., Мартынова И.Н. Грамматические и лексические ошибки при использовании систем машинного перевода “PROMT” и “GOOGLE TRANSLATE” (на материале экономического текста) // В сборнике: Вопросы переводоведения, межкультурной коммуникации и зарубежной литературы. сборник научных статей. Чувашский государственный педагогический университет им. И.Я. Яковлева. Чебоксары, 2020. С. 8-14.
3. Газизов И.Ф. Трудности перевода экономических терминов и терминологических словосочетаний в английских экономических текстах // В сборнике: Фундаментальные и прикладные научные исследования: актуальные вопросы, достижения и инновации. сборник статей VIII Международной научно-практической конференции: в 4 частях. 2017. С. 158-160.
4. Гончаров А.А., Бунтман Н.В., Нуриев В.А. Ошибки в машинном переводе: проблемы классификации // Системы и средства информатики. 2019. Т. 29. № 3. С. 92-103.
5. Дьяченко И.Н., Матыченко Ю.В. Нейронный машинный перевод: преимущества, сложности, перспективы // Языки и литература в поликультурном пространстве. 2020. № 6. С. 28-33.

6. Дроздова, К. А. Машинный перевод: история, классификация, методы / К. А. Дроздова. — Текст : непосредственный // Филологические науки в России и за рубежом : материалы III Междунар. науч. конф. (г. Санкт-Петербург, июль 2015 г.). — Санкт-Петербург : Свое издательство, 2015. — С. 139-141.
7. Карасев И. В., Артюшина Е. А. Системы машинного перевода // Успехи современного естествознания. 2011, №7, С.117-118.
8. Крашенинникова Н.А. Классификация типичных ошибок при использовании систем машинного перевода // Симбирский научный вестник. 2011. № 2 (4). С. 190-194.
9. Легостина М.С. Метрики оценки качества машинного перевода // В сборнике: ИННОВАТИКА-2019. сборник материалов XV Международной школы-конференции студентов, аспирантов и молодых ученых. 2019. С. 490-493.
10. Лядащева Е.Д. Экономический текст как особый тип текста // Филология и лингвистика. 2018. № 3 (9). С. 29-30.
11. Мошкович В.В. Оценка качества перевода и использование адекватности и эквивалентности как критериев оценки качества перевода // Вестник Челябинского государственного педагогического университета. 2013. № 10. С. 291-297.
12. Нанкевич А.А. Оценка качества результатов систем машинного перевода // Системы компьютерной математики и их приложения. 2019. № 20-1. С. 215-221.

13. Новожилова А. А. Машинные системы перевода: качество и возможности использования // Вестник ВолГУ. Серия 2: Языкознание. 2014. №3 С.67-73.
14. Павлов В.Е., Башмакова И.С. Нейронная сеть и машинный перевод // Молодежный вестник ИрГТУ. 2017. № 2. С. 16.
15. Переходько И.В., Мячин Д.А. Оценка качества компьютерного перевода // Вестник Оренбургского государственного университета. 2017. № 2 (202). С. 92-96.
16. Полякова Н.Е. О критериях оценивания качества машинного перевода // Тенденции развития науки и образования. 2020. № 61-8. С. 84-87.
17. Ртищева Н.Г. Оценка качества работы систем машинного перевода // Решетневские чтения. 2017. Т. 2. С. 704-705.
18. Рычихин А.К. О методах оценки качества машинного перевода // Системы и средства информатики. 2019. Т. 29. № 4. С. 106-118.
19. Степанов Д.А., Федорцов П.А. Машинный перевод: преимущества и недостатки // В книге: Тезисы докладов 52-й Международной научно-технической конференции преподавателей и студентов. Тезисы докладов 52-й Международной научно-технической конференции преподавателей и студентов. 2019. С. 14-15.
20. Талызина Е.В. Анализ типичных ошибок и их причин при использовании средств машинного перевода // В книге: Проблемы теории и практики современной науки. Материалы VI Международной научно-практической конференции: Сборник научных трудов. ООО «НОУ «Вектор науки». 2016. С. 120-124.

21. Улиткин И.А. Автоматическая оценка качества машинного перевода научно-технического текста // Вестник Московского государственного областного университета. Серия: Лингвистика. 2016. № 4. С. 174-182.
22. Фаустова К.И. Нейронные сети: применение сегодня и перспективы развития // Территория науки. 2017. № 4. С. 83-87.
23. Шаляпина З. М. Автоматический перевод: Эволюция и современные тенденции // Вопросы языкознания, 1996, №2, С. 105—117.
24. Colin McIntosh. Cambridge Learner's Dictionary English-Russian. Cambridge ELT, UK, 2011.
25. Papineni, K.; Roukos, S.; Ward, T.; Zhu, W. J., BLEU: a method for automatic evaluation of machine translation // ACL 2002, USA, P311-318.
26. Philipp Koehn. Statistical Machine Translation. Cambridge, UK, 2010.
27. Vilar D., Jia Xu, D'Haro L., Ney H. Error Analysis of Statistical Machine Translation Output. // Proceedings of the Fifth International Conference on Language Resources and Evaluation, Genoa, Italy, 2006, P. 697—702.

## Список источников

28. Машинный перевод от холодной войны до диплёрнинга  
URL: [https://vas3k.ru/blog/machine\\_translation/#scroll20](https://vas3k.ru/blog/machine_translation/#scroll20) (дата обращения: 20.12.2020)
29. IMF improves forecast for US GDP growth rates in 2021 and 2022  
URL: <https://tass.com/economy/1274775> (дата обращения: 29.03.2021)
30. IMF praises Russia's measures to fight economic consequences of pandemic  
URL: <https://tass.com/economy/1254455> (дата обращения: 29.03.2021)
31. Putin highlights Far East's advanced development as Russia's top priority  
URL: <https://tass.com/economy/1333593> (дата обращения: 29.03.2021)
32. Facebook, Twitter, Telegram face \$988,605 fine for prohibited content  
URL: <https://tass.com/economy/1334045> (дата обращения: 29.03.2021)
33. Sber, Gazprombank to provide \$2.7 bln loan to Russian Copper Company  
URL: <https://tass.com/economy/1333427> (дата обращения: 29.03.2021)
34. Eurasian Economic Commission projects EAEU GDP growth at 3.8% in 2021  
URL: <https://tass.com/economy/1322521> (дата обращения: 29.03.2021)
35. Russia ranks second by petroleum exports to US in June 2021 — EIA  
URL: <https://tass.com/economy/1332285> (дата обращения: 29.03.2021)
36. Google pays \$47,174 out of \$438,047 of fines in Russia  
URL: <https://tass.com/economy/1330647> (дата обращения: 29.03.2021)
37. Shanghai Stock Exchange sets up capital market service base on Hainan  
URL: <https://tass.com/economy/1330611> (дата обращения: 29.03.2021)

38. Cross-border money transfers to appear in Faster Payment System in 2022 —

Bank of Russia

URL: <https://tass.com/economy/1329133> (дата обращения: 29.03.2021)