

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РФ
Федеральное государственное автономное образовательное учреждение
высшего образования
«Южно-Уральский государственный университет
(национальный исследовательский университет)»
Институт лингвистики и международных коммуникаций
Кафедра лингвистики и перевода

РАБОТА ПРОВЕРЕНА
Рецензент, доцент

_____ /Е.И. Хабирова/

ДОПУСТИТЬ К ЗАЩИТЕ
Заведующий кафедрой,
д.филол.н., доцент

_____ /Т.Н. Хомутова/

**АВТОМАТИЧЕСКИЙ ПЕРЕВОД КАК ИНСТРУМЕНТ ПОВЫШЕНИЯ
ЭФФЕКТИВНОСТИ РАБОТЫ ПЕРЕВОДЧИКА**

ВЫПУСКНАЯ КВАЛИФИКАЦИОННАЯ РАБОТА
ЮУрГУ – 45.04.01.2019.031.ВКР

Руководитель, к.ф.н., профессор
_____ /С.О. Шереметьева/
« ____ » _____ 2019 г.

Автор
студент группы ЛМ-230
_____ /О.И. Калининкова/
« ____ » _____ 2019 г.

Нормоконтролер,
к.филол.н., доцент
_____ /О.И. Бабина/
« ____ » _____ 2019 г.

Работа защищена с оценкой

« ____ » _____ 2019 г.

Челябинск
2019

ОГЛАВЛЕНИЕ

Введение.....	3
Глава 1 Аспекты переводческой деятельности.....	7
1.1 Профессиональная деятельность переводчика	7
1.2 Инструменты повышения эффективности работы переводчика.....	12
1.2.1 История развития автоматического перевода.....	18
1.2.2 Типы систем автоматического перевода	25
1.2.3 Обзор пользовательских возможностей в системах ПРОМТ и Google Translate.....	29
1.2.3. Оценка качества автоматического перевода	38
1.2.4. Пути повышения качества автоматического перевода.....	43
Выводы по главе 1.....	52
Глава 2 Анализ эффективности использования систем автоматического перевода ПРОМТ и Google Translate на материале экономических текстов..	53
2.1 Важность предметной области экономики.....	53
2.2 Особенности англоязычных текстов по макро- и мегаэкономике.....	55
2.3 Сопоставительный анализ переводов экономических текстов в системах ПРОМТ, Google Translate	59
Выводы по главе 2.....	80
Заключение	82
Библиографический список	85
Приложение 1.....	91
Приложение 2.....	134
Приложение 3.....	136
Приложение 4.....	139

ВВЕДЕНИЕ

Современный уровень развития техники, информационных технологий и средств коммуникации обеспечивает доступ к большим объемам многоязычной информации. Ввиду постоянного расширения и усложнения информационных потоков вопрос о скорости и качестве их обработки стоит особенно остро. Обработка «вручную» связана со значительными временными и материальными затратами, поэтому в настоящее время невозможно обойтись без системы, обеспечивающей переработку информации в быстрые сроки – системы автоматического перевода. Однако качество получаемого перевода не совершенно, поскольку оно зависит от лингвистических явлений текста, с которыми системы АП не всегда могут успешно справиться, т.е. от индикаторов переводимости. В настоящей работе рассматриваются способы повышения эффективности работы переводчика, а именно, с помощью системы автоматического перевода (АП).

Актуальность исследования обусловлена необходимостью разработать рекомендации по повышению эффективности работы с данными системами. Сформулированные рекомендации призваны помочь как обычным пользователям-непрофессионалам, переводчикам-специалистам, так и разработчикам, т.к. будут выявлены явления, которые мешают получить качественный перевод.

Создание системы АП, осуществляющей перевод, качество которого соответствует профессиональным потребностям специалистов и переводчиков, остается пока недостижимым по ряду причин: технических, кибернетических, лингвистических. Разработки, направленные на улучшение качества работы систем АП, ведутся постоянно: осуществляются попытки слияния систем автоматического перевода и систем переводческой памяти, ограничивается тематическая область, расширяются автоматические словари. Тем не менее, переводы, выполняемые современными системами АП, требуют редактирования, при этом для решения сложных случаев редактору

необходимо знание и языка перевода, и исходного языка, поскольку анализ исходного текста/ предложения приходится проводить заново.

Таким образом, развитие систем АП позволит оперативно осуществлять перевод информации и обрабатывать большие массивы документов в предельно сжатые сроки, т.е. удовлетворять основному требованию сегодняшней жизни: оперативной обработке огромных массивов информации при минимальных затратах.

Научная новизна исследования заключается в том, что впервые систематизированы ошибки и неточности, возникающие при автоматическом переводе текстов экономической тематики с помощью таких систем как ПРОМТ и Google Translate, а также разработаны рекомендации по их устранению и редактированию.

Теоретическая значимость исследования заключается в том, что:

1) Установлена необходимость и целесообразность анализа особенностей перевода различных тематик с помощью систем автоматического перевода, что способствует дальнейшему развитию теории машинного перевода и вносит вклад в переводоведение;

2) Раскрыта роль систем автоматического перевода как инструмента повышения эффективности работы переводчика;

3) Составлена классификация ошибок, возникающих в процессе автоматического перевода, и индикаторов переводимости, провоцирующих их появление в переводе.

Практическая значимость исследования состоит в разработке рекомендаций по вопросу о методах установления и устранения ошибок при использовании систем автоматического перевода в процессе перевода текстов экономической тематики (подъязык макро-, мегаэкономики), следование которым позволит ускорить процесс перевода.

Целью исследования является оценка пригодности автоматического перевода для перевода текстов предметной области макро- и мегаэкономики

с английского языка на русский и разработка рекомендаций по повышению эффективности при работе с системами автоматического перевода.

Для достижения указанной цели необходимо решить следующие **задачи**:

- Описать в чем заключается работа переводчика и от чего зависит эффективность его работы;
- Провести обзор современных вспомогательных инструментов, которые используются в процессе переводческой деятельности;
- Рассмотреть понятие автоматического перевода, его историю, типы систем;
- Выявить индикаторы переводимости, характерные для текстов подъязыка макро-, мегаэкономики;
- Составить рекомендации по работе с системами автоматического перевода для повышения эффективности деятельности переводчика.

Объектом исследования является лингвистическая структура английских экономических текстов (подъязык макро-, мегаэкономики).

Предметом исследования являются лингвистические явления экономических текстов, которые вызывают трудности при автоматическом переводе (индикаторы переводимости) в онлайн-системах ПРОМТ и Google Translate и разработка рекомендаций для предредактирования англоязычных текстов по экономике.

Материалом исследования послужил корпус экономических научных статей, взятых из журнала Economics: The Open-Access, Open-Assessment E-Journal за 2015–2018 года.

С учетом объекта, целей и задач в работе использовались следующие **методы исследования**: лингвистический анализ, эксперимент, сравнительно – сопоставительный, функциональный, содержательный анализы.

Теоретической основой исследования послужили работы следующих ученых: Е. В. Андреева, Л.Н. Беляева, Ю. Н. Марчук, С. О. Шереметьева, К. Буатэ, А. Bernth, С. Gdaniec, В. Jongejan, Т. А. Kuhn, U. Reuther, N. L. Underwood.

Основные результаты исследования отражены в научной статье «Экспериментальный анализ индикаторов переводимости экономических текстов», опубликованной в сборнике «Вестник ЮУрГУ», серия «Лингвистика».

Объем и структура работы.

В первой главе рассматриваются компетенции переводчика, необходимые для осуществления переводческой деятельности; освещаются этапы развития автоматического перевода, типы систем, основные преимущества и недостатки в процессе осуществления АП.

Во второй главе проводится экспериментальный анализ эффективности использования систем автоматического перевода ПРОМТ и Google Translate в ходе осуществления перевода текстов экономической тематики (подъязык макро-, мегаэкономики), выявляются закономерные ошибки, провоцирующие их индикаторы переводимости, разрабатываются правила контролируемого языка для повышения качества получаемого перевода с помощью систем АП.

В заключении подводятся итоги исследования и намечаются перспективы применения описанных теоретических положений на практике.

ГЛАВА 1 АСПЕКТЫ ПЕРЕВОДЧЕСКОЙ ДЕЯТЕЛЬНОСТИ

1.1 Профессиональная деятельность переводчика

Развитие связей между странами в эпоху глобализации было бы невозможно без знания иностранных языков. Владение иноязычной речью особенно актуально для установления партнерских отношений между предприятиями, деловыми структурами, общественными объединениями, учебными заведениями, некоммерческими организациями.

Большая роль в международном коммуникативном процессе принадлежит переводчикам.

Переводчик – специалист, занимающийся переводом, то есть созданием письменного или устного текста на определенном языке, эквивалентного письменному или устному тексту на другом языке.

Современный человек пользуется плодами трудов переводчика, когда читает инструкцию к стиральной машине иностранного производства, когда узнает по телевизору о том, что пишут сегодня во французских газетах, когда ему нужно провести переговоры с иностранным партнером, когда он берет в руки интересный роман. К этому можно добавить и труд переводчиков прошлых веков, потому что Библию и многие художественные произведения мы читаем и в старых переводах.

Труд переводчиков способствует открытости общества. Не случайно все диктаторские режимы ставили переводческую деятельность под жесточайший контроль, а современные «закрытые» государства (например, Иран) организуют гонения на переводчиков.

В наши дни самая большая потребность в переводчиках наблюдается в технических областях, в которых трудятся более 70% переводчиков в мире [1].

Профессиональный переводчик, осуществляющий перевод как письменно, так и устно, должен знать все нюансы переводческой деятельности «изнутри», исходя из типа текста, выбирать правильную

стратегию перевода, учитывать прагматический, лингвоэтнический и другие аспекты перевода, соблюдать переводческую этику.

Среди основных функциональных обязанностей переводчика можно выделить следующие: перевод текстов определенной тематики, общение с иностранными партнерами, ведение корреспонденции, осуществление устного перевода при встрече, синхронный перевод, подготовка аннотаций и рефератов иностранных источников.

Для осуществления профессиональной деятельности переводчик должен владеть переводческой компетенцией. Переводческая компетенция – это сложная, многоаспектная лингвокогнитивная категория, включающая профессиональные навыки и умения, позволяющие переводчику осуществлять акт межъязыковой и межкультурной коммуникации. Согласно В.Н. Комиссарову, переводческая компетенция включает в себя языковую, коммуникативную, текстообразующую, техническую компетенции и личностную характеристику переводчика.

Языковая компетенция заключается в знании всех аспектов владения языком, которые присущи любому носителю языка. Переводчик должен всегда помнить о системе, нормах и узусе языка, о его лексическом составе и грамматическом строе, о правилах употребления языковых единиц для построения речевых выражений, о территориальных, социальных и профессиональных различиях в использовании слов. Учитывая, что перевод – это акт межъязыковой коммуникации, переводчик должен обладать достаточной языковой компетенцией в области не одного, а двух языков, причем как в рецептивном, так и в продуктивном плане.

Важное место в профессиональной компетенции переводчика занимает коммуникативная компетенция, то есть способность переводчика подбирать наиболее подходящее речевое поведение, которое соответствует целям и ситуации общения, формировать правильные выводы из содержания исходного текста на основе фоновых знаний. Переводчик не просто интерпретирует смысл текста, но и при необходимости вводит недостающую

для понимания информацию в само высказывание или выносит ее в примечания и сноски. Коммуникативная компетенция предполагает «умение проецировать на высказывания в тексте оригинала инференциальные возможности рецепторов перевода» [13].

Текстообразующую компетенцию составляет умение создавать тексты различного типа в соответствии с коммуникативной задачей и ситуацией общения, обеспечивать надлежащую структуру текста, использовать языковые единицы текста по правилам построения речевых единиц в языке, оценивать место и соотношение отдельных частей текста и воспринимать текст как связное речевое целое. Кроме того, переводчик должен знать различия в общей стратегии построения текста в двух языках.

Профессиональная компетенция переводчика включает некоторые личностные характеристики, которые помогают ему успешно выполнять свои профессиональные функции. Перевод представляет собой сложный вид умственной деятельности, осуществление которой предполагает особую психическую организацию. Поэтому, с одной стороны, переводчик должен проявлять гибкость, пластичность, быстро переключать внимание, переходить от одного языка к другому, от одной культуры к другой, от одной коммуникативной ситуации к другой. С другой стороны, от переводчика требуется умение сосредотачиваться, мобилизовать ресурсы своей памяти, весь свой интеллектуальный и эмоциональный потенциал.

Переводчику приходится переводить тексты разнообразной тематики, принадлежащих к различным сферам употребления, которые зачастую требуют специальных знаний: знания терминологии, стилистических особенностей текстов. Поэтому от переводчика требуется эрудированность, начитанность, широта интересов, непрерывное саморазвитие, умение находить и воспринимать новую информацию, продуктивно использовать справочники, словари и другие источники информации.

Также стоит отметить морально-этический компонент профессиональной компетенции переводчика, ведь он несет ответственность за качество

выполненной им работы, моральный и материальный ущерб, который может быть вызван недобросовестно выполненным переводом. Переводческая деятельность полностью основана на доверии рецепторов перевода к результатам работы переводчика. Оправдать это доверие переводчик может лишь за счет большой собранности, работоспособности и исключения любых элементов несерьезного, халатного отношения к делу.

Скромность, тактичность, умение держаться, хранить тайну – обязательная часть профессионального поведения переводчика, поскольку переводчик работает в самых разных ситуациях, от работы с секретными политическими и коммерческими документами до присутствия при закрытых переговорах на самом высоком уровне.

Техническая компетенция включает специфические знания, умения и навыки, которые необходимы для эффективного выполнения перевода. В данную категорию входят, как умения работать с компьютером и современной техникой, применять технические средства в процессе перевода, так и знания о стратегиях перевода, переводческих приемах и трансформациях.

В настоящее время переводоведение занимается исследованием и разработкой методов и способов перевода, которые помогли бы переводчикам справиться с требованиями качественного перевода, выполняемым в наиболее короткие сроки. Существует целый спектр возможностей, к которым может прибегать переводчик: начиная от тех, которые требуют от него исключительно логических умозаключений, и, заканчивая теми, которыми можно пользоваться только при наличии специального оборудования.

К первой категории относится, например, теория уровней эквивалентности В. Н. Комиссарова [13], теория закономерных соответствий Я. И. Рецкера [19] и другие теории и модели перевода, характеризующиеся тем, что требуют от переводчика в первую очередь наличие крепкого фундамента теоретических знаний. Ещё одной их особенностью можно

назвать элемент сравнительного анализа, присущий алгоритму перевода в рамках вышеприведённых теорий, так как они учат обращающегося к ним отличать «норму» от отклонения, схематично показывая и то, и другое.

Иногда то, что в одной переводческой ситуации, может и должно называться «нормой», в другой ситуации может приниматься за отклонение. К примеру, В. Г. Гак [8], известный лингвист и преподаватель, разработал свою типологию номинаций, в которой разбил все рассматриваемые им явления по категориям – или номинациям. Одним из признаков его типологии является разделение на субъект и объект номинации и соответствующее отношение к типам, которые выделяются по способу обозначения. Учитываются такие факторы, как отношение говорящего к предмету, отношение говорящего к адресату, общие условия речевого акта и отношение адресата к предмету. Такие параметры имеют большое значение в типологии В. Г. Гака. Соответственно, в том случае, если посредник перевода будет обращаться к теоретическому материалу, такому, как данная типология, он будет обязан рассматривать переводимые им предложения с точки зрения отношений, таким же образом разделять «субъекта» и «объекта» и изучать общие условия речевого акта. Данное условие необходимо для того, чтобы переводчик имел полное представление о том отрывке текста, который ему приходится переводить, и, на основании сделанных выводов, имел возможность выполнить перевод правильно.

На первых порах, данная схема может представлять существенную трудность для переводчика, однако, с практикой, она становится таким же естественным и привычным ритуалом, как использование в процессе перевода словаря или лингвистического справочника [24].

Вторая категория методов, с помощью которых переводчики могут повысить эффективность работы, включает в себя специальные инструменты и программы. Более подробно она рассмотрена в следующем пункте.

1.2 Инструменты повышения эффективности работы переводчика

Прежде всего, эффективность – это достижение каких-либо определенных результатов с минимально возможными затратами. Применительно к профессиональной деятельности переводчика повышение эффективности его работы означает сокращение времени на обработку определенного объема текста при сохранении или улучшении качества перевода.

Скорость перевода и последующего редактирования текста зависят от способностей и опыта переводчика. Если улучшение качества перевода осуществимо при постоянном самосовершенствовании, то на скорость работы это влияет незначительно. Особенно остро встает проблема рационального использования времени при освоении новых тем, при переходе к работе в новых областях, когда темп работы значительно снижается из-за нехватки опыта и необходимости постоянного обращения к словарям и источникам информации. Для повышения эффективности необходимо провести оптимизацию работы.

Процесс оптимизации работы начинается с анализа всех этапов выполнения перевода и поиска наиболее проблемных факторов, которые вызывают трудности.

Процесс перевода состоит из трех основных этапов:

- 1) Предпереводческий анализ текста;
- 2) Аналитический вариативный поиск;
- 3) Анализ результатов перевода.

Предпереводческий анализ – это анализ исходного текста, предваряющий создание переводного текста и направленный на выявление доминант перевода, т. е. тех черт, которые входят в инвариант при переводе. И. С. Алексеева выделяет следующие доминанты: тип текста, реципиенты текста перевода, источник текста оригинала, конвенции текста оригинала и текста перевода, виды информации в тексте оригинала (когнитивная, оперативная, эмоциональная, эстетическая), коммуникативное задание текста оригинала/перевода [1]. Главной трудностью данного этапа является

отсутствие экстралингвистических сведений, поэтому всем аспектам предпереводческого анализа должно предшествовать полноценное понимание содержания текста и знакомство с той предметной областью, к которой он относится. При подготовке к переводу переводчик, во-первых, изучает другие тексты по той же тематике, во-вторых, подбирает справочную литературу и специальные словари и, в-третьих, с помощью создания плана или краткой аннотации к тексту уточняет для себя его содержание.

На этапе аналитического вариативного поиска главной трудностью является определение единиц перевода в подлиннике – наименьших семантических единств, подлежащих переводу, и нахождение для них функционально эквивалентных соответствий, а именно независимых от контекста однозначных эквивалентных соответствий, контекстуально-зависимых вариантных соответствий или трансформационных соответствий. К тому же, на данном этапе переводчик сталкивается с синтаксическими трудностями, вызванными различным строением языков. Среди них можно выделить: порядок слов в предложении, разные синтаксические функции частей речи (передача функций артикля, герундия, конструкций сложного дополнения и сложного подлежащего, неличные формы глагола и т.д.), несовпадение грамматических категорий, отсутствие запятых, выделяющих смысловые группы. Как правило, грамматическая форма и синтаксическая конструкция воспринимаются при переводе как неразрывное целое с их лексическим содержанием. Если соответствующая грамматическая форма или синтаксическая конструкция отсутствует в ПЯ, то переводчик прибегает к различным трансформациям.

Для подбора лексических единиц переводчик использует различные источники информации, которые помогают ему произвести перевод на уровне морфем, слов, словосочетаний. Данные источники носят название рабочие источники информации переводчика. Они подразделяются на два типа: общие источники информации и специальные источники информации.

С общими источниками информации работают все переводчики, в то время как специальными источниками информации преимущественно пользуются переводчики при осуществлении перевода текстов узкой специализации, например, технический перевод, юридический перевод и т. п.

Общие источники информации представлены словарями общего назначения и общими энциклопедиями. Словари общего назначения имеют следующее деление:

- Двухязычные словари:
 - Двухязычные неспециальные словари;
 - Двухязычные фразеологические словари;
- Одноязычные словари:
 - Толковые словари:
 - Толковые словари русского языка;
 - Толковые словари иностранного языка;
 - Словари иностранных слов;
 - Вспомогательные словари:
 - Словари синонимов;
 - Словари антонимов;
 - Словари омонимов;
 - Орфографические словари;
 - Словари сокращений;
 - Словари пословиц и поговорок и др.;
- Энциклопедические словари.

Специальные источники информации помогают переводчику справиться с проблемой нехватки знаний в какой-либо профессиональной области. К ним относят:

- Специальные словари:
 - Двухязычные специальные словари:
 - Политехнические словари;
 - Отраслевые словари;

- Вспомогательные словари;
- Одноязычные специальные словари;
- Специальные энциклопедии:
 - Политехнические энциклопедии;
 - Отраслевые энциклопедии;
- Справочники;
- Специальная литература (глоссарии, тезаурусы);
- Прочие источники информации:
 - Предшествующий опыт;
 - Консультации специалистов.

В современном мире большую роль играет Интернет, поскольку в нем содержится огромное количество информации и онлайн-программ, которые могут существенно облегчить процесс «ручного» перевода. Среди электронных источников можно выделить: электронные словари (общие и специализированные), онлайн-словари, параллельные и квазипараллельные корпуса текстов, память переводчика, системы автоматического перевода.

Электронные словари отличаются от онлайн-словарей тем, что они являются оцифрованной версией бумажных словарей. Как правило, они существуют в форматах .pdf, .djvu. Онлайн-словари снабжены пользовательским интерфейсом и основаны на базе данных разной глубины. Помимо перевода лексической единицы, они могут давать морфологическую информацию, а также перевод лексемы в разных контекстах и учитывать предметную область [22]. Наиболее популярными онлайн-словарями являются Abbyu Lingvo [54] и MultiTran [55]. Они относятся к словарям открытого типа, что значит, что пользователи могут вносить в базу данных свои переводы и эквиваленты, увеличивая покрываемость лексикона. При этом главным недостатком является наличие многочисленных вариантов перевода лексемы, что не всегда помогает при выборе нужного значения слова. Классификация терминов по предметным областям не всегда содержит требуемый пользователем раздел и поэтому бывает достаточно сложно

принимать решения относительно предложенных вариантов. Также не стоит забывать, что словари открытого типа пополняются не только профессиональными переводчиками, но и обычными пользователями, поэтому переводы терминов могут быть некорректны.

Параллельные тексты (корпусы тестов) – это тексты на разных языках одного и того же содержания. Квазипараллельные тексты имеют сходное, но не полностью совпадающее содержание, т. е. относятся к одной теме. Данные тексты позволяют переводчику правильно определить значение лексемы в контексте для поиска корректного эквивалента. Недостатком использования параллельных и квазипараллельных текстов является дефицит корпусов, поскольку сложно, а иногда и совсем невозможно найти тексты конкретной узкоспециальной области. При этом для эффективного применения корпусного метода определения межъязыковых эквивалентов часто требуется вспомогательный инструментарий [22].

К электронным источникам также относятся инструменты автоматизированного перевода, или же САТ-инструменты. САТ расшифровывается как Computer-Aided (Assisted) Translation – «перевод с помощью компьютера». Данные технологии не стоит отождествлять с автоматическим переводом, в котором система сама осуществляет перевод всего текста. Программы автоматизированного перевода – это программное обеспечение, которое лишь помогает переводчику сделать процесс перевода более эффективным на основе ранее полученных от пользователя переводов похожих текстов.

САТ-инструменты включают в себя различные лингвистические ресурсы, которые облегчают труд переводчиков с однотипными текстами, содержащими стандартные фразы и предложения – технические, юридические и медицинские термины, описания товаров и многое другое. Одними из самых распространенных ресурсов являются базы Translation Memory – базы памяти переводов, которые содержат ранее переведенные сегменты текста (словосочетания и предложения). Они создаются и

пополняются на основе пар параллельных текстов. Другой важный ресурс – глоссарии, которые содержат термины и понятия, аббревиатуры, принятые в той или иной области.

Полностью автоматизированный перевод всего текста или его отдельных единиц способны обеспечить только системы автоматического перевода. Они увеличивают скорость работы переводчика, поскольку ему предлагается готовый переведенный текст. Пользователь может оставить предложенный перевод или отредактировать его при наличии неправильного перевода.

На последнем, третьем этапе перевода осуществляется анализ результатов перевода: переводчик контролирует свои собственные решения и производит необходимую правку, а именно сверку полноты текста, орфографическую коррекцию, стилистическую коррекцию (редактирование), коррекцию опечаток, проверку полноты воплощения стратегии перевода в полученном тексте. Трудность, которая может появиться на данном этапе перевода, связана с человеческим фактором, окружающей обстановкой. Переводчик может не заметить собственных ошибок, неточностей и искажений, что приведет к снижению адекватности перевода. Избежать этой трудности помогут специальные программы, проверяющие орфографические, грамматические и стилистические ошибки.

На сегодняшний день только человек может создать грамотный адекватный перевод, но при этом он может использовать все возможности, которые ему предоставляет современное программное обеспечение. Поиск перевода слов при помощи электронных словарей намного быстрее, чем при использовании обычных источников информации. Перевод на уровне всего текста способны выполнять лишь системы автоматического перевода. Однако, качество получаемого перевода не совершенно, поскольку оно зависит от лингвистических явлений текста, с которыми системы АП не всегда могут успешно справиться. Учитывая популярность систем АП среди пользователей, рассмотрение данного вопроса особенно актуально.

1.2.1 История развития автоматического перевода

Начало развития автоматического перевода было положено в конце 1940-х – начале 1950-х годов. Теоретической основой этого периода выступал взгляд на язык как кодовую систему. Пионерами автоматического перевода были математики и инженеры. Описания их первых опытов, связанных с использованием только что появившихся ЭВМ для решения криптографических задач, были опубликованы в США в конце 1940-х годов. Датой рождения автоматического перевода как исследовательской области обычно считают март 1947; именно тогда специалист по криптографии Уоррен Уивер в своем письме Норберту Винеру впервые поставил задачу автоматического перевода, сравнив ее с задачей дешифровки [18]. У. Уивер писал: «У меня перед глазами текст, написанный по-русски, но я собираюсь сделать вид, что на самом деле он написан по-английски и закодирован при помощи довольно странных знаков. Всё, что мне нужно – это взломать код, чтобы извлечь информацию, заключенную в тексте» [25].

Аналогия между переводом и дешифрованием была естественной в контексте послевоенной эпохи, если учитывать успехи, которых достигла криптография в годы Второй мировой войны.

Идеи Уоррена Уивера легли в основу подхода к АП и были основаны на концепции *interlingva*. Согласно данной концепции стадия передачи информации делится на два этапа:

- Исходное предложение переводится на язык-посредник (созданный на базе упрощенного английского языка);
- Результат этого перевода представляется средствами выходного языка.

В 1949 году У. Уивер составил меморандум, в котором теоретически обосновал принципиальную возможность создания систем автоматического перевода.

В 1952 году в Массачусетском технологическом университете состоялась первая конференция по автоматическому переводу, а в 1954 году в Нью-Йорке была представлена первая система АП – IBM Mark II, разработанная

компанией IBM совместно с Джоржтаунским университетом (это событие вошло в историю как Джорджтаунский эксперимент). Данная система, использовавшая словарь в 250 единиц и грамматику из 6 правил, осуществила перевод 49 заранее отобранных предложений с русского на английский язык.

В том же 1954 году первый эксперимент по автоматическому переводу был осуществлен в СССР в Институте точной механики и вычислительной техники Академии наук СССР. И.К. Бельская отвечала за лингвистическую часть эксперимента, и Д.Ю. Панов был ответствен за программную часть. Первый промышленно пригодный алгоритм автоматического перевода и система автоматического перевода с английского языка на русский на универсальной вычислительной машине были разработаны коллективом под руководством Ю. А. Моторина. После этого работы начались во многих информационных институтах, научных и учебных организациях страны.

Идея автоматического перевода стимулировала развитие исследований в теоретическом и прикладном языкознании во всем мире. Появились теории формальных грамматик, большое внимание стало уделяться моделированию языка и отдельных его аспектов, языковой и мыслительной деятельности, вопросам языковой формы и количественных распределений лингвистических явлений. Возникли новые направления лингвистической науки – вычислительная, математическая, инженерная, статистическая, алгоритмическая лингвистика и ряд других отраслей прикладного и теоретического языкознания. В течение 1950-х годов в учебных центрах многих стран мира были открыты отделения прикладной лингвистики и автоматического перевода [18].

В 1959 г. философ Й. Бар-Хиллел (Yohoshua Bar-Hillel) выступил с утверждением, что высококачественный полностью автоматический АП (FANQMT) не может быть достигнут в принципе. В качестве примера он привел проблему нахождения правильного перевода для слова *pen* в следующем контексте: *John was looking for his toy box. Finally he found it. The*

box was in the pen. John was very happy. (Джон искал свою игрушечную коробку. Наконец он её нашёл. Коробка была в манеже. Джон был очень счастлив.). *Pen* в данном случае должно переводиться не как "ручка" (инструмент для письма), а как "детский манеж" (*play-pen*). Выбор того или иного перевода в этом случае и в ряде других обусловлен знанием внеязыковой действительности, а это знание слишком обширно и разнообразно, чтобы вводить его в компьютер. Однако Бар-Хиллел не отрицал идею АП как таковую, считая перспективным направлением разработку машинных систем, ориентированных на использование их человеком-переводчиком.

Это выступление самым неблагоприятным образом отразилось на развитии АП в США. Правительство США было озабочено тем, что прогресс был несоразмерен со значительными затратами на разработку проекта. В 1966 г. специально созданная Национальной Академией наук комиссия ALPAC (Automatic Language Processing Advisory Committee), основываясь, в том числе и на выводах Бар-Хиллела, пришла к заключению, что автоматический перевод нерентабелен: было установлено, что автоматический перевод дороже, медленнее и менее точный в сравнении с переводом, выполненным человеком, и несмотря на большие капиталовложения, в ближайшее время автоматический перевод не достигнет того же качества, что и перевод, выполненный человеком.

Однако отчет советовал развитие вспомогательных программ для переводчиков – например, автоматических словарей – и поддержку исследований в области компьютерной лингвистики.

Публикация отчета в большей степени повлияла на исследования автоматического перевода в США и в гораздо меньшей в СССР и Великобритании. По крайней мере, в США такого рода исследования были остановлены на целое десятилетие. В Канаде, Франции и Германии исследования все-таки продолжались. В США главными исключениями стали основатели компаний Systran.

Новый подъем исследований в области АП начался в 1970-х годах. Он был связан с развитием вычислительной техники и серьезными достижениями в области компьютерного моделирования интеллектуальной деятельности. Главной целью исследователей стало развитие «реалистических» систем АП, предполагавших участие человека на различных стадиях процесса перевода. Системы АП становятся незаменимыми помощниками переводчиков, способствующие экономии времени и человеческих ресурсов.

Одной из новых разработок стала технология ТМ (translation memory), работающая по принципу накопления: в процессе перевода сохраняется исходный сегмент (предложение) и его перевод, в результате чего образуется лингвистическая база данных; если идентичный или подобный исходному сегмент обнаруживается во вновь переводимом тексте, он отображается вместе с переводом и указанием совпадения в процентах. Затем переводчик принимает решение (редактировать, отклонить или принять перевод), результат которого сохраняется системой. В конечном итоге не нужно дважды переводить одно и то же предложение. В настоящее время разработчиком известной коммерческой системы, основанной на технологии ТМ, является система TRADOS (основана в 1984 г.).

Дальнейшее развитие систем АП обуславливается развитием информационных технологий, широким использованием сети Интернет. Появляются новые подходы, основанные на использовании нейронных сетей, концепции коннекционизма, статистических методах [50].

В ноябре 2016 года компания Google представила систему нейронного машинного перевода (GNMT – Google Neural Machine Translation), которая использует искусственную нейронную сеть для повышения беглости и точности перевода в Google Переводчике. Благодаря большой сквозной структуре, система со временем учится создавать более качественные и естественные переводы. GNMT способна переводить предложения целиком, а не по частям [48].

В настоящее время можно выделить два основных стимула к развитию работ по автоматическому переводу. Первый – собственно научный; он определяется комплексностью и сложностью компьютерного моделирования перевода. Как вид языковой деятельности перевод затрагивает все уровни языка – от распознавания графем (и фонем при переводе устной речи) до передачи смысла высказывания и текста. Кроме того, для перевода характерна обратная связь и возможность сразу проверить теоретическую гипотезу об устройстве тех или иных языковых уровней и эффективности предлагаемых алгоритмов.

Второй стимул – социальный, и обусловлен он возрастающей ролью самой практики перевода в современном мире как необходимого условия обеспечения межкультурной коммуникации, объем которой возрастает с каждым годом. Другие способы преодоления языковых барьеров на пути коммуникации – разработка или принятие единого языка, а также изучение иностранных языков – не могут сравниться с переводом по эффективности. С этой точки зрения можно утверждать, что альтернативы переводу нет, так что разработка качественных и высокопроизводительных систем автоматического перевода способствует разрешению важнейших социально-коммуникативных задач.

Кристиан Буатэ выделяет три вида современного автоматического перевода:

1) «информативный» автоматический перевод – грубый автоматический перевод, пословный, достаточный, чтобы ознакомиться с текстом на незнакомом языке;

2) «профессиональный» автоматический перевод, качество которого сравнимо с качеством «человеческого» перевода и при небольшом редактировании является полностью удовлетворительным. Выгоден для больших объемов текста (свыше 10 тыс. страниц в год);

3) «персональный» автоматический перевод. Авторы текстов заранее избавляют их от неоднозначностей и работают в режиме диалога с

компьютером. Таким образом можно значительно повысить качество перевода [15].

Среди преимуществ использования автоматического перевода можно выделить:

- Экономия времени: за короткий срок можно получить перевод большого текста;
- Доступ к услуге: программа-переводчик всегда находится под рукой;
- Перевод с любого языка: загруженные двуязычные словари позволяют осуществлять перевод для любой языковой пары в любом направлении;
- Низкая стоимость перевода (по сравнению с профессиональным переводом);
- Конфиденциальность: введенная информация не распространится.

К недостаткам автоматического перевода можно отнести то, что:

- Необходимо осуществлять редактирование текста для получения качественного перевода;
- В автоматическом переводе не всегда учитываются правила грамматики, языковые приемы, стилистика текста;
- Система АП не всегда способна выбрать правильный перевод, тем самым снижается удобочитаемость текста;
- Если в словарной базе отсутствует какое-то слово, оно не переводится.

Современный автоматический перевод следует отличать от использования других вспомогательных инструментов в помощь переводчику. Различные электронные словари, глоссарии, тезариусы помогают подбирать нужный переводной эквивалент. В то время как термин «автоматический перевод» подразумевает, что большая часть работы по переводу и отысканию переводных соответствий принадлежит системе АП. Переводчик лишь осуществляет контроль и исправляет ошибки при необходимости. Таким образом, автоматический перевод – это процесс

передачи содержания текста с одного естественного языка на другой с помощью системы автоматического перевода.

В своем докладе в Пенанге на международной конференции по актуальным проблемам вычислительной лингвистики в Малайзии, Кристиан Буатэ сформулировал двенадцать проблем современного автоматического перевода, которые подразделяются на 4 класса: концептуальные проблемы; проблемы архитектуры; инженерные проблемы; технические проблемы.

К концептуальным проблемам АП Кристиан Буатэ относит:

- Разрешение неоднозначности разного рода с помощью «неоднозначного программирования», предусматривающего разные возможности разрешения неоднозначности;

- Обеспечение точного переводного соответствия между абстрактной структурой и текстом, поскольку конкретное высказывание всегда отличается от его абстрактной модели;

- Создание эффективных средств анализа из чистых формализмов, добавляя к ним процедурные или эвристические (приближенные, эмпирические) знания.

Среди проблем архитектуры наиболее важные относятся к неполноте описания языка в целом:

- Конструирование и использование грамматических формализмов, которые позволили бы описать разные варианты языка («микроязыки» или «подязыки») в рамках единого языка, приспособив конкретную систему автоматического перевода для данного языка во всей его общности к конкретному подязыку посредством настройки;

- Обучение на основе текстов: статистически-основанные САП могут быть эффективными для перевода коротких сообщений. Создание массивов параллельных текстов также может быть выгодным для разработки таких систем. Что касается грамматических правил, то их гораздо труднее выводить автоматически из сравнения параллельных текстов;

- «Степень понимания»: обучение системы АП истолкованию высказывания частично по содержанию, частично по структуре, а затем гармонично их соединять.

Технические проблемы заключаются в организации многоязычных лексических баз данных:

- Организация больших многоязычных, многотеоретических баз лексических данных, на основе которых можно было бы строить автоматически конкретные словари для автоматического перевода;

- Работать с целым текстом, а не с отдельными фразами (эта задача не является технической, поскольку ее решение требует алгоритмического установления анафорических связей, обращения к фрагментам предыдущего и последующего текстов для разрешения лексической многозначности и т.п.).

Следующий класс проблем К. Буатэ называет инженерными, хотя они также являются достаточно сложными и существенно затрагивают лингвистику:

- Кодирование многоязычных текстов: обучение системы АП распознавать буквы, буквенные сочетания, обозначающие разные звуки;

- Увязать «архитектуру» с конкретными целями применения;

- Измерение прогресса в автоматическом переводе.

Таким образом, несмотря на то, что некоторые из этих проблем получили приемлемые для практики решения, улучшение качества автоматического перевода требует учета перечисленных теоретических и практических моментов, сформулированных достаточно давно. Причина такого положения заключается, прежде всего, в том, что перевод является сложной проблемой с самыми разнообразными аспектами, а естественные языки настолько разнообразны и неоднозначны, что невозможно формализовать всё многообразие их явлений.

1.2.2 Типы систем автоматического перевода

В настоящее время различают три типа систем автоматического перевода:

- Системы на основе грамматических правил (Rule-Based Machine Translation, RBMT);

- Статистические системы (Statistical Machine Translation, SMT);
- Нейронный автоматический перевод;

Системы автоматического перевода на основе правил (RBMT) были первыми коммерческими системами. Эти системы позволили уйти от пословного перевода. На основе разработанных лингвистических правил системы автоматического перевода стали способны менять порядок слов в предложении, выбирать значение в зависимости от контекста и т.д. [29].

Системы автоматического перевода на основе правил производят анализ текста, который используется в процессе перевода. Перевод производится на основе встроенных словарей для данной языковой пары, а так же грамматик, охватывающих семантические, морфологические, синтаксические закономерности обоих языков. На основе всех этих данных исходный текст последовательно, предложение за предложением, преобразуется в текст на требуемом языке. Основной принцип работы таких систем – связь структур исходного и конечного текстов. Наиболее ярким представителем данного типа является система автоматического перевода ПРОМТ.

Преимуществами систем на основе грамматических правил являются:

- Грамматическая и синтаксическая точность;
- Стабильность и предсказуемость результата;
- Возможность настройки на специфическую предметную область.

К недостаткам систем на основе грамматических правил относят:

- Необходимость создания, поддержки и обновления лингвистических баз данных;
- Трудоемкость создания такой системы;
- Высокая стоимость [12].

Статистические системы является подвидом метода, основанного на корпусах текста (corpus-based machine translation, CBMT) [40]. При своей

работе они используют статистический анализ. В систему загружается двуязычный корпус текстов (содержащий большое количество текста на исходном языке и его «ручной» перевод на требуемый язык), после чего система анализирует статистику межъязыковых соответствий, синтаксических конструкций и т.д. Предполагается, что слова с наибольшей частотностью дают лучший перевод [42]. Чем больший словарь внутри языковой пары и чем точнее он составлен, тем лучше результат статистического автоматического перевода. С каждым новым переведенным текстом улучшается качество последующих переводов. Представителями данного типа до недавнего времени являлись Google Translate [51] и Yandex [53].

К преимуществам статистических систем относятся:

- Быстрота настройки;
- Легкость добавления новых направлений перевода.

Среди недостатков наиболее значительными являются:

- Наличие многочисленных грамматических ошибок;
- Нестабильность перевода;
- «Дефицит» параллельных корпусов [12].

Нейронный автоматический перевод – это относительно новый подход к решению проблемы автоматического перевода, получивший широкое распространение в последние годы. Несмотря на то, что нейронный перевод похож на статистический, поскольку также использует анализ параллельных данных и формирует на основе этого анализа определённые зависимости и закономерности, он имеет совершенно иные принципы работы. Функционирование данного подхода основано на использовании нейронных сетей, вычислительных моделей, по своей структуре напоминающих строение человеческого мозга, в которых сигнал распространяется через последовательные слои элементов, имитирующих нейроны.

На сегодняшний день большинство систем НАП имеют примерно схожую архитектуру: кодер, декодер, и механизм внимания [49]. На стороне

исходного предложения кодер преобразует каждое отдельное слово в контекстный вектор, который декодер, в свою очередь, превращает в слово целевого предложения. Механизм внимания позволяет системе сфокусироваться на отдельных частях исходного предложения и контролировать точность перевода.

Сервисы перевода компаний Google, Яндекс, Microsoft и ПРОМТ уже используют нейронный перевод. Google использует нейронный автоматический перевод Google (GNMT) вместо ранее используемых статистических методов. Майкрософт использует похожую технологию для перевода речи (в том числе в Майкрософт Переводчике и Skype Переводчике). Яндекс.Переводчик имеет гибридную модель: свой вариант перевода предлагает и статистическая модель, и нейросеть. После этого технология CatBoost, в основе которой лежит машинное обучение, будет выбирать лучший из полученных результатов.

Основная причина, по которой системы НАП не получали широкого распространения до недавнего времени – это относительно низкие показатели точности перевода в сравнении с самыми совершенными системами статистического автоматического перевода. По большей части недостатки систем НАП определяются архитектурой нейронных сетей, используемых в данных системах. К примеру, одна из наиболее общих проблем – это неспособность к переводу редко встречающихся слов [32, 45, 49]. Данная проблема обусловлена тем, что системы НАП отбирают только n -е количество самых частотных слов из корпусов текстов, опуская при этом менее частотные. Это может вызывать пропуски в переведенном тексте. Невозможность включения в словарь системы всех возможных переводов конкретного слова продиктована необходимостью излишне не усложнять систему.

Второй ключевой недостаток – это относительно низкая скорость самообучения данных систем, особенно при работе с большими объемами

корпусов текстов [43, 49]. Данная особенность, в основном, обуславливается ограничениями, существующими на уровне аппаратного обеспечения.

С одной стороны, нейронная сеть, при должном количестве входных данных способна выдать практически идеальный или близкий к таковому перевод, поскольку она не просто обучается, она «понимает» принципы, по которым строится перевод. С другой стороны, во-первых, она требует куда больше вычислительных мощностей для своего обучения, и, во-вторых, в условиях недостатка параллельных данных, коих требует на порядок больше, чем статистическая сеть, нейронная сеть выдаст перевод крайне низкого качества, то есть нейронные переводчики ещё более требовательны к объёму входных данных, чем статистические системы.

Несмотря на очевидные перспективы нейронных сетей в качестве основной технологии автоматического перевода в будущем, на текущий момент недостаток в первую очередь должного объёма данных в различных предметных областях не позволяет говорить о безоговорочном преимуществе нейронных сетей во всех сферах применения автоматического перевода. Поэтому статистические системы, требующие меньший объём данных для обучения, и системы на основе правил, дающие стабильный и точный результат в узких областях, по-прежнему сохраняют актуальность.










1.2.3 Обзор пользовательских возможностей в системах ПРОМТ и Google Translate

В настоящее время системы автоматического перевода ПРОМТ и Google Translate наиболее популярны среди пользователей интернета.

Онлайн-переводчик ПРОМТ – сервис перевода, предназначенный для быстрого перевода небольших фрагментов текста с одного языка на другой. За один раз разрешается переводить тексты объемом не более 3 000 символов, включая пробелы и знаки конца параграфа (10 000 символов для зарегистрированных пользователей) [52]. В ПРОМТ доступен перевод на 19 языках. ПРОМТ позволяет выбрать тематику, к которой относится текст:

автомобили, гаджеты, гуманитарные науки, деловая переписка, естественные науки, знакомства, изучение языков, кино, компьютеры и технологии, медицина и фармацевтика и т.д.

Основные функции ПРОМТ:

-  Вставить текст – перенести текст в поле ввода из буфера обмена;
-  Копировать текст – скопировать перевод в буфер обмена;
-  Виртуальная клавиатура – вставить символы, отсутствующие на клавиатуре;
-  Проверить орфографию – исключить орфографические ошибки во введенном тексте;
-  Прослушать – узнать, как произносится искомое слово и все его переводы;
-  Поделиться – отправить перевод с помощью другого приложения или предложенной ссылки;
-  Оценить – оценить перевод (хорошо/плохо);
-  Распечатать – распечатать перевод введенного слова, словосочетания и примеры употребления;
-  Улучшить – предложить свой вариант перевода.
- Интерредактирование (описание функции в п. 1.2.4);
- Перевод веб-страницы: онлайн-переводчик ПРОМТ позволяет переводить текстовое содержимое сайтов в Интернете с одного языка на другой, включая гиперссылки и подписи к картинкам, с полным сохранением HTML-форматирования на странице. Для этого нужно ввести полный адрес веб-страницы для перевода в поле для ввода и результат перевода автоматически отобразится в новой вкладке (см. рис. 1.1).

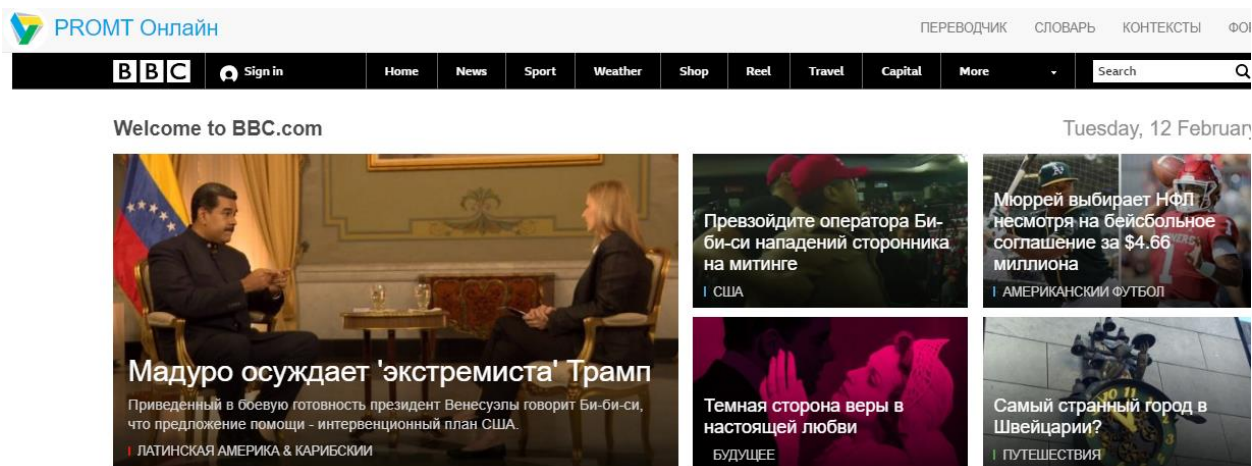



Рисунок 1.1 – Результат перевода веб-страницы в системе ПРОМТ

Для зарегистрированных пользователей предоставляется возможность сохранять выполненные переводы лексем, словосочетаний, предложений с помощью функции «Мои переводы» .

Сервис перевода можно использовать не только как переводчик связных и сложных предложений, но и как справочный словарь, в котором можно посмотреть переводы отдельных слов и словосочетаний и полезную грамматическую информацию, которую можно получить с помощью команды «Словарь» верхнего меню. В словаре доступны 7 языков.

В поле «Перевод» помимо собственно перевода входного слова в соответствующей форме будет представлена полная словарная справка о входном слове или о том слове из словаря, от которого оно образовано (см. рис. 1.2):

1. основные грамматические формы слова (в зависимости от части речи);
2. все переводы, сгруппированные по частям речи;
3. краткая грамматическая информация и комментарии;
4. транскрипция (для английского языка);
5. возможность прослушать произношение исходного слова или перевода.

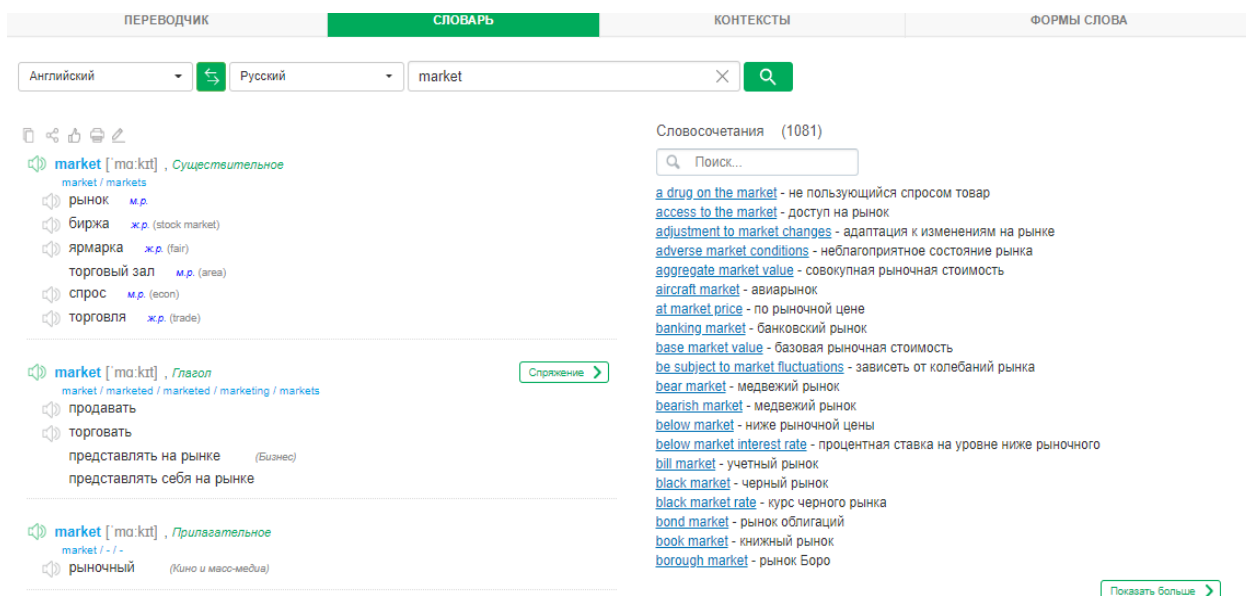


Рисунок 1.2 – Словарная справка в системе ПРОМТ

Если слово или словосочетание является существительным, прилагательным или глаголом, то можно посмотреть список всех грамматических форм данного слова с помощью команд «Склонение» (для существительных и прилагательных), «Спряжение» (для глаголов). Откроется страница «Формы слова» (см. рис. 1.3). В данном разделе можно посмотреть варианты словоизменения введенного слова: род и склонение существительных и прилагательных, степени сравнения прилагательных, спряжение глаголов.

Сервис перевода ПРОМТ обладает обширной базой примеров для различных направлений перевода и различных тематик, содержащей тексты и их переводы, выполненные человеком. Посмотреть примеры употребления введенного слова или выражения можно в разделе «Контексты» (см. рис. 1.4).

Английский market Словоизменение

market, Существительное market, Глагол market, Прилагательное

market ['mɑ:kt] - Глагол

ВСЕ ФОРМЫ | INDEFINITE | CONTINUOUS | PERFECT | PERFECT CONTINUOUS | INFINITIVES | PARTICIPLES

Present Indefinite I market you market he/she/it markets we market you market they market	Present Perfect I have marketed you have marketed he/she/it has marketed we have marketed you have marketed they have marketed	Present Continuous I am marketing you are marketing he/she/it is marketing we are marketing you are marketing they are marketing	Present Perfect Continuous I have been marketing you have been marketing he/she/it has been marketing we have been marketing you have been marketing they have been marketing
Past Indefinite I marketed you marketed he/she/it marketed we marketed you marketed they marketed	Past Continuous I was marketing you were marketing he/she/it was marketing we were marketing you were marketing they were marketing	Past Perfect I had marketed you had marketed he/she/it had marketed we had marketed you had marketed they had marketed	Past Perfect Continuous I had been marketing you had been marketing he/she/it had been marketing we had been marketing you had been marketing they had been marketing
Future Indefinite I will market you will market he/she/it will market	Future Continuous I will be marketing you will be marketing he/she/it will be marketing	Future Perfect I will have marketed you will have marketed he/she/it will have marketed	Future Perfect Continuous I will have been marketing you will have been marketing he/she/it will have been marketing

Рисунок 1.3 – Вкладка «Формы слова» в системе ПРОМТ

ПЕРЕВОДЧИК СЛОВАРЬ **КОНТЕКСТЫ** ФОРМЫ СЛОВА

Английский Русский market

Поиск...

Переводы: все²⁶⁴⁵³ Другие переводы²⁸³⁵ рынка³⁸⁴⁷ рынки³¹⁵⁶ рынок³⁰³⁶ рынке²⁷⁷⁸ рынков²²⁸⁸ рынках¹⁸⁸⁵ рынкам⁸⁸⁵ рынк

Тематика: все²⁶⁴⁵³ Политика и экономика¹⁴⁴³⁰ Общая лексика⁸²¹¹ Финансы⁴⁹¹⁸ Software³¹⁷ Разговорная речь²⁸³ Интернет²¹⁸ Наука

Some bankers have tried to play down the affair by saying the vast and highly liquid foreign exchange market is almost impossible to manipulate, but senior traders are saying this is not necessarily true.

Некоторые банкиры пытались приуменьшить серьезность ситуации, говоря, что обширным и высоколиквидным валютным рынком практически невозможно манипулировать, но старшие трейдеры говорят, что это не всегда верно.

A further and important conclusion is that structural change in the contracting organisation may invoke structural change in the market.

Еще одно важное следствие состоит в том, что структурные изменения в организации-партнере могут вызвать структурные изменения на рынке.

My thesis is that the degree of dynamism in a nation's economy hinges on its development of some key economic institutions - company law and corporate governance, the population's preparation for business life, the development of financial instruments such as the stock market and so forth.








Мой тезис заключается в том, что уровень динамизма в экономике страны тесно связан с развитием некоторых основных экономических институтов, таких как законодательство о компаниях, корпоративное управление, подготовленность населения к деловой жизни, развитие финансовых инструментов, как, например,

Рисунок 1.4 – Раздел «Контексты» в системе ПРОМТ

Таким образом, перечисленные функции системы автоматического перевода ПРОМТ дают расширенные возможности переводчику для повышения эффективности его работы.

Google Переводчик (англ. Google Translate) [51] – веб-служба компании Google, предназначенная для автоматического перевода отдельных слов, выражений, предложений и текста или веб-страницы на другой язык. Переводить можно тексты объемом до 5 000 символов. Материалы большего размера следует переводить по частям или можно воспользоваться отличительной возможностью Google Translate и загрузить документ целиком. Поддерживаемые форматы: DOC, DOCX, ODF, PDF, PPT, PPTX, PS, RTF, TXT, XLS, XLSX. При переводе Google Translate формирует отдельную страницу, на которой пытается сохранить исходное форматирование документа. Некоторые особенности форматирования исходного документа могут быть не отражены в переводе, например, в таблицах столбцы текста сбиваются и наезжают друг на друга. Однако это не влияет на качество перевода. В онлайн-переводчике Google Translate доступно больше языков, чем в ПРОМТ – 103 языка.

Среди основных функций Google Translate можно выделить:

-  Перевести – выполнить перевод для выбранной языковой пары;
-  Прослушать – узнать, как произносится перевод;
-  Включить голосовой ввод – ввода текста при помощи обычной речи, передаваемой посредством микрофона;
-  Поделиться – отправить перевод с помощью другого приложения;
-  Копировать – скопировать перевод в буфер обмена;
-  Виртуальная клавиатура – ввод текста с помощью мыши;
- Выбрать способ ввода: рукописный ввод, клавиатура с раскладкой DVORAK, виртуальная клавиатура;
-  Предложить исправление – внести свой вариант перевода вместо неправильного. Прежде чем вариант перевода, добавленный пользователем, будет добавлен в сервис, он должен пройти проверку в Сообществе Google Переводчика (см. рис. 1.5) (URL: <https://translate.google.com/community>).

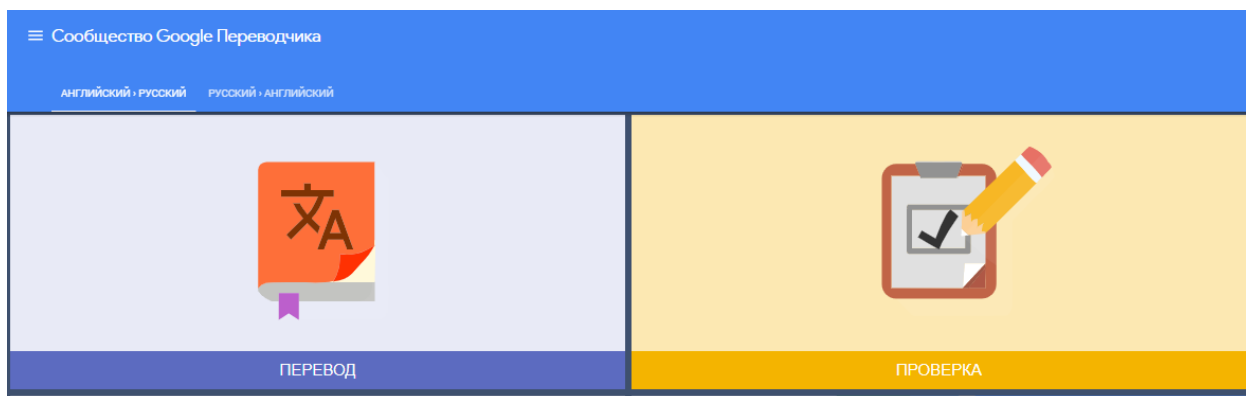


Рисунок 1.5 – Сообщество Google Translate

В разделе «Перевод» предлагаются слова или фразы, которые следует перевести на выбранный язык. В разделе «Проверка», предлагаются слова и фразы с несколькими вариантами перевода. Участникам сообщества нужно пометить их как правильные или неправильные. Участники сообщества помогают не только улучшать качество переводов, но и добавлять в этот сервис новые языки. Для этого требуется только переводить с этого языка и на него, а также проверять готовые переводы.

- Интерредактирование (описание функции в п. 1.2.4);
- Перевод веб-страницы: для осуществления перевода сайта необходимо выбрать языковую пару, вставить URL страницы в поле ввода текста. Google Translate автоматически переключится в режим перевода сайтов и отобразит указанную страницу уже в переведенном виде. Возможность переходить по ссылкам этого сайта сохранится. Новые страницы будут открываться уже переведенными (см. рис. 1.6).

К дополнительным возможностям платформы Google Translate относятся: формирование «Разговорника», поиск определений и синонимов, просмотр истории переводов.

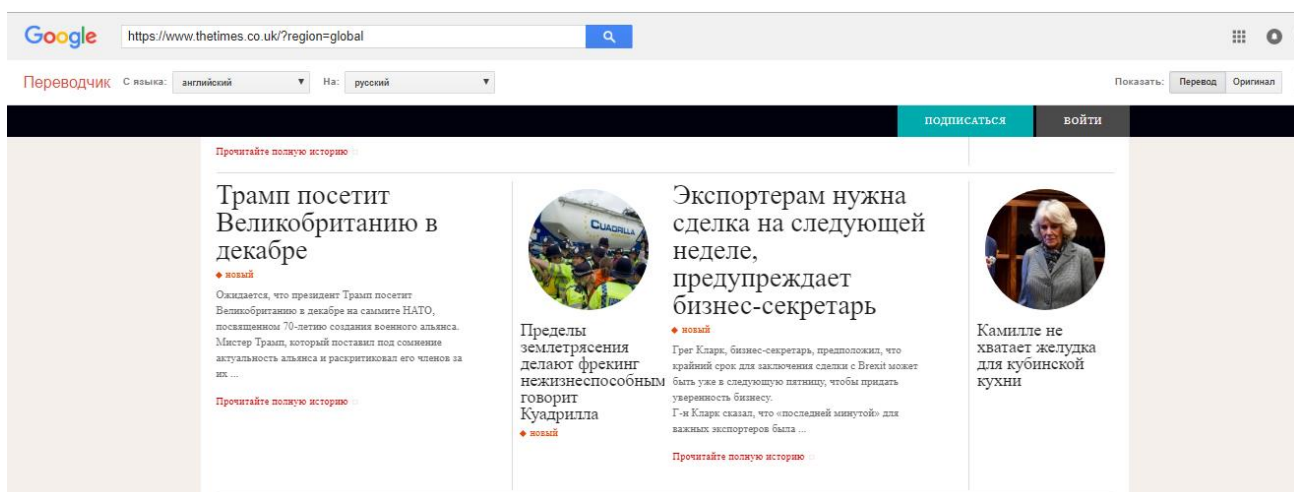


Рисунок 1.6 – Результат перевода веб-страницы в системе Google Translate

Как и ПРОМТ, Google Translate для зарегистрированных пользователей предоставляет возможность вести пользовательский словарь с помощью функции «Разговорник». Система API Google Translate сохраняет перевод введенного слова, фразы или предложения в «Разговорник». Для этого нужно нажать на иконку ☆ в поле с выполненным переводом (см. рис. 1.7).

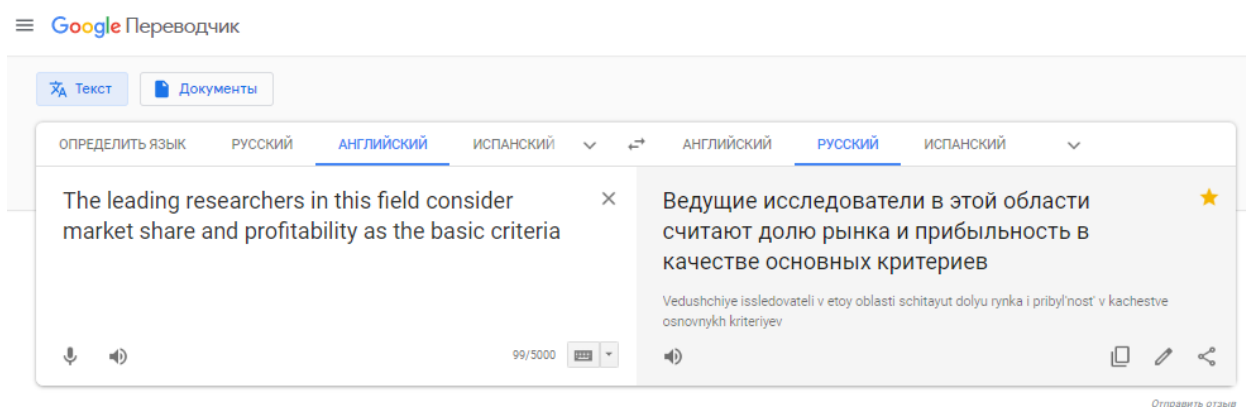



Рисунок 1.7 – Добавление перевода предложения в «Разговорник» Google Translate

Чтобы открыть составленный разговорник, нужно нажать на кнопку  Сохранено в нижней части экрана. В разговорнике можно осуществлять поиск по лексемам, сортировать переводы по языковым парам. Данная функция может

быть полезна при изучении иностранных языков, при составлении терминологического словаря.

С пользовательской точки зрения, недостатками разговорника является невозможность добавления в «Разговорник» отдельного слова, нужной фразы из выполненного перевода текста, а также невозможность интеграции собственного варианта перевода для исправления перевода, выполненного системой автоматического перевода Google Translate.

Хотя основная функция Google Translate – перевод, сервис можно использовать и в качестве толкового словаря или словаря синонимов. Для этого достаточно ввести в поле перевода не предложение или текст, а отдельное слово. Под полем исходного текста будет отображено определение этого слова и примеры его использования. Под полем перевода – варианты перевода, упорядоченные по частоте употребления, с огромным количеством синонимов для каждого из возможных контекстов (см. рис. 1.8).

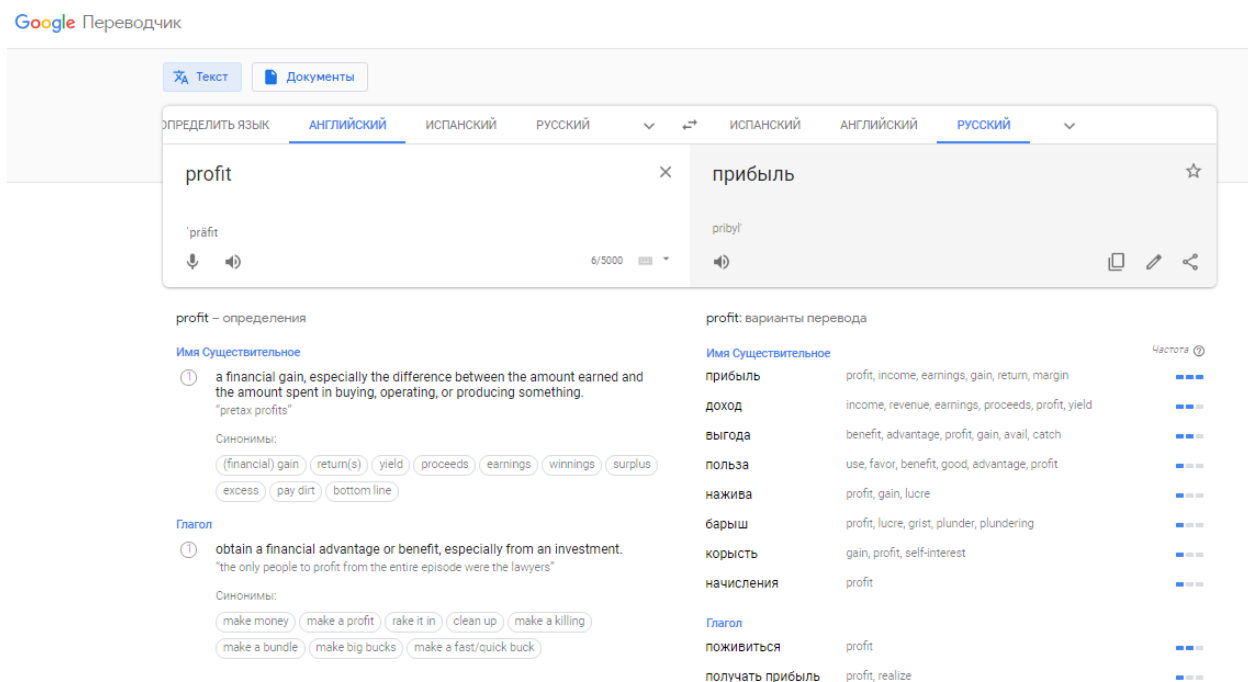


Рисунок 1.8 – Функция словаря системы Google Translate

При необходимости можно найти ранее выполненные переводы с помощью функции «История переводов».

Перечисленные возможности систем автоматического перевода ПРОМТ и Google Translate представлены в табл. 1.1.

Таблица 1.1 – Сравнительная таблица возможностей систем АП

Критерий	ПРОМТ	Google Translate
Объем вводимого текста	< 3 000 символов (10 000 – для зарегистрированных пользователей)	< 5 000
Перевод документов	-	+
Количество доступных языков	19	103
Выбор тематики текста	+	-
Проверка орфографии	+	+
Возможность улучшить перевод	+	+
Возможность добавления новых языков, отсутствующих в системе АП	-	+
Перевод сайтов	+	+
Ведение пользовательского словаря	+	+
Толковый словарь	-	+
Синонимы, примеры употребления	+	+
Грамматическая информация	+	-
Просмотр истории переводов	-	+

Обе системы автоматического перевода обладают схожими возможностями, поэтому выбор более предпочтительной системы АП для работы остается за пользователем.

1.2.3 Оценка качества автоматического перевода

В связи с тем, что качество автоматического перевода неидеально и требует редактирования, возник вопрос об оценке текстов, полученных с помощью систем АП.

Под качеством перевода понимается оценочный показатель сделанного перевода, определяемый точностью, сжатостью, ясностью, литературностью переведенного текста. Для художественных переводов их качество оценивается также такими показателями, как эмоциональная и художественная выразительность текста и т.п. [16].

Стоит отметить, что качество перевода зависит от стиля и тематики исходного текста. Автоматический перевод художественных текстов практически всегда оказывается неудовлетворительного качества, поскольку в текстах данного стиля используется огромное количество разнообразных средств художественной выразительности: от эпитетов, игры слов до ритмико-синтаксического построения фразы.

Тем не менее для технических документов при наличии специализированных машинных словарей и некоторой настройке системы на особенности того или иного типа текстов возможно получение перевода приемлемого качества, который нуждается лишь в небольшой редакторской корректировке. Чем более формализован стиль исходного документа, тем большего качества перевода можно ожидать. Самых лучших результатов при использовании автоматического перевода можно достичь для текстов, написанных в техническом (различные описания и инструкции) и официально-деловом стиле.

В настоящее время существует два основных направления оценки качества автоматического перевода: ручная и автоматическая оценки.

В основе ручного метода лежит оценка качества автоматического перевода человека по определенным критериям и определенной шкале, например, грамотность и адекватность перевода с помощью человеческих сил [21].

Оценка качества автоматического перевода ручным способом является очень дорогим, субъективным и трудоемким процессом [34]. Кроме того, у эксперта во время процесса оценки может появиться усталость, сонливость, испорченное настроение, что негативно скажется на его суждениях. Тем не

менее, автоматические метрики не могут заменить оценку, выполненную человеком.

Макото Нагао предложил шкалу оценки автоматического перевода на основе лингвостилистического анализа, от пяти до одного балла.

1 балл – смысл предложения понятен и не возникает никаких вопросов, грамматика, словоупотребление и стиль соответствуют общей структуре текста и не требуют постредктирования

2 балла – смысл предложения понятен, но возникают большие проблемы с грамматикой, словоупотреблением и стилем.

3 бала – общий смысл предложения понятен, но смысл некоторых его частей вызывает сомнение из-за неправильного грамматического строя.

4 балла – присутствуют ошибки словоупотребления и стилистики, требуется обращение к оригиналу.

5 баллов – в предложении имеется большое количество грамматических, словоупотребительных и стилистических ошибок, смысл предложения с трудом можно понять после внимательного изучения [35].

Каллисон-Берч с соавторами считают, что при оценке текста можно использовать рейтинговую систему, т.е. эксперты должны оценить качество перевода по ряду критериев и составить рейтинг результатов разных систем автоматического перевода, расположив их по порядку от 1 до 5, где 1 – хуже всего и 5 – лучше всего [28].

Еще одним способом оценки качества перевода является анализ ошибок, предложенный Виларом с соавторами [42]. Под анализом ошибок подразумевается идентификация и классификация ошибок в переводном тексте, полученном при автоматическом переводе.

Таким образом, ручная оценка качества слишком дорога, требует существенных временных затрат. Также недостатком этого метода оценки является то, что человек не может быть объективен, так как уровень понимания текста реципиентом, его мнения относительно текста всегда зависит от индивидуальных факторов.

Автоматические системы оценки машинного перевода основываются на методе N-грамм, на сравнении строк (string matching).

Метод N-грамм основан на использовании вероятности появления цепочки букв N-го порядка (N-грамм) в анализируемых текстах. Он был введен Макото Нагао и Шинсукэ Мори [35]. N-грамма – это последовательность из n-слов или знаков. По своему составу N-граммы могут делиться на униграммы (одно слово или один знак), биграммы (два), триграммы (четыре) и так далее.

Наиболее популярные метрики:

- U.S National Institute of Standards & Technology (NIST) metric;
- Bilingual Evaluation Understudy (BLEU);
- Metric for Evaluation of Translation with Explicit Ordering (METEOR).
- Translation Error Rate (TER).

Первые три метрики основываются на методе N-грамм, последняя – на основе сравнения строк.

Метрика BLEU была разработана сотрудниками компании IBM и в настоящее время является одним из наиболее широко используемых автоматизированных методов оценки качества автоматического перевода. Данная метрика отличается простотой и не зависит от языка перевода. Она подсчитывает пересечения по N-граммам (от 1 до 4), усредняет точность для разных N-грамм и начисляет штраф за краткость перевода (вместо полноты), который учитывает разную длину переводов. Показатель BLEU колеблется от 0 до 1, и обычно отображается в процентах. Чем больше перевод коррелирует с человеческим переводом, тем ближе оценка к единице (100%). Иными словами, метрика BLEU измеряет, сколько слов совпадает в автоматическом переводе и эталонном переводе, при этом фразам с правильной последовательностью слов присваиваются более высокие баллы. Значения ниже 15% указывают на то, что механизм автоматического перевода не может обеспечить качественный перевод. Для получения качественных переводов требуется высокий уровень постредактирования.

Значения более 30% показывают то, что переводы должны быть понятны без каких-либо проблем. Значения выше 50% говорят о хорошем качестве перевода.

Метрика NIST является улучшенной версией метрики BLEU. Она использует среднее арифметическое и среднее геометрическое N-грамм совпадений, чтобы оценить правильный выбор лексики. За правильный перевод редких N-грамм присуждает больший вес. Метрика NIST отображает выходные значения в диапазоне от 0 до 15, при этом большее значение указывает на лучшее качество перевода [37].

Представленная Институтом языковых технологий университета Карнеги-Меллона, метрика METEOR также является модифицированной версией метрики BLEU. METEOR отдает предпочтение полноте перевода, а не штрафу за краткость как метрика BLEU. Кроме того, данная метрика использует модули стемминга (выделение основы слова) и синонимичности, что позволяет преодолеть проблему поверхностной оценки схожести. Также METEOR оценивает правильность порядка слов с помощью фрагментационного штрафа. Он выражается в количестве пропусков или прыжков между словами, которые необходимо совершить, чтобы восстановить верный порядок слов в переводе. Этот показатель, как и BLEU, дает оценки в диапазоне от 0 до 100.

Наиболее современной метрикой является TER. Она оценивает минимальное количество человеческих исправлений, которые необходимо внести в перевод, чтобы полностью соответствовать эталонному переводу с точки зрения восприятия содержания. Среди правок, вносимых в полученный перевод можно выделить: замена слов или переформулирование фраз, вставка и удаление и т.д. В отличие от других метрик, более низкое значение TER указывает на большее сходство с эталонным переводом. Его баллы варьируются от 0 до 100 [26].

Рассмотрев ручную и автоматическую оценки, можно прийти к выводу, что метод автоматической оценки машинного перевода более объективен,

так как он позволяет оценить перевод с помощью формул, и тем самым исключить субъективность. С другой стороны, автоматические метрики оценки не могут дать полностью адекватный результат, так как при оценке с помощью N-грамм, не ставится задача понимания семантики текста, что и ведет к некоторой неточности такой оценки.

1.2.4 Пути повышения качества автоматического перевода

Главная цель систем автоматического перевода – выполнение высококачественного преобразования исходного языка в эквивалентный текст переводного языка.

Однако, перевод, выполненный системой автоматического перевода, далек от совершенства и требует редактирования. Иногда к переводным текстам не предъявляется высоких требований точности перевода, конкретного стиля изложения; достаточным является «донесение» до реципиента общего смысла, содержания. В таком случае «сырой» переводной текст является допустимым.

В случаях, когда переводной текст должен удовлетворять высоким критериям качества перевода (точность, адекватность, понятность и т.д.), редактирование является обязательным.

Существует несколько форм организации взаимодействия системы АП и человека в процессе перевода: постредактирование, предредактирование, интерредактирование.

Постредактирование – это редакция текста, полученного на выходе системы автоматического перевода, с целью устранения переводческих ошибок [23]. В зависимости от цели использования переведенного текста выделяют следующие типы постредактирования:

- Идеальное постредактирование – текст редактируется до уровня качественного перевода, выполненного профессиональным переводчиком;
- Быстрое или «грязное» постредактирование – полученный текст редактируется до уровня «понимаемости» перевода с возможными

стилистическими и даже грамматическими ошибками, если они не искажают смысл текста;

- Целевое постредактирование – редактированию подлежат отдельные фрагменты текста, содержащие интересующую заказчика информацию.

Для быстрого постредактирования характерно проведение как можно меньшего числа операций над текстом, среди которых можно выделить: исправление наиболее очевидных опечаток и грамматических ошибок, изменение сложных предложений, фиксирование ошибок системы автоматического перевода, удаление ненужных или альтернативных вариантов перевода, создания глоссария без углубленной проработки терминов.

Идеальное постредактирование включает в себя: проверку соответствия терминологии; сбор информации, связанной с текстом перевода; синтаксические изменения в соответствии с правилами языка перевода; работа над стилистикой текста; перевод и адаптация культурологических явлений (фразеологизмы, идиомы и др.); попытка добиться полного соответствия с оригинальным текстом; выполнение форматирования, в соответствии с оригиналом текста; исправление всех грамматических, пунктуационных и орфографических ошибок.

Интерредактирование – взаимодействие человека и системы автоматического перевода в процессе осуществления перевода, в котором человек разрешает трудные случаи, неопределенности, например, в случае многозначных слов.

Например, системы АП Google Translate и ПРОМТ предоставляют доступ ко всем лексическим значениям выбранного слова, которые распределены от наиболее часто используемых до редких вариантов перевода. Данная функция представлена на рис. 1.9, 1.10.

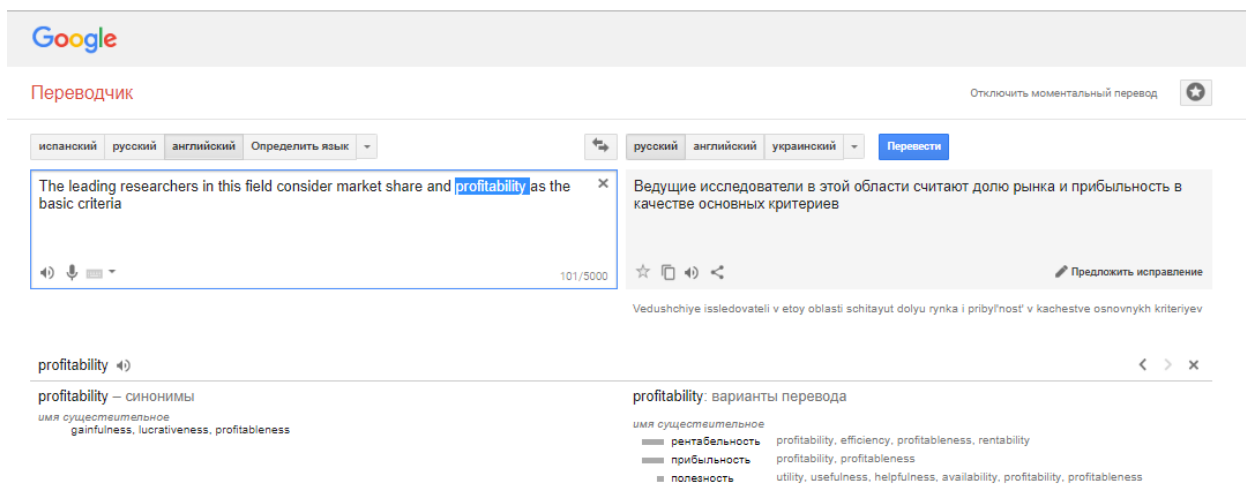


Рисунок 1.9 – Интерредактирование в Google Translate

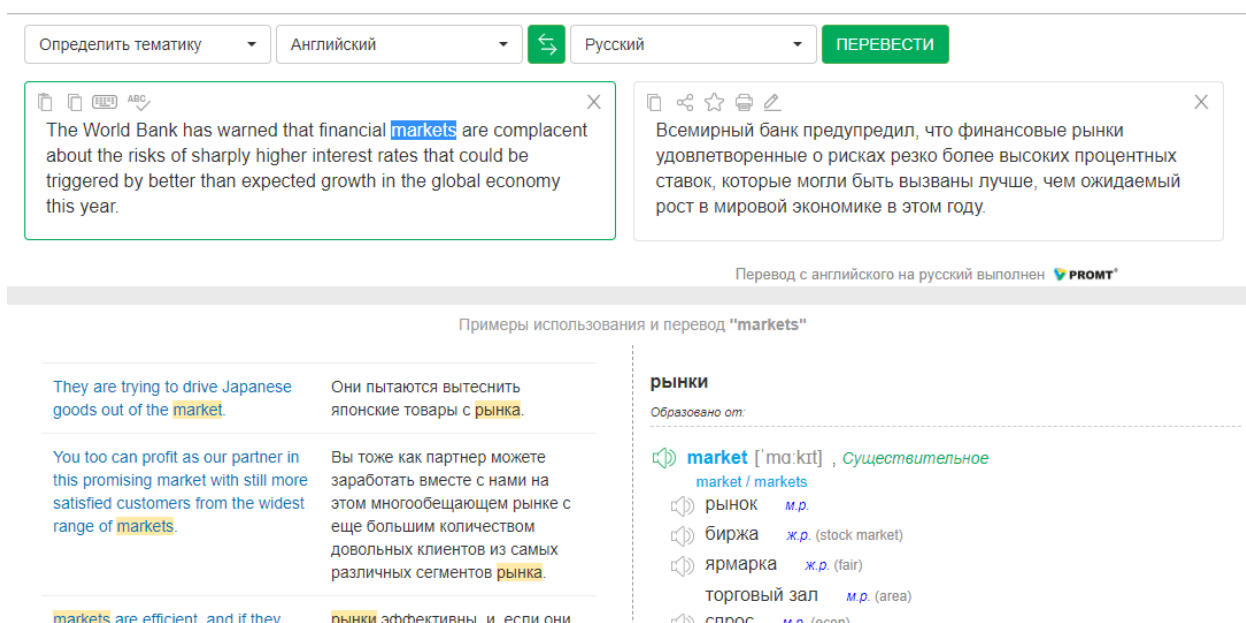


Рисунок 1.10 – Интерредактирование в ПРОМТ

При обнаружении опечаток или других орфографических ошибок в исходном тексте система АП Google Translate предлагает исправленный вариант вводимого текста, а переводчик принимает решение, какой из них является более корректным. Данное взаимодействие переводчика и системы АП представлено на рис. 1.11.

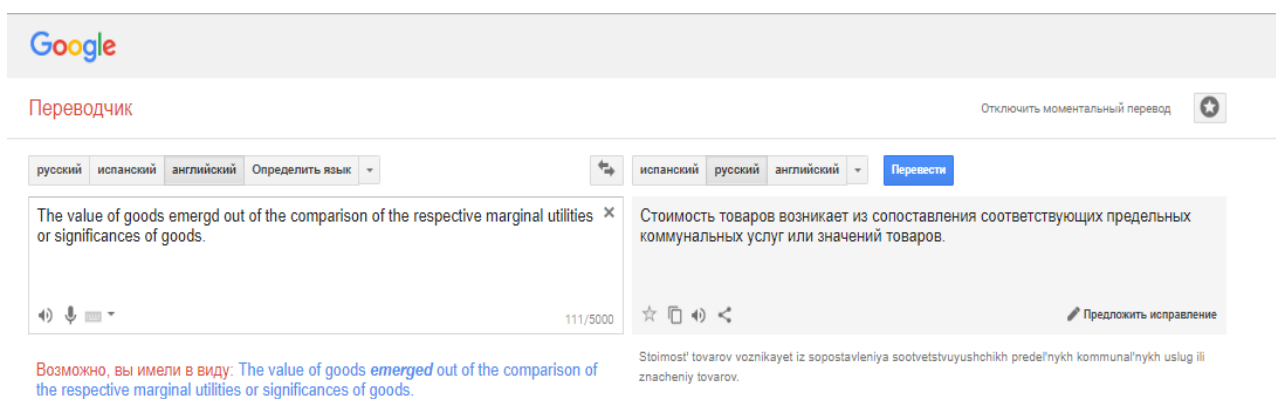


Рисунок 1.11 – Взаимодействие человека и системы АП Google Translate

Предредактирование – это преобразование фрагментов текста перед его вводом в систему автоматического перевода с целью повышения качества автоматического перевода.

Переводимость текста, или же возможность адекватного перевода текста на другой язык, зависит от качества входного текста. Уровень пригодности текста для автоматического перевода определяется с помощью индикаторов переводимости.

Индикаторы переводимости – это графические, лексические и синтаксические элементы текста, провоцирующие ошибки при автоматическом переводе. Чем меньше индикаторов переводимости в тексте, тем лучше качество автоматического перевода.

Среди индикаторов переводимости выделяют два основных класса: универсальные и специфические.

Универсальные индикаторы переводимости – это явления в исходном тексте, вызывающие ошибки при автоматическом переводе, вне зависимости от пары языков, направления перевода и конкретной системы автоматического перевода [23].

К универсальным индикаторам переводимости относятся следующие лингвистические явления в тексте:

- Большая длина предложения;

- Семантическая многозначность лексем;
- Лексическая многозначность лексем;
- Синтаксическая многозначность лексем;
- Пассивный залог;
- Дистантное расположение зависимых слов;
- Координация – синтаксическая связь главных членов двусоставного (двухкомпонентного) предложения – подлежащего и сказуемого – с полным или частичным формальным употреблением словоформ, занимающих позиции этих членов;

- Идиомы.

Специфические индикаторы переводимости зависят от конкретных пар языков, направления перевода и используемой системы перевода [23].

Для повышения уровня переводимости текста необходимо проводить работы по устранению индикаторов переводимости до ввода текста в систему автоматического перевода. Повышение пригодности текста для автоматического перевода достигается с помощью:

- Предредактирования;
- Правил контролируемого языка.

Процесс предредактирования состоит из двух этапов:

1. Обнаружение фрагментов входного текста, содержащих индикаторы переводимости;
2. Переформулирование обнаруженных фрагментов по правилам контролируемого языка и исключение индикаторов переводимости.

Контролируемый язык – это подвид естественного языка, который получен с помощью ограничений в использовании грамматики, терминологии и речевых оборотов посредством регламентирующих правил для того, чтобы снизить или искоренить его многозначность и сложность [31, 33].

К контролируемому языку предъявляются следующие требования:

- Он должен быть основан на естественном языке;
- Главное отличие контролируемого языка от естественного языка заключается в определенных ограничениях лексики, синтаксиса и/или семантики;
- Должен сохранять признаки языка-основы так, что люди говорящие на языке-основе могли бы правильно понять текст, созданный на контролируемом естественном языке;
- Контролируемый язык – созданный язык, что означает, что он не является продуктом естественных процессов.

Контролируемые языки не стоит путать с искусственными языками, такими как Эсперанто или Бейсик-инглиш, поскольку они не обладают вышеперечисленными характеристиками, например, не имеют языка-основы, а также не являются продуктом естественных процессов, поскольку полностью созданы человеком.

В общем виде правила контролируемого языка совпадают с правилами редактирования текста, но формулируются в виде предписаний использовать или не использовать определенные языковые средства.

Универсальные правила контролируемого языка для создания текстов можно сформулировать следующим образом:

- Не использовать многозначные слова или выражения, подбирать однозначные синонимы;
- Не использовать фраз и предложений с синтаксической омонимией;
- Не употреблять длинные и сложные предложения, разбивать на более простые предложения;
- Использовать прямой порядок слов в предложении;
- Избегать инфинитивных и деепричастных конструкций;
- Избегать междометий, вводных фраз и идиоматических выражений;
- Вместо местоимений использовать полнозначные слова;
- Не использовать пассивный залог;

- Не использовать длинные определительные конструкции в препозиции;
- Обстоятельственные слова в предложении ставить так, чтобы порядок слов предложения по возможности совпадал с порядком слов языка перевода [23].

Среди преимуществ использования контролируемого языка можно выделить:

- Повышение удобочитаемости и переводимости текстов;
- Точность автоматического перевода текста на контролируемом языке выше, чем на неупрощенном;
- Отсутствие многозначности в текстах.

К недостаткам использования контролируемого языка относится:

- Значительные временные затраты на приспособление к контролируемому языку;
- Разработка контролируемого языка требует много времени и средств;
- Язык теряет свою эстетичность.

Для улучшения качества АП компания ПРОМТ рекомендует придерживаться следующим шагов:

- Определение тематики подлежащего переводу текста и подбор подходящих (специализированных) словарей;
- Получение списка наиболее частых терминов-оборотов в переводимых текстах, определение их переводных эквивалентов и занесение в словарь;
- Получение списка наиболее часто встречающихся слов и проверка их текущего перевода системой АП;
- Установка требуемых правил перевода.

В результате данных шагов объем редактирования автоматического перевода переводчиком снизится с 80% до 50% за счет настройки предметной области, с 50% до 30% за счет настройки терминологии. Как следствие повысится общее качество полученного перевода.

Для оценки применимости автоматического перевода выделяются такие критерии как:

- Эквивалентность по числу компонентов предложения и эквивалентность перевода по каждому компоненту;
- Знание предметной области;
- Устойчивость соблюдения терминологии;
- Связность: слова-связки, служебные части речи, окончания.

Выводы по оценке применимости АП представлены в табл. 1.2.

Таблица 1.2 – Выводы по оценки применимости АП

Критерий	Оценка	Вывод
Эквивалентность	Нет	АП придется использовать с особой осторожностью, проверяя каждое слово
	Да	Можно обратить внимание на более значимые моменты
Предметная область	Нет	Воспользоваться специализированным словарем, найти консультантов
	Да	Сосредоточиться на более точной передаче смысла второстепенных понятий
Терминология	Нет	Выделить время на проверку каждого термина по глоссарию
	Да	Найти более удачные решения по сочетаемости, тщательно выстроить логические связи между предложениями
Связность	Нет	Воспользоваться справочником по грамматике и тезаурусом
	Да	Разнообразить текст синонимами, проработать стиль

В ходе постредактирования АП можно выделить следующие стратегии:

- Полное редактирование с проверкой каждого компонента текста по исходному фрагменту (объем постредактирования – 100%);
- Использование автоматического перевода исключительно в качестве подсказки по терминологии (объем постредактирования – 82%);
- Использование АП в качестве грамматического подстрочника с подбором синонимичных выражений (объем постредактирования – 65%);
- Чтение АП и исправление вызывающих сомнения мест с обращением к исходному тексту (объем постредактирования – 34%);
- Чтение АП и исправление очевидных погрешностей без обращения к исходному тексту (объем постредактирования – 14%) [52].

Основные критерии для выбора стратегии редактирования:

- знание предметной области;
- степень владения исходным и целевым языком;
- сроки и предполагаемый объем постредактирования;
- опыт работы с данной системой АП;
- степень доверия автоматическому переводу;
- наличие технических средств для анализа внутренней структуры АП.

Подводя итог вышесказанному, стоит отметить, что уже сегодня автоматический перевод в сочетании с дальнейшим редактированием текста становится достаточно конкурентоспособной переводческой работой на основании стратегических направленностей и рекомендаций, разработанных специалистами в данной области.

Для получения качественного перевода с помощью систем автоматического перевода необходимо придерживаться правил контролируемого языка, как на этапе создания текста, так и в процессе его предредактирования, что позволит сделать текст более пригодным для автоматического перевода. Системы АП совершают разные ошибки, поэтому практическая часть нашей работы посвящена повышению качества перевода

систем автоматического перевода ПРОМТ и Google Translate с помощью предредактирования на материале англоязычных экономических текстов, принадлежащих к подобласти макроэкономики и мегаэкономики.

Выводы по главе 1

Повышение эффективности работы переводчика означает сокращение времени на обработку определенного объема текста при сохранении или улучшении качества перевода.

Существует ряд вспомогательных инструментов в помощь переводчику: электронные словари, глоссарии, системы переводческой памяти, корпуса параллельных и квазипараллельных текстов и др. Эти инструменты чрезвычайно популярны в переводческом сообществе, но не позволяют получить перевод всего текста при минимальных временных затратах – для этих целей предназначены системы АП.

Автоматический перевод – это процесс передачи содержания текста с одного естественного языка на другой с помощью системы автоматического перевода. Качество автоматического перевода неидеально и требует взаимодействия с человеком в форме предредактирования, постредактирования или интерредактирования. Для получения качественного перевода с помощью систем автоматического перевода необходимо придерживаться правил контролируемого языка, как на этапе создания текста, так и в процессе его предредактирования, что позволит сделать текст более пригодным для автоматического перевода. Умение работать с системами автоматического перевода, знание специфичных индикаторов переводимости выбранной языковой пары и выбранной системы АП, может помочь пользователям получить перевод более высокого качества, сократить время, затрачиваемое на осуществление перевода.

ГЛАВА 2 АНАЛИЗ ЭФФЕКТИВНОСТИ ИСПОЛЬЗОВАНИЯ СИСТЕМ АВТОМАТИЧЕСКОГО ПЕРЕВОДА ПРОМТ И GOOGLE TRANSLATE НА МАТЕРИАЛЕ ЭКОНОМИЧЕСКИХ ТЕКСТОВ

2.1 Важность предметной области экономики

В настоящее время экономика приобретает абсолютную значимость и ценность для существования и жизнеобеспечения не только каждого из государств мира, но и всей цивилизации. Глобализация рынка, интеграция международных организаций, появление транснациональных компаний, развитие информационных технологий обуславливают бурный рост и расширение экономической сферы.

Важность экономики возрастает с каждым днем, ведь именно экономическая сфера оказывает значительное влияние на все аспекты жизни общества, являясь его фундаментом. Этот вывод вытекает из следующих положений:

- Существование современного общества невозможно без постоянного производства материальных благ;
- Сложившееся разделение труда, распределение рабочих мест определяют развитие социальной структуры общества;
- Экономические отношения активно влияют на политическую сферу, определяя вектор развития страны;
- В процессе производства создаются необходимые материальные условия для развития духовной жизни общества (строительство библиотек, школ, театров, оборудование для издания книг, газет и т. п.).

Экономика находится в постоянном развитии. Соответственно, появляются новые термины, наименования, аббревиатуры и сокращения, перевод которых требует не только специальные знания, но и осведомленность о текущих событиях и нововведениях, происходящих в мире.

Рост научно-технических знаний в наши дни отражается в том, что свыше 90% новых слов, появляющихся в современных словарях, составляет специальная лексика. Рост числа терминов экономической науки обгоняет рост числа общеупотребительных слов языка, и поэтому изучение специальной лексики становится все более актуальным для развития языка [5].

В настоящее время возрастает необходимость перевода экономической информации в связи с развитием международного бизнеса и зарубежных рынков. Учитывая, что деловое взаимодействие многоязычных компаний в разных странах предполагает постоянное общение и ведение отчетности, перевод экономических текстов стал одним из самых популярных видов перевода. Как подвид технических текстов, перевод такой экономической информации востребован банками, профессиональными и транснациональными корпорациями.

На самом деле экономические тексты охватывают многие области, такие как финансы, бухгалтерский учет, маркетинг и менеджмент. Перевод текстов по экономике может рассматриваться в качестве подвида делового перевода. Он подразумевает перевод таких документов, как бухгалтерские отчеты, бухгалтерские балансы, статистические данные, банковские выписки, гарантийные соглашения, отчеты, счета, счета-фактуры, тендерный пакет, гарантийные письма, а также кредитные заявки, технико-экономические отчеты и т. д.

Следует отметить, что перевод всех видов деловых писем, заявлений, предложений, отчетов, годовых отчетов, договоров купли-продажи, финансовых отчетов, статей экономического и финансового характера также связан с данной областью перевода.

Таким образом, экономический перевод необходим как организациям, так и физическим лицам. Организации сталкиваются с необходимостью экономического перевода в связи с расширением своей внешнеэкономической деятельности, открытием зарубежных филиалов,

заключением новых партнерских отношений. Физическим лицам, как правило, экономический перевод необходим для успешного осуществления инвестиций в зарубежную недвижимость, ценные бумаги, для совершения различных финансовых операций.

Экономические знания стали столь же необходимы, как и общеобразовательная подготовка. Знание экономических процессов не только упрощает жизнь людям, но и помогает рационально использовать денежные средства, понимать взаимосвязь происходящих событий. Поэтому изучение предметной области экономики в наши дни крайне актуально.

2.2 Особенности англоязычных текстов по макро- и мегаэкономике

Макроэкономика изучает совокупности связей между отдельными элементами национальной экономики, в то время как мегаэкономика исследует связи между различными национальными хозяйствами [44]. Исходя из этого, языковые средства, используемые в данных предметных областях, могут различаться. Для выявления специфики англоязычных текстов подязыка макро- и мегаэкономике был составлен и проанализирован корпус научных статей данной предметной области, взятых из журнала *Economics: The Open-Access, Open-Assessment E-Journal* за 2015–2018 года. Его объем составляет 202 445 словоупотреблений.

На основе проведенного анализа были выявлены следующие характеристики подязыка. В общем виде лексический компонент языка экономической специальности можно представить следующим образом:

- слова, обозначающие узкоспециальные понятия и составляющие узкоспециальную терминологическую лексику, обычно отличную в разных языках (безэквивалентная лексика):

the parent of Airbus – управляющая компания «Аэробуса»;

publicly traded company – компания, акции которой свободно обращаются на открытом рынке ценных бумаг;

listed company – компания, акции которой зарегистрированы на фондовой бирже.

- слова, обозначающие понятия, входящие в состав общенаучной экономической терминологии:

contract – контракт, loss – убытки, sales – продажи;

- общеупотребительные слова:

company – компания, office – офис, work – работа.

- неологизмы:

dotcoms – компании, осуществляющие свою деятельность в сети Интернет;

glocalization – создание товаров и услуг для глобализированного рынка с учетом национальных особенностей;

ghost brand – в прошлом знаменитая торговая марка, чьи продукты не пользуются былым спросом.

- метафоры:

golden parachute – «золотой парашют» – крупная денежная компенсация, выдаваемая руководителю высшего звена корпорации при уходе на пенсию или при увольнении, в том числе в случае поглощения данной компании;

safe haven – зона безопасности; безопасная гавань для инвестиций (или для инвесторов);

road map – поэтапный план развития, направленный на достижения конкретной цели.

- латинские выражения, которые часто появляются в деловом дискурсе в целом:

a priori – заранее, предварительно;

ad hoc – для данного случая, для данной цели;

contra – против;

de facto – фактически;

de jure – по праву, юридически;

per se: по существу;

pro bono: ради общественного блага;

quasi – якобы, будто бы;

sic: так, таким образом.

Все эти составляющие являются подвижными, т. е. находятся в постоянном развитии и оказывают друг на друга большое влияние. Именно составляющая лексического компонента стала причиной выделения языка экономической специальности из общеупотребительного языка.

Важнейшей проблемой достижения эквивалентности перевода экономических текстов является передача исходного содержания текста с помощью терминологии языка перевода. Для преодоления этих сложностей важно обратить внимание на ряд особенностей терминов:

- Многие термины являются интернациональными словами: business, firm, product;

- Широкое использование многокомпонентных терминов – терминологических сочетаний, состоящих из двух и более компонентов:

discrete-time survival model – дискретная модель выживаемости по времени;

import demand elasticity – эластичность спроса на импорт;

nominal exchange rate movements – изменения номинального обменного курса.

- Широко распространены аббревиатуры и сокращения:

CPI (Consumer Price Index) – индекс потребительских цен;

NPV (Net Present Value) – чистая приведенная стоимость;

WTO (World Trade Organization) – Всемирная Торговая Организация;

tn (trillion) – триллион (трлн.);

- Использование разных терминов для обозначения идентичных понятий в различных областях экономики, финансов, менеджмента. В этом случае выбор эквивалента обусловлен сферой его употребления, макро- и микроконтекстом. Например, термин «sales» может быть переведен как объем продаж, выручка от реализации; товарооборот. В бухгалтерском учете

применяется термин «выручка от реализации»; в маркетинге – термин «объем продаж»; все они могут фигурировать в тексте финансовой отчетности, обозначая одно и то же явление; термин «interest» может быть переведен термином «процентная ставка» или нетерминологической лексемой «интерес» [14].

Необходимо отметить, что экономические термины создаются на английском языке или заимствуются через английский язык как язык-посредник. Русскоязычная экономическая терминология активно развивается и пополняется. Однако, как отмечено в статье А.А. Гнидченко [9], в русскоязычной научной среде встречаются случаи отсутствия общепринятых терминов, соответствующих англоязычным терминам. Например, *extensive margin* и *intensive margin*. В своей статье автор использует термины «маржа экспорта», «вклад в экспорт» и «составляющая экспорта» как взаимозаменяемые.

Среди синтаксических особенностей англоязычных экономических текстов можно выделить:

- Широкое распространение пассивных, безличных и неопределенно-личных конструкций:

«As indicated in above equation, an increase of REER is associated with a depreciation of the RMB against foreign currency»;

«Firm size is measured by the number of employees»;

- Преобладание сложносочиненных и сложноподчиненных предложений, в которых преобладают существительные, прилагательные и неличные формы глагола:

«Other sources of heterogeneity, such as extensive margins, import demand elasticity, different destinations, U.S. dollar peg, and the liberalization of trading rights is also important to the effects of exchange rate changes»;

- Ряды однородных членов предложений:

«There are always new threats on the horizon in the form of pioneer substitute products, bankruptcy of suppliers, changes in tax regulations, rates of exchange, countries entering WTO, world crisis etc.»;

- Длина предложения (средняя длина предложения составляет 26 слов);
- Обилие предложных фраз в одном предложении:

«Changes in consumer behavior can be connected with descent of shopping ability (for example amid crisis), with appearance of more cheap and/or qualitative substitutes, competitor's new advertising campaign etc. »;

- Наличие многокомпонентных сочетаний, что вызывает синтаксическую многозначность:

«The responses of exporters to exchange rate movements may be different across goods with different import demand elasticity»;

«The high churning rate of exporting firms likely reflects their more intense dynamics on exporting markets».

- Широкое использование эллиптических конструкций:

«Complex modern societies, characterised by specialisation, division of labour and social coordination, emerged through adaptation in different environments, based on the capacity and disposition of people to cooperate when influenced by political instruments».

Основная задача экономических текстов заключается в предельно ясном и точном донесении определенной информации до читателей. Это достигается логически обоснованным изложением фактического материала, без применения эмоционально окрашенных слов, выражений и грамматических конструкций [4].

2.3 Сопоставительный анализ переводов экономических текстов в системах ПРОМТ, Google Translate

С целью оценки качества перевода и пригодности использования систем автоматического перевода при осуществлении перевода экономических текстов (предметной области макро- и мегаэкономики) был проведен

экспериментальный анализ переводов предложений (с английского языка на русский), выполненных с помощью систем автоматического перевода – Google Translate и ПРОМТ. Выбор данных систем автоматического перевода был обусловлен тем, что ПРОМТ и Google Translate являются наиболее доступными и популярными в настоящее время. Также ПРОМТ и Google Translate являются представителями разных типов систем автоматического перевода, т.е. эти системы построены в рамках диаметрально различных подходов и по-разному осуществляют процесс перевода. Система автоматического перевода ПРОМТ выполняет перевод на основе встроенных словарей для выбранной языковой пары, а так же грамматик, охватывающих семантические, морфологические, синтаксические закономерности обоих языков. В свою очередь Google Translate в процессе осуществления перевода использует статистический анализ с элементами технологии нейронного машинного перевода. Система подбирает наиболее частотные межъязыковые соответствия, синтаксические конструкции на основе загруженных двуязычных корпусов текстов.

Эксперимент проводился в несколько этапов:

1. Анализ перевода предложений из составленного корпуса текстов подъязыка макро- и мегаэкономики, выполненного в системах автоматического перевода ПРОМТ и Google Translate с целью выявления английских индикаторов переводимости, лингвистических явлений, вызывающих трудности при переводе, для англо-русского направления перевода;
2. Систематизация ошибок, возникающих в русском АП;
3. Систематизация английских индикаторов переводимости, провоцирующих ошибки;
4. Проверка общих правил предредактирования и разработка новых, специфичных для подъязыка макро-, мегаэкономики.

Для реализации первого этапа эксперимента был составлен подкорпус из 100 предложений, содержащих характерные для экономического языка

лингвистические явления, перечисленные в пункте 2.2. Данные предложения были переведены с английского языка на русский в онлайн-переводчиках ПРОМТ и Google Translate. Полученные переводы анализировались с точки зрения лексической и грамматической правильности перевода, соблюдения стилистических норм, правил пунктуации, покрываемости системы. В табл. 2.1 и 2.2 представлены фрагменты первого этапа эксперимента, полная версия таблицы приведена в приложении 1. В первом столбце указан номер предложения. Во втором столбце представлено исходное англоязычное предложение (индикаторы переводимости подчеркнуты), в третьем – его перевод на русский язык (неправильный перевод входного предложения, вызванный подчеркнутыми индикаторами переводимости английского текста). В четвертом столбце перечислены типы выявленных индикаторов переводимости, содержащиеся в данном предложении. Для большей наглядности каждому индикатору переводимости присвоено оформление. Термины, как наиболее значимые элементы экономического текста, выделены серым цветом. Многокомпонентные термины – подчеркиванием линией. Синтаксическая многозначность – подчеркиванием двойной линией. Лексическая многозначность – волнистой линией. Лексическая неоднозначность языковых единиц – курсивом. Место пропущенного слова обозначено символом у.

Стилистические ошибки не выделяются в данных таблицах, поскольку являются следствием уже совершенных переводческих ошибок. Как правило, среди стилистических ошибок можно выделить:

- лексическую несочетаемость слов: *export volume* – громкость экспорта вместо объем экспорта;
- тавтологию: *this study explores* – это исследование исследует; *to be complemented by additional development finance* – быть дополнена дополнительным финансированием развития;

Таблица 2.1 – Анализ перевода с помощью системы ПРОМТ

№	Исходное предложение	Перевод	Типы индикаторов переводимости
1	<p>The present study <u>intends</u> to examine the idea and factors of <u>business competitive ability</u>, so that from this point on <u>offers</u> a <u>range</u> of tools which will be the most effective in <u>competitive abilities control</u> on the level of companies in the <u>near</u> decade.</p>	<p>Данное исследование <u>намеревается</u> исследовать идею и факторы <u>деловой конкурентоспособной способности</u>, так, чтобы с этого момента <u>предложил диапазон</u> инструментов, которые будут самыми эффективными при <u>конкурентоспособном контроле за способностями</u> на уровне компаний в <u>близкое десятилетие</u>.</p>	<p><u>Многозначность лексемы</u> («intend», «range», «near»); <u>Многокомпонентный термин</u>, провоцирующий появление <u>синтаксической многозначности</u>; <u>Грамматическая категория времени</u>.</p>
2	<p>The positive <u>dynamics</u> of these <u>rates in the longer term</u> mean that the product of a company finds a market, i.e. <u>it</u> is competitive, that the <u>branch</u> wherein the firm is functioning is <u>consolidating</u> on domestic or foreign markets, <u>that it</u> contributes to national economy by way of <u>work positions</u>, taxes, <u>levies</u> and other payments.</p>	<p>Положительные <u>движущие силы</u> этих <u>ставок в дальнейшей перспективе</u> означают, что продукт компании <u>находит рынок</u>, т.е. <u>это конкурентоспособно</u>, что <u>отделение</u>, где фирма функционирует, <u>объединяется</u> на внутренних или иностранных рынках, <u>которые это вносит</u> <u>у</u> в национальную экономику посредством <u>положений работы</u>, налогов, <u>налогов</u> и других платежей.</p>	<p><u>Многозначность лексемы</u> («dynamics», «rates», «it», «branch», «consolidate», «that»); <u>Многокомпонентный термин</u>; <u>Термин</u>; Местоимение it; Пропущено слово (<u>v</u>); Ряд однородных придаточных предложений, длина предложения.</p>

Таблица 2.2 – Анализ перевода с помощью системы Google Translate

№	Исходное предложение	Перевод	Индикатор переводимости
1	<i>Firm</i> age is measured by the difference between firm <u>establishing</u> year and the current year.	<i>Твердый</i> возраст измеряется разницей между годом <u>установления</u> фирмы и текущим годом.	<u>Многозначность лексемы</u> («establishing»); <i>Лексическая</i> <i>неоднозначность языковых</i> <i>единиц («firm»).</i>
2	The higher hazard of exiting and the lower propensity of entering for younger firms caused by exchange rate appreciation may be explained by their <u>poor</u> market experience and lower <u>financial ability</u> .	Более высокая опасность выхода и более низкая склонность к вступлению <u>для</u> молодых фирм, вызванная повышением курса валют, могут быть объяснены их <u>плохим</u> опытом работы на рынке и более низкой <u>финансовой способностью</u> .	<u>Многозначность лексемы</u> («poor»); Сложная синтаксическая структура (лишний предлог «для» – разница в передаче Р.п.); <u>Многокомпонентный термин.</u>
3	The Convention recognizes that a large influx of refugees may overburden individual host countries and calls for international cooperation.	В Конвенции признается , что большой приток беженцев может перегрузить отдельные страны пребывания и призывает к международному сотрудничеству.	Дистантное расположение однородных членов.

- буквальный перевод: based on the capacity and disposition of people to cooperate when influenced by – на основе способности и расположения людей, чтобы сотрудничать когда под влиянием.

В первых двух примерах стилистическая ошибка возникла вследствие неправильно выбранного значения лексемы, а в последнем – особенностей технологии перевода системы ПРОМТ.

Среди прочих ошибок стоит отметить пунктуационные ошибки. Они не искажают содержание полученного перевода, но незначительно влияют на его качество. Пунктуационные ошибки возникают вследствие различного строения языков, выполняемых функций знаков препинания в предложении.

В результате анализа переводов, выполненных в системах автоматического перевода ПРОМТ и Google Translate, выявленные в русскоязычном тексте ошибки, вызванные индикаторами переводимости англоязычного текста, были систематизированы и классифицированы по группам:

- Семантические ошибки:

- Неправильный перевод экономической терминологии:

- margin – край (маржа);

- competition – соревнование (конкуренция);

- demand – требование (спрос);

- Неправильный перевод многокомпонентных терминов:

- foreign trading rights – иностранные торговые права (права на внешнюю торговлю);

- import demand elasticity – эластичность требования импорта (эластичность спроса на импорт);

- business competitive ability – деловая конкурентоспособная способность (конкурентоспособность бизнеса);

- Неправильный выбор значения при переводе, что приводит к искажению смысла предложения:

- movements – движения (изменения);

appreciation – оценка (повышение);

response – ответ, отклик (реакция);

- Неправильный перевод служебных частей речи (предлогов, союзов, частиц и т.д.):

exit *from* export markets – выйти *от* внешних рынков (выйти *с* внешних рынков);

risk of failure *in* export markets – риск неудачи *во* внешних рынках (риск неудачи *на* внешних рынках);

to hunt *for* greater returns *from* riskier assets – охотиться *для* большей прибыли *из* более опасных активов (охотиться *за* большей прибылью *от* более опасных активов);

- Лексические ошибки:

- Неправильный выбор части речи:

capital mobility – капитальная подвижность (мобильность капитала);

firm size – устойчивый размер (размер фирмы);

public good – хорошей общественности (общественное благо);

- Синтаксические ошибки:

- Пропущен предлог:

likelihood of export market entering – вероятность входа в внешнего рынка;

more intense export market competition – более интенсивное соревнование в внешнего рынка (более жесткая конкуренция на экспортном рынке);

for long-term market presence – для долгосрочного присутствия в рынка (для долгосрочного присутствия на рынке).

- Использование лишнего предлога:

effective tools for business responses – эффективные инструменты *для* реагирования бизнеса (инструменты реагирования бизнеса);

better than expected growth in the global economy – лучше, чем ожидаемый рост в мировой экономике (лучшим, чем ожидаемый рост мировой экономики);

change in the general price level – изменение в общем уровне цен (изменение общего уровня цен).

– Нарушение согласования по числу, падежу:

on domestic or foreign markets – на внутреннем или внешнем рынке (рынках);

one concern is that... – одна из опасений заключается в том, что... (одно из опасений заключается в том, что...);

weighted Chinese RMB – взвешенный китайский юаней (торгово-взвешенный курс юаня);

the influence of exchange rate appreciation on export market dynamics – влияние оценки обменного курса на динамике внешнего рынка (влияние повышения обменного курса на динамику экспортного рынка).

– Неверно выбрана категория времени, залог:

the World Bank *has warned*... – Мировой банк *предупреждает*, что ... (предупредил, что...);

growth in China's manufacturing sector *slowed* in December – рост в промышленном секторе Китая, *который замедляют* в декабре (рост в промышленном секторе Китая *замедлился* в декабре);

the policies *implemented* by government – политика, *проводившая* правительством (политика, *осуществляемая* правительством).

– Пунктуационная ошибка:

in addition, labour taxes, or at least their contribution... – кроме того, трудовые налоги, или по крайней мере их вклад.

- Стилистические ошибки (были рассмотрены ранее);

- Покрываемость системы: невыполненный перевод слова, сокращения, аббревиатуры:

\$9 tn (£6.6 tn) – 9 трлн \$ (6,6 трлн. £);

International Financial Statistics (IFS) – Международная финансовая статистика;

other U.S. dollar peggers – другой доллар США peggers (валюты, привязанные к доллару США).

Сравнительные характеристики перечисленных типов ошибок и их частотность для каждой из систем автоматического перевода представлены в табл. 2.3 и на рис. 2.1.

Таблица 2.3 – Типы ошибок при англо-русском АП экономических текстов в системах ПРОМТ и Google Translate

№	Ошибка	Частота появления	
		ПРОМТ	Google Translate
Семантические ошибки:			
1*	Неправильный перевод многокомпонентного термина	61	23
1.	Неправильный перевод термина	43	19
2.	Неправильный выбор значения	41	22
3.	Неправильный перевод предлога	16	7
4.	Неправильный перевод союза	3	-
5.	Неправильный перевод вводного слова	1	1
6.	Неправильный перевод частицы	1	-
Итого:		166	72

Продолжение таблицы 2.3

№	Ошибка	Частота появления	
		ПРОМТ	Google Translate
Лексические ошибки:			
7.	Неправильный выбор части речи	20	5
Итого:		20	5
Синтаксические ошибки:			
8.	Лишний предлог	9	6
9.	Пропущен предлог	13	2
10.	Многокомпонентное прилагательное	1	1
11.	Нарушение согласования	11	5
12.	Незаконченный перевод (пропущено слово)	2	6
13.	Неверное согласование по числу	3	3
14.	Неверно выбранная категория времени	9	1
15.	Неправильно переведен залог	1	1
16.	Пунктуационная ошибка	7	1
17.	Добавление лишнего слова	-	2
18.	Порядок слов	1	-
Итого:		57	28
Стилистические ошибки:			
19.	Стилистическая ошибка	24	8
Итого:		24	8

№	Ошибка	Частота появления	
		ПРОМТ	Google Translate
Покрываемость системы:			
20.	Невыполненный перевод	3	-
21.	Невыполненный перевод сокращения	1	-
22.	Неспособность системы АП перевести аббревиатуру	1	1
23.	Пропущена отрицательная приставка	-	2
Итого:		5	3
ВСЕГО:		272	116

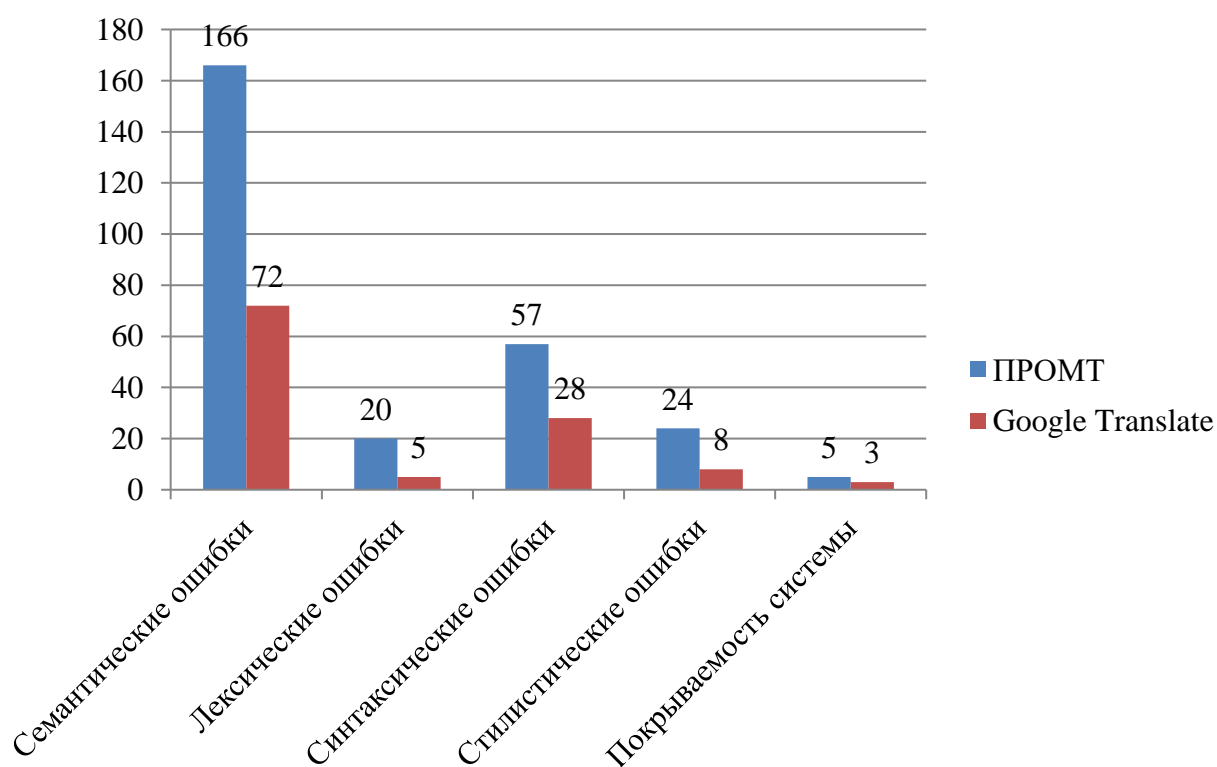


Рисунок 2.1 – Типы ошибок при англо-русском АП экономических текстов в системах ПРОМТ и Google Translate

Из таблицы 2.3 и рисунка 2.1 видно, что наиболее распространенными ошибками при англо-русском переводе экономических текстов, относящихся к подязыку макро-, мегаэкономики, в системах ПРОМТ и Google Translate являются семантические ошибки, в особенности перевод терминологии. Пример неправильного перевода терминологии представлен в табл. 2.4.

Таблица 2.4 – Пример неправильного перевода терминологии

Оригинал	ПРОМТ	Google Translate
Exchange rate is a significant institutional factor affecting <u>aggregate trade</u> and <u>individual exporting behavior.</u>	Обменный курс – значительный установленный фактор, затрагивающий <u>совокупную торговлю</u> и <u>отдельное поведение экспорта.</u>	Обменный курс является значительным институциональным фактором, влияющим на <u>совокупную торговлю</u> и <u>индивидуальное экспортное поведение.</u>

Термины «aggregate trade» («общий объем торговли») и «individual exporting behavior» («поведение отдельных экспортеров») были неверно переведены обеими системами АП. Термины, в частности многокомпонентные терминологические сочетания, являются опасным индикатором переводимости. При их переводе пользователь должен прибегать к использованию специализированных словарей, глоссариев, а также проверять правильность выполненного перевода с помощью поисковых систем по методике, описанной в [22]. Использование предложенной методики является предпочтительнее, поскольку «традиционные» словари и их оцифрованные версии имеют ограниченную покрываемость, и не содержат достаточное количество примеров перевода многокомпонентной терминологии. Также стоит учитывать, что экономическая терминология находится в постоянном развитии, поэтому перевод необходимого термина может попросту отсутствовать в словаре. В сравнении с «бумажными» словарями электронные ресурсы обеспечивают

большую информативность и актуальность, но они могут содержать переводческие ошибки, поскольку пополняются не только непрофессиональными переводчиками и лексикографами, но и самими пользователями.

Среди семантических ошибок также очень частотными являются ошибки, связанные с неправильным выбором значения лексемы, что приводит к искажению смысла. Данные ошибки вызваны такими индикаторами переводимости как лексическая и синтаксическая многозначность. Пример приведен в табл. 2.5.

Таблица 2.5 – Пример влияния ИП лексическая и синтаксическая многозначность на корректность перевода

Оригинал	ПРОМТ	Google Translate
<p><u>High productivity firms</u> are less likely to exit from export markets and more likely to enter and survive in export markets in the period of exchange rate <u>appreciation</u>.</p>	<p><u>Высокие фирмы</u> <u>производительности,</u> менее вероятно, выйдут от внешних рынков и более вероятно войти и выжить на внешних рынках в период <u>оценки</u> обменного курса.</p>	<p><u>Компании с высокой</u> <u>производительностью</u> менее склонны к выходу с экспортных рынков и более склонны к выходу и выходу на экспортные рынки в период <u>повышения</u> обменного курса.</p>

Лексическая единица «appreciation» имеет несколько значений: признательность, оценка, понимание, одобрение, повышение, удорожание. Учитывая контекст, подходящим эквивалентом является вариант «повышение».

Синтаксическая многозначность «high productivity firms» в данном предложении была разрешена системой АП Google Translate. ПРОМТ, в свою очередь, выбрал неверное опорное слово, что привело к некорректному переводу.

Чтобы избежать подобных ошибок, необходимо заменять многозначные слова на их однозначные синонимы, а цепочки существительных разбивать на группы с соответствующим предлогом.

Второй по частности является группа синтаксических ошибок, которые провоцируются такими индикаторами переводимости как: сложная структура предложения и его длина, предложные фразы, дистантное расположение однородных и зависимых членов предложения, бессоюзное присоединение придаточных предложений, цепочки существительных, союз «and» и т.д. Примеры представлены в табл. 2.1 и 2.2.

Лексические ошибки возникают в силу наличия в тексте лексической неоднозначности языковых единиц, когда лексема может принадлежать к разным частям речи. Пример приведен в табл. 2.6.

Таблица 2.6 – Пример влияния ИП лексическая неоднозначность на корректность перевода

Оригинал	ПРОМТ	Google Translate
Firm age is measured by the difference between firm establishing year and the current year.	Устойчивый возраст измерен различием между устойчивым годом установления и в текущем году.	Твердый возраст измеряется разницей между годом установления фирмы и текущим годом.

В обоих случаях системы автоматического перевода ПРОМТ и Google Translate неверно распознали частеречную принадлежность лексемы «firm», используя в переводе прилагательное, а не существительное.

Покрываемость системы – это способность системы АП перевести все слова исходного текста. Данные ошибки являются менее распространенными.

На следующем этапе эксперимента индикаторы переводимости английских текстов предметной области макро-, мегаэкономики при переводе на русский язык были проанализированы относительно их

частотности в системах автоматического перевода ПРОМТ и Google Translate. Результаты этого анализа представлены в табл. 2.7.

Таблица 2.7 – Сравнительная статистика индикаторов переводимости при англо-русском МП текстов по макро-, мегаэкономике

Индикатор переводимости	ПРОМТ	Google Translate
Многозначность лексемы	91	41
Многокомпонентная терминология	61	28
Синтаксическая многозначность	11	4
Лексическая неоднозначность	19	5
Сложная синтаксическая структура	23	10
Местоимение it	3	2
Дистантное расположение однородных членов предложения	5	1
Предложные фразы	33	17
Грамматическая категория времени	9	1
Покрываемость системы	5	2
Стиль	21	7
Форма слова	7	7
Ряд однородных членов предложения	3	-

Наиболее частотные индикаторы переводимости английского языка, характерные для английских текстов предметной области макро-, мегаэкономики при англо-русском переводе в системах автоматического перевода ПРОМТ и Google Translate, представлены на рис. 2.2 и 2.3 соответственно.

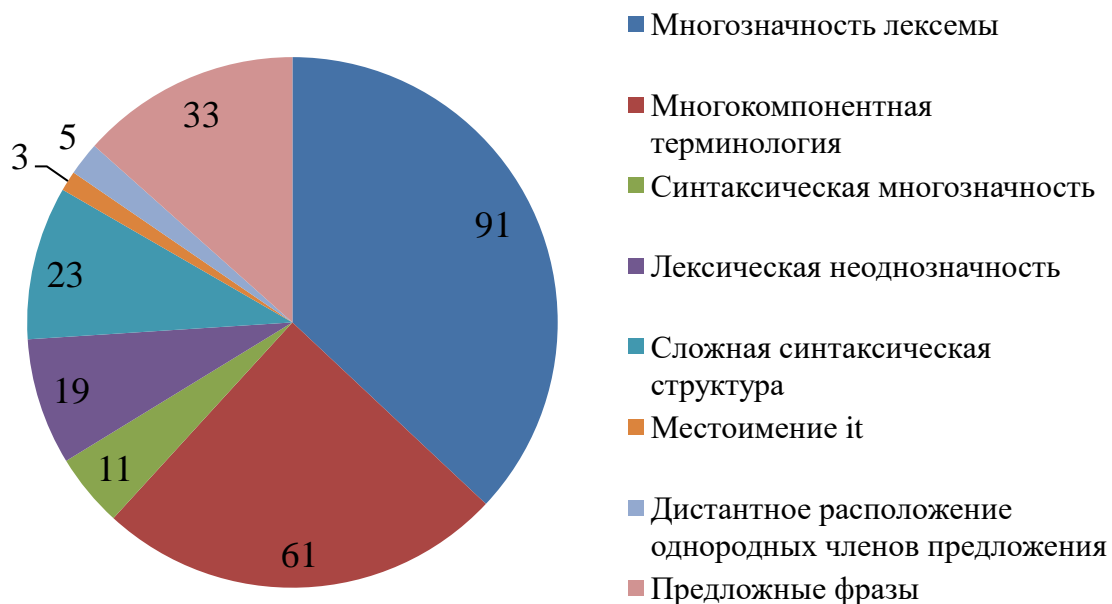


Рисунок 2.2 – Наиболее частотные ИП, характерные при англо-русском переводе в системе АП ПРОМТ



Рисунок 2.3 – Наиболее частотные ИП, характерные при англо-русском переводе в системе Google Translate

На последнем этапе эксперимента осуществлялась проверка общих правил предредактирования и разработка новых, специфичных для подъязыка макро-, мегаэкономики для систем автоматического перевода PROMT и Google Translate с целью повышения качества получаемого перевода. Правила предредактирования предусматривают переформулирование фрагментов текста, содержащих индикаторы переводимости, по следующим правилам:

1. Цепочки существительных разбить на группы с предлогами (of/to/in/with);
2. При наличии ряда однородных членов предложения, связанных союзно/бессоюзно, после второго компонента цепочки дублировать предлог, указывающий на образование необходимой словоформы;
3. Существительные, имеющие окончание –ing, заменить синонимами или поставить в начало словосочетания с предлогом of в постпозиции;
4. Заменить многозначные слова на их однозначные синонимы;
5. Исправить авторские ошибки (опечатки, неверное использование предлогов и т.д.);
6. Перед существительным всегда должен присутствовать артикль или указательное/притяжательное местоимение;
7. Предложения, содержащие дистантную связь, разбивать на более простые;
8. Сложные предложения, содержащие множественную сочинительную связь, разбивать на простые;
9. Длина предложения должна составлять не более 20 слов;
10. Заменить инфинитивный оборот-сказуемое на синонимичное прилагательное;
11. Не употреблять многокомпонентные прилагательные;
12. Избегать использование многокомпонентных терминов;
13. Обращать особое внимание на предложные фразы;

14. Снять синтаксическую многозначность, чтобы избавиться от неоднозначности;
15. Пассивный залог заменить на активный залог;
16. Восстановить пропущенное подлежащее и/или сказуемое;
17. Добавить союз перед бессоюзно присоединенным придаточным предложением;
18. Проверять перевод терминологии по методике, описанной в [22].

На последнем этапе эксперимента из исходного корпуса текстов (202 450 словоупотреблений) был составлен второй подкорпус из 50 предложений (отличный по содержанию от первого подкорпуса). В него вошли предложения с индикаторами переводимости, которые были выявлены на предыдущих этапах эксперимента. Данные предложения были отредактированы (перефразированы) по разработанным правилам контролируемого языка, а затем вновь переведены с помощью систем автоматического перевода ПРОМТ и Google Translate. Фрагмент таблицы с результатами редактирования предложений в Google Translate и ПРОМТ указаны в табл. 2.8 и 2.9 соответственно. Полные варианты таблиц представлены в приложениях 2 и 3.

По результатам проведенной работы с учетом того, что системы автоматического перевода ПРОМТ и Google Translate совершают наибольшее число ошибок в терминологии, мы составили глоссарий терминов, в том числе и многокомпонентных, принадлежащих к подобласти макро-, мегаэкономики. Данный глоссарий призван помочь переводчикам избежать ошибок при переводе терминологии данного подязыка. Фрагмент глоссария представлен в табл. 2.10. Полная версия глоссария находится в приложении 4.

Таблица 2.8 – Результаты предредактирования предложений в Google Translate

	Исходное предложение	Первоначальный перевод	№ Правила	Перифраз	Перевод после предредактирования
1	The responses of exporters to exchange rate movements may be different across goods with...	Отклики экспортеров на движение обменного курса могут быть разными в разных товарах с ...	4 5 6 12	The reaction of exporters to exchange rate changes may be different among the goods with ...	Реакция экспортеров на изменения обменного курса может быть разной среди товаров с ...
2	... that could be triggered by better than expected growth in the global economy this year.	... которые могут быть вызваны лучшими, чем ожидалось, ростом мировой экономики в этом году.	4 11	... that could be triggered by the growth which was higher than expected in the global economy this year.	... это может быть вызвано ростом, который был выше, чем ожидалось в мировой экономике в этом году.
3	... but also switch their products mix or even export market dynamics	... но также переключают свою смесь товаров или даже динамику экспорта .	1 4 12	... but also change their range of goods or even dynamics of the export market но и поменять ассортимент товаров или даже динамику экспортного рынка .

Таблица 2.9 – Результаты предредактирования предложений в ПРОМТ

	Исходное предложение	Первоначальный перевод	№ Правила	Перифраз	Перевод после предредактирования
1	Recently, the New York Declaration has again called for ■ more equitable sharing of burden and responsibility, to be achieved through a new global compact on refugees by 2018.	Недавно, нью-йоркская Декларация снова призвала к большему количеству равноправного разделения бремени и ответственности, чтобы быть достигнутой через новый Глобальный договор на беженцах к 2018.	6 7 13 17 18	Recently, the New York Declaration has again called for a more equitable sharing of burden and responsibility. The sharing should be achieved by means of a new global compact on refugees by 2018.	Недавно, нью-йоркская Декларация снова призвала к более равноправному разделению бремени и ответственности. Разделение должно быть достигнуто посредством нового Глобального договора на беженцах к 2018.
2	The higher hazard caused by exchange rate appreciation may be explained by their poor market experience and lower financial ability .	Более высокая опасность, вызванная оценкой обменного курса, может быть объяснена их бедным рынком, испытывают и понижают финансовую способность .	1 4 6 11 18	The higher hazard caused by the rise of exchange rate may be explained by their insufficient work experience in the market and their lower financial opportunity .	Более высокая опасность, вызванная повышением обменного курса, может быть объяснена их недостаточным опытом работы на рынке и их более низкой финансовой возможностью .

Таблица 2.10 – Фрагмент глоссария

№	Оригинал	Правильный перевод термина
1.	additional food demand	дополнительный спрос на продовольствие
2.	aggregate economic benefits	совокупная экономическая выгода
3.	aggregate trade	общий объем торговли
4.	annual inflation	годовая инфляция
5.	appreciation	повышение
6.	business profitability	рентабельность бизнеса
7.	capital mobility	мобильность капитала
8.	competitive abilities control	управление конкурентоспособностью
9.	competitive ability	конкурентоспособность
10.	complementary log-log model (cloglog)	модель с дополняющим двойным логарифмированием (cloglog)
11.	cooling property market	охлаждение рынка недвижимости
12.	descent of shopping ability	снижение покупательской способности
13.	donors	доноры, инвесторы
14.	exchange rate movements	изменения обменного курса
15.	individual exporting behavior	поведение отдельных экспортеров

Проведенный эксперимент показал, что разработанные правила контролируемого языка для предредактирования англоязычных экономических текстов, относящихся к подобласти макро-, мегаэкономики, значительно повышают качество получаемого перевода на русский язык, выполненного с помощью систем автоматического перевода ПРОМТ и Google Translate. Но для достижения более высокого качества перевода данные системы требуют разные правила предредактирования и в разном

количестве. Поскольку системы автоматического перевода ПРОМТ и Google Translate являются представителями различных технологий перевода, то и ошибки, возникающие при переводе, отличаются по типу и частотности.

Умение работать с системами автоматического перевода, знание их «слабых мест», то есть специфичных индикаторов переводимости выбранной языковой пары и выбранной системы АП, может помочь пользователям получить перевод более высокого качества, сократить время, затрачиваемое на осуществление перевода и постредактирования.

Результаты эксперимента показывают, что система автоматического перевода Google Translate лучше осуществляет перевод англоязычных текстов подъязыка макро-, мегаэкономики и требует значительно меньшего количества правил контролируемого языка, чем ПРОМТ. Главным недостатком онлайн-переводчика ПРОМТ является недостаточная покрываемость, в то время как система АП Google Translate при отсутствии лексемы в базе знаний опускает её перевод и выдает предложение с неправильным, искаженным смыслом. Поэтому необходимо проверять перевод термина с помощью составленного нами глоссарий или с помощью поисковых систем по методике, описанной в [22].

Таким образом, система АП Google Translate, с пользовательской точки зрения, может быть рекомендована к использованию в качестве приоритетной.

Выводы по главе 2

Необходимость перевода экономической информации возрастает в связи с развитием международного бизнеса и зарубежных рынков. Учитывая, что деловое взаимодействие многоязычных компаний в разных странах предполагает постоянное общение и ведение отчетности, перевод экономических текстов стал одним из самых популярных видов перевода.

Проблемы перевода экономических текстов связаны, прежде всего, со спецификой предметной области, которая отражается в терминологии, синтаксисе, лексике и стиле изложения.

По результатам экспериментального анализа были выявлены следующие характеристики подязыка макро-, мегаэкономики: насыщенность терминологией, в том числе многокомпонентной; большое количество аббревиатур и сокращений; большая длина предложения (средняя длина предложения составляет 26 слов); широкое распространение пассивных, безличных и неопределенно-личных конструкций; обилие предложных фраз в одном предложении; преобладание сложносочиненных и сложноподчиненных предложений, с длинными рядами однородных членов в которых превалируют существительные, прилагательные и неличные формы глагола; наличие многокомпонентных сочетаний, что вызывает синтаксическую многозначность.

Наиболее опасными индикаторами переводимости, специфичными для подязыка макро-, мегаэкономики при англо-русском направлении перевода оказалась терминология, в особенности многокомпонентные терминологические сочетания. Другими специфичными индикаторами переводимости являются: лексическая неоднозначность, многозначность лексем, синтаксическая многозначность, предложные фразы, дистантное расположение однородных членов предложения, длина предложения, сложная структура предложения и т.д.

Проведенный эксперимент показал, что индикаторы переводимости можно устранить из исходного текста с помощью предредактирования на основе контролируемого языка, учитывающего индикаторы переводимости, и таким образом значительно повысить качество автоматического перевода без обращения к разработчикам систем.

ЗАКЛЮЧЕНИЕ

Повышение эффективности работы переводчика обозначает сокращение времени на обработку определенного объема текста при сохранении или улучшении качества перевода.

В процессе осуществления перевода переводчик сталкивается с различными трудностями. На этапе предпереводческого анализа главной трудностью является отсутствие экстралингвистических сведений, поэтому переводчик вынужден прибегать к справочной литературе, текстам по схожей тематике, а также созданию аннотации для уточнения содержания текста. На этапе аналитического вариативного поиска главной трудностью является определение единиц перевода в подлиннике и нахождение их эквивалентных соответствий. Для подбора лексических единиц переводчик использует различные источники информации: толковые словари, словари синонимов, узкоспециальные словари, энциклопедии, онлайн-словари, память переводчика, параллельные и квазипараллельные корпуса текстов, системы автоматического перевода. На последнем этапе перевода переводчик проводит анализ результатов перевода: производит правку при необходимости, сверяет полноту текста, орфографическую, пунктуационную коррекцию, коррекцию опечаток, стилистическое редактирование. Для оптимизации данного этапа могут использоваться специальные программы, проверяющие орфографические, грамматические и стилистические ошибки, т.е. редакторы.

На сегодняшний день существующие вспомогательные инструменты в помощь переводчику: электронные словари, глоссарии, системы переводческой памяти, корпуса параллельных и квазипараллельных текстов и др. чрезвычайно популярны в переводческом сообществе, но не позволяют получить перевод всего текста при минимальных временных затратах – для этих целей предназначены системы АП.

Автоматический перевод – это процесс передачи содержания текста с одного естественного языка на другой с помощью системы автоматического перевода.

В настоящее время выделяется три типа систем автоматического перевода: системы на основе грамматических правил (RBMТ), статистические системы (SMT), нейронный автоматический перевод (NMT). Последний тип систем АП является наиболее перспективным и начал развиваться относительно недавно, но уже демонстрирует хорошие результаты, хоть и уступает переводу, выполненному профессиональным переводчиком.

Наиболее популярные системы автоматического перевода ПРОМТ и Google Translate в процессе перевода применяют нейронные сети.

Бесспорно, качество автоматического перевода неидеально и требует взаимодействия с человеком в форме предредактирования, постредактирования или интерредактирования. Однако умение работать с системами автоматического перевода, знание специфичных индикаторов переводимости выбранной языковой пары и выбранной системы АП, может помочь пользователям получить перевод более высокого качества, сократить время, затрачиваемое на осуществление перевода.

В ходе проведенного нами экспериментального анализа были выявлены особенности подъязыка макро-, мегаэкономики: насыщенность терминологией, в том числе многокомпонентной; большое количество аббревиатур и сокращений; большая длина предложения (средняя длина предложения составляет 26 слов); широкое распространение пассивных, безличных и неопределенно-личных конструкций; обилие предложных фраз в одном предложении; преобладание сложносочиненных и сложноподчиненных предложений, с длинными рядами однородных членов; наличие многокомпонентных сочетаний, что вызывает синтаксическую многозначность.

С целью оценки качества перевода и пригодности использования систем автоматического перевода при переводе экономических текстов (предметной области макро- и мегаэкономики) был проведен экспериментальный анализ переводов предложений (с английского языка на русский), выполненных с помощью систем автоматического перевода – Google Translate и ПРОМТ. В результате эксперимента нами были выявлены типичные ошибки, возникающие в русском переводе, а так же специфичные индикаторы переводимости, провоцирующие появление данных ошибок. К ним относятся: терминология, в том числе многокомпонентная, лексическая неоднозначность, многозначность лексем, синтаксическая многозначность, предложные фразы, дистантное расположение однородных членов предложения, длина предложения, сложная структура предложения и т.д. Системы автоматического перевода ПРОМТ и Google Translate в разной степени справляются с переводом данных лингвистических явлений, поскольку осуществляют перевод по разным технологиям. С пользовательской точки зрения, система АП Google Translate может быть рекомендована к употреблению, поскольку она лучше справляется с переводом терминологии, в том числе многокомпонентной.

С целью повышения эффективности деятельности переводчика были разработаны правила контролируемого языка для предредактирования исходного текста. Проведенный эксперимент показал, что составленные правила для систем автоматического перевода ПРОМТ и Google Translate значительно повышают качество получаемого перевода. Индикаторы переводимости могут быть выявлены и устранены до момента их появления в тексте перевода, то есть не потребуется выполнять постредактирование.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

- 1 Алексеева, И. С. Введение в перевод введение: учеб. пособие для студ. филол. и лингв, фак. высш. учеб. заведений / И. С. Алексеева. – СПб.: Филологический факультет СПбГУ. – М.: Издательский центр «Академия», 2004. – 352 с.
- 2 Андреева, А. Д. Обзор систем машинного перевода / А. Д. Андреева, И. Л. Меньшиков, А. А. Мокрушин // Молодой ученый. – 2013. – №12. – С. 64–66.
- 3 Бархударов, Л. С. Язык и перевод / Л. С. Бархударов. – М.: Международные отношения, 1975. – 240 с.
- 4 Беляева, М. Знаковая природа экономического термина как основа перевода / М. Беляева, В. Д. Табанакова // Язык и литература. – Тюмень: Тюмен. гос. ун-т, 2001. – Вып. 9. – С. 12–15.
- 5 Винтова, Н. Л. Тенденции развития терминологических интернационализмов в экономической лексике английского языка / Н. Л. Винтова. – М.: Изд-во Панина, 2006. – Вып. 2. – С. 184–188.
- 6 Витушина, В. В. Экономические тексты как объект перевода / В. В. Витушина. – СПб.: НиваЛит, 2005. – 210 с.
- 7 Волкова, Л. Л. Машинная лингвистика: от перевода со словарём к нелинейным динамическим системам / Л. Л. Волкова // Новые информационные технологии в автоматизированных системах. – 2013. – №16. – С. 317–328.
- 8 Гак, В. Г. К типологии лингвистических номинаций / В. Г. Гак // Языковые номинации. Общие вопросы. – М.: Наука, 1977. – С. 230–294.
- 9 Гнидченко, А. А. Декомпозиция роста экспорта на экстенсивную и интенсивную составляющие с учетом сравнительных преимуществ / А. А. Гнидченко // Журнал Новой экономической ассоциации. – №4 (24). – С. 38–64.

10 Дудина, И. А. Практикум по переводу терминологии теории новой институциональной экономики: English-Russian / И.А. Дудина. – Волгоград: Волгогр. науч. изд-во, 2004. – 74 с.

11 Калинин, С. М. Обзор современных подходов к улучшению точности нейронного машинного перевода / С. М. Калинин // Rhema. Рема. – 2017. – №2. – С. 70–79.

12 Колганов, Д. С. Обзор аналитической, статистической и нейронной технологий машинного перевода / Д. С. Колганов, Е. А. Данилов // Международный студенческий научный вестник. – Пенза, 2018. – № 3. – С. 301–303.

13 Комиссаров, В. Н. Теоретические основы методики обучения переводу / В. Н. Комиссаров. – М.: Рема, 1997 – 112 с.

14 Лашкова, Г. В. К вопросу о морфологической структуре английских заимствований в экономической терминологии русского и немецкого языков / Г. В. Лашкова, И. А. Ибатуллина, Л. П. Салеева // Романо-германская филология. – Саратов, 2003. – Вып. 3. – С. 145–152.

15 Марчук, Ю. Н. Компьютерная лингвистика: учеб. пособие / Ю. Н. Марчук. – М.: АСТ: Восток-Запад, 2007. – 317 с.

16 Нелюбин, Л. Л. Толковый переводоведческий словарь / Л. Л. Нелюбин – 3-е издание, переработанное. – М.: Флинта: Наука, 2003. – 320 с.

17 Переходько, И. В. Оценка качества компьютерного перевода / И. В. Переходько, Д. А. Мячин // Вестник ОГУ. – 2017. – № 2. – С. 92–96.

18 Пиванова, Э. В. Теория и практика машинного перевода: учебное пособие / Э. В. Пиванова – Ставрополь: Изд-во СКФУ, 2014. – 114 с.

19 Рецкер, Я. И. О закономерных соответствиях при переводе на родной язык / Я. И. Рецкер // Вопросы теории и методики учебного перевода: сборник статей. – М.: Изд-во Академии педагогических наук РСФСР, 1950. – С. 156–183.

20 Рыбков, М. В. Эффективность работы переводчика при использовании САТ-систем / М. В. Рыбков, Н. Ф. Паникарова // Молодёжь и наука: Сборник

материалов VI Всероссийской научно-технической конференции студентов, аспирантов и молодых учёных. – Красноярск: Сибирский федеральный ун-т, 2011. – С. 78–82.

21 Улиткин, И. А. Автоматическая оценка качества машинного перевода научно-технического текста / И. А. Улиткин // Вестник Московского государственного областного университета. Сер. Лингвистика. – 2016. – № 4. – С. 174–182.

22 Шереметьева, С. О. К вопросу об электронных ресурсах профессиональной лексики / С. О. Шереметьева, П. Г. Осминин, Е. С. Щербаков // Вестник ЮУрГУ. Сер. Лингвистика – Челябинск: Изд-во ЮУрГУ, 2014. – №2 (37). – С. 230–232.

23 Шереметьева, С.О. Информационные технологии в помощь переводчику / С.О. Шереметьева, П.Г. Осминин. – Челябинск: Издательский центр ЮУрГУ, 2014. – Ч. 3. – 43 с.

24 Шичкина, М. Г. Современные способы и средства повышения эффективности процесса перевода / М. Г. Шичкина // Молодой ученый. – 2016. – №28. – С. 1036–1041.

25 Слокум, Дж. Обзор разработок по машинному переводу / Дж. Слокум // Новое в зарубежной лингвистике. – М.: Наука, 1989. – 134 с.

26 Banerjee, S. METEOR: An Automatic Metric for MT Evaluation with Improved Correlation with Human Judgments / S. Banerjee, A. Lavie // Proc of ACL Workshop on Intrinsic and Extrinsic Evaluation Measures for Machine Translation and/or Summarization. – Ann Arbor, USA, 2005. – P. 65–72.

27 Bernth, A. MTranslatibility / A. Bernth, C. Gdaniec // AMTA-2000 Tutorial. – 2000. – P. 175–218.

28 Callison-Burch, C. Evaluation of machine translation / C. Callison-Burch, C. Fordyce, P. Koehn, C. Monz, J. Schroeder // Proc. Second Workshop on Statistical Machine Translation. – Prague, 2007. – P. 136–158.

29 Costa-Jussa, M. R. Study and comparison of rule-based and statistical Catalan-Spanish machine translation systems / M. R. Costa-Jussa, M. Farrus,

J. B. Marino, J. A. R. Fonollosa // *Computing and Informatics*. – 2012. – Vol. 31. – P. 245–270.

30 Doyon, J. Automated Machine Translation Improvement through Post-editing Techniques: Analyst and Translator Experiments / J. Doyon, C. Doran, C. D. Means, D. Parr // *Proceedings of AMTA, 2008*. – P. 346–353.

31 Hartley, A. Readability and Translatability Judgments for ‘Controlled Japanese’ / A. Hartley, M. Tatsumi, H. Isahara, K. Kageura, R. Miyata // *Proceedings of the 16th EAMT Conference*. – Trento, Italy, May 2012. – P. 237–244.

32 He, W. Improved neural machine translation with SMT features / W. He, H. Wu, H. Wang // *Thirtieth AAAI conference on artificial intelligence*. – 2016. – P. 151–157.

33 Kittredge, R. Sublanguages and Controlled Language / R. Kittredge // *The Oxford Handbook of Computational Linguistics*. – Oxford, 2003. – P. 430–447.

34 Koehn, P. Manual and automatic evaluation of machine translation between european languages / P. Koehn, C. Monz // *Proceedings on the Workshop on Statistical Machine Translation, 2006*. – P. 102–121.

35 Nagao, M. A New Method of N-gram Statistics for Large Number of n and Automatic Extraction of Words and Phrases from Large Text Data of Japanese / M. Nagao, S. Mori // *Proceedings of the 15th International Conference on Computational Linguistics (COLING 1994)*. – Kyoto, Japan, 1994. – P. 611–615.

36 O'Brien, S. Controlling Controlled English: An Analysis of Several Controlled Language Rule Sets / S. O'Brien // *Joint conference combining the 8th International Workshop of the European Association for Machine Translation and the 4th Controlled Language Applications Workshop (EAMT/CLAW)*. – Dublin City University, May 2003. – P. 105–114.

37 Papineni, K. BLEU: a Method for Automatic Evaluation of Machine Translation / K. Papineni, S. Rouskos, T. Ward, WJ. Zhu // *40th Annual Meeting of the Assoc for Computational Linguistics*. – Philadelphia, 2002. – P. 311–318.

38 Reuther, U. Two in one – Can it work? Readability and Translatability by means of Controlled Language / U. Reuther // Proceedings of EAMT-CLAW03, Controlled Language Translation (Dublin), 2003. – P. 124–132.

39 Sennrich, R. Linguistic input features improve neural machine translation / R. Sennrich, R. Haddow // Research papers of the first conference on machine translation. – Berlin, 2016. – Vol. 1 – P. 83–91.

40 Tripathi, S. Approaches to machine translation / S. Tropathi, J. K. Sarkhel // Annals of Library and Information Studies, 2010. – Vol. 57. – P. 388–393.

41 Underwood, N.L. Translatability Checker: A Tool to Help Decide Whether to Use MT / N. L. Underwood, B. Jongejan // Proceedings of MT Summit VIII (Santiago de Compostela, Spain), 2001. – P. 125–133.

42 Vilar, D. Error Analysis of Statistical Machine Translation Output / D. Vilar, J. Xu, L. F. D’Haro, H. Ney // Proceedings of the Fifth International Conference on Language Resources and Evaluation (LREC), 2006. – P. 697–702.

43 Zhou, J. Deep recurrent models with fast-forward connections for neural machine translation / J. Zhou, Y. Cao, X. Wang et al. // Computational linguistics, 2016. – Vol. 4. – P. 371–383.

44 Агентство Лангуст: <http://www.langust.ru/etc/history.shtml>.

45 Google Переводчик: <https://translate.google.ru>.

46 Онлайн-переводчик ПРОМТ: <https://www.translate.ru>.

47 Онлайн-словарь АБВУ Lingvo.Pro: <http://lingvopro.abbyyonline.com/ru>

48 Словарь МультиТран: <http://www.multitran.ru>

49 Микро, макро и мегаэкономика. Высшая школа экономики. <http://ecopos.ru/mikro,-makro-i-megaekonomika.php> (дата обращения 20.02.2019).

50 Luong, M. Effective approaches to attention-based neural machine translation / M. Luong, H. Pham, C. Manning. – 2015. – URL: <https://arxiv.org/abs/1508.04025> (дата обращения: 25.11.2018).

51 Muscan, M. On the translatability of cultural specifics in economic texts. URL: https://www.researchgate.net/publication/298973806_On_the_translatability_of_cultural_specifics_in_economic_texts (дата обращения 19.02.2019).

52 Siranush, M. Translation Peculiarities of Economic Texts. URL: http://publications.yasu.am/wp-content/uploads/2015/12/Siranush_Meloyan.pdf (дата обращения 19.02.2019).

53 Turovsky, B. Found in translation: More accurate, fluent sentences in Google Translate / B. Turovsky // Blog, 2016. – URL: <https://www.blog.google/products/translate/found-translation-more-accurate-fluent-sentences-google-translate/> (дата обращения: 18.11.2018).

54 Wu, Y. Google's Neural Machine Translation System: Bridging the Gap between Human and Machine Translation / Y. Wu, M. Schuster, Z. Chen, Q. V. Le, M. Norouzi, W. Macherey, M. Krikun, Y. Cao, et al. – 2016. – URL: https://www.researchgate.net/publication/308646556_Google's_Neural_Machine_Translation_System_Bridging_the_Gap_between_Human_and_Machine_Translation (дата обращения 20.11.2018).