

Ефимова Ольга Васильевна

канд. филол. наук, доцент

Институт международных отношений

ФГБОУ ВО «Пятигорский государственный университет»

г. Пятигорск, Ставропольский край

ФУНДАМЕНТАЛЬНЫЕ ОСНОВЫ СОВРЕМЕННОГО МАШИННОГО ПЕРЕВОДА В СВЕТЕ ОБЩЕЙ ТЕОРИИ ПЕРЕВОДЧЕСКОЙ ДЕЯТЕЛЬНОСТИ

Аннотация: в статье рассматриваются вопросы зарождения и становления современного машинного перевода, проблемы автоматической обработки текста, задачи анализа, синтеза и моделирования языковой структуры. Также раскрывается сущность трансформационной переводческой модели как теоретической составляющей машинного перевода.

Ключевые слова: прикладная лингвистика, машинный/автоматический перевод, общая теория перевода, трансформационная модель перевода.

На настоящем этапе развития цивилизации переводческая деятельность, несомненно, является одной из самых продуктивных и востребованных. Среди объективных факторов такой популярности следует упомянуть процессы глобализации и интеграции, а также укрепление международных контактов и сотрудничества. Практическая значимость данного вида активности обуславливает научный интерес и стремление описать и изучить теоретическую составляющую рассматриваемого явления.

Сегодня объектом внимания лингвистов становится специфика перевода текстов разных жанров, в особенности научно-технических [1], кроме того, предпринимаются попытки использовать основные понятия теории перевода в процессе обучения иностранному языку [4]. Данная статья посвящена анализу теоретических аспектов наиболее популярного вида переводческой деятельности, а именно, автоматического перевода.

В сфере перевода компьютеры стали широко использоваться приблизительно с 1980-х годов, но не совсем тем образом, который предполагался изначально. С начала проведения разработок в области машинного перевода считалось, что компьютеры составят конкуренцию людям-переводчикам. Однако с разработкой и развитием систем машинного перевода они заполнили тот сегмент, который не затрагивал интересов профессиональных переводчиков. Напротив, в лице компьютеров и соответствующих программных продуктов переводчики, продолжая играть ключевую роль в процессе создания перевода, обрели надежных и продуктивных помощников. На настоящем этапе отношения компьютеров и переводчиков можно охарактеризовать как взаимодействие, а не конкуренцию.

Для того, чтобы понять основные принципы функционирования современных систем компьютерного перевода, прежде всего, следует рассмотреть теоретические основы данного вида деятельности. Под переводом в нашей работе понимается «лингвистический процесс, межъязыковое преобразование или трансформация текста на одном языке в текст на другом языке; средство обеспечить возможность общения или коммуникации между людьми, говорящими на разных языках» [6, с. 3] Широко известны определения этого рода деятельности В.Н. Комиссарова, Ю. Найды, Федорова, А.Д. Швейцера. Некоторые ученые подчеркивают также не просто важность передачи смысла с исходного языка на язык перевода, а возможность содержательного и стилистического осмыслиния оригинала в процессе перевода.

Системный анализ переводческой практики и понятийной сферы современного переводоведения позволил сформировать единую типологизацию переводов. Существующая типологизация построена с учетом основных компонентов переводческой деятельности как разностороннего процесса в целом и своей целью имеет обобщение различных аспектов подготовки, выполнения, презентации и функционирования перевода. В современной науке о переводе выделение различных типов перевода происходит на основании определенных параметров, важнейшими из которых являются:

-
- основные функции перевода;
 - тип адекватности;
 - полнота и тип передачи смыслового содержания оригинала;
 - соотношение языка перевода и оригинала по типу;
 - форма презентации текста перевода и текста оригинала;
 - жанрово-стилистические особенности;
 - характер соответствия текста перевода тексту оригинала;
 - тип переводческой сегментации;
 - характер субъекта переводческой деятельности и его отношение к автору переводимого текста и др [6].

На основании предложенных принципов выделяются известные нам письменный и устный перевод, в рамках которого различаются последовательный и синхронный, односторонний и двухсторонний. Разная степень квалификации специалиста предполагает пословный, пофразовый, абзацно-фразовый и цельнотекстный переводы. Интересными представляются такие разновидности межъязыкового перевода, как интерсемиотический перевод, выполняющийся с естественного языка на искусственный или наоборот, а также трансмутация, предполагающая перевод с одного искусственного языка на другой. Среди многообразия жанрово-стилистической принадлежности выделяются общественно-политический, военный, юридический, а также художественный и научно-технический виды перевода.

Приоритетный интерес для нашего дальнейшего исследования представляют переводы, выделяемые по общей характеристике субъекта переводческой деятельности и по его отношению к автору переводимого текста. Эти виды представлены традиционным переводом, который еще называют человеческим или ручным, машинным (автоматическим) переводом и смешанным видом перевода, с использованием значительной доли традиционной или машинной переработки текста.

Современная наука о языке расширяет свои границы исследования, и одним из важнейших направлений ее развития становится прикладная лингвисти-

ка, в сферу главных задач которой входит решение и оптимизация традиционных лингвистических задач, как, например, перевод или обучение языкам. Данная наука носит практический, междисциплинарный характер и решает конкретные экспериментальные задачи в обучении языкам – в лингводидактике; при статистическом анализе языка и текстов, занимаясь построением и использованием структурно-вероятностных моделей языков в квантитативной математической лингвистике; при исследовании текстов на незнакомом коде, или языке, для получения информации – т.е. дешифровке, а также в ряде других вопросов.

Сам термин «прикладная лингвистика» в российской науке употребляется в достаточно широком понимании и отражает практические цели в многообразных ситуациях, таких, например, как машинный перевод, техническая коммуникация, распознавание и синтез речи, информационный поиск и др. Что касается зарубежных ученых, изначально термин «Applied Linguistics» трактовался достаточно узко, в применении по отношению к сфере подготовки преподавателей иностранных языков, изучению методики лингводидактики. Однако такое положение вещей изменилось с течением времени и в данный момент зарубежные и отечественные подходы к этой дисциплине практически идентичны.

Ряд исследователей, в том числе известный отечественный учёный В.А. Звегинцев, [2] считают, что настоящий период развития языкознания характеризуется серьезным влиянием прикладной лингвистики на общую лингвистическую теорию. И действительно, прикладная лингвистика представляет собой ту сферу, где реально проводятся лингвистические эксперименты, имеющие целью верификацию положений теории языкознания и проверку эффективности лингвистических продуктов, создаваемых разработчиками.

Особенно подобные процессы заметны в русле молодой отрасли научного знания под названием *компьютерная лингвистика*, которая сначала была известна как математическая лингвистика, потом структурная лингвистика и вычислительная лингвистика. Данная дисциплина изучает лингвистические основы информатики и все аспекты связи языка и мышления, моделирования языка

и мышления в компьютерной среде с помощью определенных программ. Фундаментальные интересы этой науки лежат в области оптимизации коммуникации на основе лингвистических знаний, создания естественно-языкового интерфейса и технологий понимания языка для общения человека с машиной в рамках решения основных проблем искусственного интеллекта, а также создания и моделирования информационных компьютерных систем. В ее сферу деятельности входят:

- машинный перевод (Machine Translation);
- компьютерная лексикография (подготовка электронных словарей);
- компьютерная лингводидактика (CALL/T – Computer Assisted Language Learning and Teaching);
- корпусная лингвистика;
- автоматическая обработка естественных языков (NLP – Natural Language Processing).

Большинство ученых в этой области придерживаются такой точки зрения, что основным направлением компьютерной лингвистики являются проблемы автоматической обработки языка, лежащей в основе автоматического перевода, который включает задачи анализа и моделирования языковой структуры, а именно:

- графематический/фонематический анализ языка;
- морфологический анализ;
- лексико-грамматический анализ языка;
- синтаксический анализ, или парсинг;
- анализ и моделирование семантической структуры;
- задачи синтеза языковых элементов, в т. ч. генерации текстов.

Для нашей работы наибольший интерес представляет лишь один из аспектов изучения компьютерной лингвистики, а именно машинный перевод.

В переведоведении машинный перевод известен также как автоматический перевод, machine translation, mechanical translation, automatic translation, MT.

Рассмотрим существующие дефиниции данного вида переводческой деятельности. Так, согласно Лингвистическому энциклопедическому словарю под редакцией В.Н. Ярцевой, *автоматический (машинный) перевод* – это «выполняемое на ЭВМ действие по преобразованию текста на одном естественном языке в эквивалентный по содержанию текст на другом языке, а также результат такого действия» [3, с. 15].

Как известно, для осуществления автоматического перевода в компьютер не необходимо установить программу, определенный алгоритм, затем словари исходного и переводного языков, при этом они должны содержать разноуровневую лингвистическую информацию морфологического, синтаксического, лексического и на современном этапе, даже фонетического характера. Затем автоматически выполняется последовательность формальных операций, направленных на анализ и синтез информации. Весь процесс сводится к [7]:

- вводу текста и поиску входных словоформ в словаре с сопутствующим морфологическим анализом;
- переводу устойчивых выражений и фразеологизмов;
- определению основных грамматических, лексических и семантических признаков, необходимых для перевода в рамках данной пары языков;
- анализу омографов и многозначных слов с учетом контекста;
- затем финальному грамматическому анализу;
- и в итоге – синтезу словоформ, предложений и текста в целом на ПЯ.

В различных существующих на сегодняшний день системах машинного перевода анализ может проходить как на уровне отдельных фраз, так и на уровне всего текста.

В начале своего создания автоматизированные системы перевода были ориентированы на поиск соответствий на поверхностном уровне, теперь же современные аналоги производятся с учетом возможности использования глубинных уровней представления смыслового содержания текста. Качество данного вида переводческой деятельности зависит от объема, структуры и качества общелингвистической информации к лексическим единицам входного и вы-

ходного языков, настройки алгоритмов анализа на специфику переводимых текстов – как правило, перевод с помощью современных компьютерных технологий ориентирован на тексты научно-технического или официально-делового характера.

Теоретической базой начальных разработок в области автоматического перевода стал взгляд исследователей на язык как на систему определенных кодов. С дальнейшим развитием прогресса и научной мысли возникла идея множественности вариантов анализа и синтеза информации. Следует также отметить, что машинный перевод как деятельность стимулировал исследования по теоретическому языкознанию в аспекте различения языка и речи, теории формальных грамматик, статистическом и информационном измерении речевых характеристик, а также в таких областях знания, как информационные технологии и искусственный интеллект.

Разработкой проблематики и возможностей автоматического перевода занимались как зарубежные исследователи – Д.Г. Хэйз (1967), И.С. Батори (1977), Н. Мейровиц (1982), Г. Брудерер (1978), У.Дж. Хитчинс (1986), У. Вивер, Б. Сигурд, так и отечественные ученые. В частности в нашей стране первый опыт в этой области принадлежит И.К. Бельской, занимавшейся лингвистическими основами построения алгоритма для машинного перевода, и Д.Ю. Панову, воплотившему в реальность программу реализации проекта (1954). Следует также отметить исследования А.А. Ляпунова, Н.П. Трифонова и А.Н. Королева (1955), И.А. Мельчук, Ю.Д. Апресян (1960), Р.Г. Пиотровского (1979), Р.Г. Котова (1983), Л.Л. Нелюбина (1983), Ю.Н. Марчук (1985), Н.К. Рябцевой (1986), В.М. Андрющенко (1985), Г.Г. Белоногова, Б.А. Кузнецова (1984), А. Борковского, Г. Хельбиг (1984).

Очевидно, что начало разработок и исследований этой проблематики было обусловлено появлением первых компьютеров в 50-х годах 20 века. Фундаментом работы первых автоматических систем стал пословно-пооборотный метод на основе достаточно простого алгоритма, учитывающего в основном синтаксические особенности структуры, при этом семантическая и грамматическая парадигма ис-

ходных единиц не учитывалась, что приводило к значительному количеству ошибок и низкому качеству конечного перевода. В то время об учете различного рода экстралингвистической информации не было и речи.

Хотя в начале зарождения идеи об автоматическом переводе звучало мнение о том, что машины в будущем полностью заменят людей-переводчиков, по прошествии некоторого времени в связи с возникшими определенными трудностями и не слишком высоким качеством конечного продукта ряд ученых высказали мнение о том, что сфера применения машинного перевода будет ограничена так называемым промышленным переводом. Б.Сигурд назвал это направление *переводом для специальных целей*, и подчеркнул, что машинный перевод не разрабатывается для перевода художественных текстов, поэзии или Библии. Как показывает практика, такой вид переводческой деятельности может быть эффективен в рутинной переводческой работе, когда быстрота выполнения перевода важнее его качества. Так, специалист в какой-либо области, не владеющий тем или иным иностранным языком, но следящий за последними зарубежными разработками в своей области, может воспользоваться машинным переводом, чтобы понять, нужен ли ему впоследствии этот источник, заказывать ли более качественный перевод или предпринять попытку отредактировать уже полученный. С.В. Тюленев отмечает, что ценность машинного перевода заключается в возможности получить то, что он условно называет сигнальным переводом.

Переводческая деятельность осуществляется на основе определенных моделей перевода, и перевод, выполняемый машиной, не является исключением. Под моделью перевода мы будем понимать условное изображение процедуры осуществления процесса перевода, основанное на попытке распространить на перевод некоторые общие постулаты языкознания и психологии, а в случае с автоматическим переводом – еще математики и информатики. В современной науке о переводе наиболее известными являются ситуативная, семантическая и трансформационная модели перевода. Наибольшей популярностью пользуется трансформационная модель, на основе которой базируется так называемая

трансформационная теория перевода, функционирующая в современном переводоведении наряду с общелингвистической и денотативной теорией.

Автоматический перевод основывается на трансформационном методе, в связи с этим нам представляется целесообразным рассмотреть сущность трансформационной теории перевода более подробно.

Данная идея и разработка самой теории осуществлялись представителями американской школы перевода Ю. Найдой, и нашли свое отражение в более поздних работах О. Каде и В. Коллера. Основой трансформационной теории является генеративная, или порождающая грамматика Наома Хомского. В отечественной науке теоретическим фундаментом для систем машинного перевода стала формальная грамматика Ф.Ф. Фортунатова. В структуре самой модели лежит взгляд на деятельность переводчика как на работу определенной системы, словно на входном этапе берется оригинал, а на выходе уже имеется перевод как продукт, в целом, переводческая деятельность представлена преобразованием или трансформацией оригинала в текст перевода. Трансформационная грамматика рассматривает правила порождения синтаксических структур, которые характеризуются общностью лексем и основных логико-синтаксических связей. Например: «*Девочка сорвала цветок*», «*Цветок был сорван девочкой*», «*Сорванный девочкой цветок*», «*Срывание цветка девочкой*». Трансформационная грамматика не просто соотносит друг с другом структуры, связанные правилами трансформации. Она считает некоторые из таких структур исходными, т.е. «ядерными», а остальные структуры, так называемые «трансформы», выводит из небольшого числа ядерных структур или, наоборот, сводит к ядерным структурам, соответственно, в вышеприведенном примере ядерной считается структура «деятель-действие», а остальные структуры являются производными от нее.

Общая теория перевода использует положения генеративной грамматики в двух различных аспектах. Во-первых, сам процесс перевода приравнивается к процессу межъязыковых трансформаций. Во-вторых, предполагается наличие полных эквивалентных, или так называемых ядерных структур в обоих языках,

и в этом случае переводческая деятельность заключается в поиске идеальных в данной конкретной ситуации эквивалентов.

Согласно этой модели, сам процесс перевода осуществляется в три этапа. Сначала, производная структура в оригинале соотносится с ядерной структурой в исходном языке, это непосредственно анализ или обратная трансформация: *He is a poor public speaker – He speaks in public poorly*. На следующем этапе происходит переход от ядерной структуры языка оригинала к аналогичной ядерной структуре языка перевода, этот процесс носит название перенос: *Он говорит на публике плохо*. На заключительном этапе ядерная структура в языке перевода преобразуется в производную в соответствии с нормой языка: *Он говорит на публике плохо – Он не умеет выступать публично*. Этот процесс называется синтезом или реструктурированием.

Трансформационная модель является весьма продуктивной для анализа переводческой деятельности в целом, ведь рассматривая формальные единицы оригинала и перевода как взаимосвязанные трансформы, трансформационная теория тем самым отводит важное место сопоставительному изучению разноязычных форм, между которыми могут устанавливаться отношения переводческой эквивалентности. Подобный подход создает теоретическую базу для описания системы переводческих отношений двух конкретных языков, что составляет одну из главных задач переводоведения. На основе сопоставления исходных и конечных форм переводческих преобразований оказывается возможным выявить различные типы трансформации при переводе, именно так переводчик приобретает знание так называемых переводческих приемов, которые широко используются на практике.

Однако у трансформационной модели имеются и некоторые недостатки. Если учитывать лишь грамматические категории поверхностных структур, их использование является одним из основных источников буквализма при переводе, когда копируется план выражения оригинала и механически переносится в язык перевода, здесь ярко проявляется неоднозначность поверхностных структур исходного языка. В связи с этим, в новейших исследованиях транс-

формационная модель дополняется семантической моделью перевода, базирующейся на процедуре компонентного анализа, который позволяет разбивать значения языковых единиц на более мелкие элементарные частицы – семы. В этом случае в процессе перевода вначале определяется семный состав отрезка оригинала, и выделяются релевантные для коммуникации семы, а затем происходит подбор единиц, в значения которых входят наибольшее количество релевантных сем.

Итак, трансформационная модель легла в основу зародившегося в середине 20 века машинного или автоматического перевода, базирующегося на максимальной формализации процесса распознавания и порождения языковых знаков, ведь именно на таком принципе может работать машина. Ряд ученых, в частности, И.И. Ревзин и В.Ю. Розенцвейг, [5] в своих работах заострили внимание на принципиальном отличии автоматизированного перевода от того же процесса, осуществленного живым человеком. Ими даже было предложено два принципиально разных термина: действия, осуществляемые человеком, предлагалось называть интерпретацией, так как человек в своей деятельности производит анализ различных экстралингвистических факторов, в то время, как автоматический процесс предлагалось называть непосредственно переводом, потому что эти операции сводятся к преобразованию языковых единиц и структур исходного языка в соответствующие им единицы и структуры языка перевода с учетом определенных принципов и правил, но без обращения к каким-либо внеязыковым данным.

Список литературы

1. Ефимова О.В. О способах перевода словосокращений и аббревиатур в текстах-инструкциях [Текст] / О.В. Ефимова // «Philology». – Волгоград: Научное обозрение, 2016. – №2 (2). С. 8–10.
2. Звегинцев В.А. Мысли о лингвистике [Текст] / В.А. Звегинцев. – ЛКИ, 2008. – 336 с.
3. Лингвистический энциклопедический словарь [Текст] // Главный редактор Ярцева В.Н. – М.: Советская энциклопедия, 1990. – 683 с.

4. Михиенко С.А. Использование понятия эквивалентности перевода в процессе обучения иностранному языку (на примере подготовки студентов-конфликтологов) [Текст] / С.А. Михиенко // Современные образовательные технологии в мировом учебно-воспитательном пространстве. – Новосибирск, 2016. – №5. – С. 84–88.
5. Ревзин И.И. Основы общего и машинного перевода [Текст] / И.И. Ревзин, В.Ю. Розенцвейг. – М.: Изд-во Высшая школа, 1964. – 243 с.
6. Утробина А.А. Основы теории перевода. Конспект лекций [Текст] / А.А. Утробина. – М.: Приор-издат, 2006. – 144 с.
7. Энциклопедия-Он-Лайн («Кругосвет») [Электронный ресурс]. – Режим доступа: http://krugosvet.ru/enc/gumanitarnye_nauki/lingvistika/