

Секция 2

ПЕРЕВОД В АСПЕКТЕ МЕЖКУЛЬТУРНОЙ КОММУНИКАЦИИ

М. Г. Богова, А. А. Кравченко (Минск, МГЛУ, Беларусь)

УСТРАНИМЫ ЛИ ПРОБЛЕМЫ МАШИННОГО ПЕРЕВОДА?

Прежде всего необходимо понять, как работают системы машинного перевода (МП).

Для осуществления машинного перевода на вычислительное устройство устанавливается специализированное программное обеспечение с соответствующим *алгоритмом перевода*, под которым понимается последовательность строго определенных действий над текстом с целью выявления переводных соответствий в имеющейся паре языков.

На большинстве устройств используется следующий алгоритм машинного перевода: ввод текстовых данных и поиск входных словоформ; перевод идиоматических словосочетаний, фразеологизмов, штампов, присущих языку или данной предметной области; финальный грамматический анализ; синтез выходных словоформ и предложений в целом на языке-цели.

Существует несколько методов перевода: система может базироваться на лингвистических правилах, когда слова из языка-источника сопоставляются с подобными в целевом языке. Имеется утверждение, что для решения проблемы МП сначала нужно понять текст на естественном языке. Такой метод используется посредством символического представления, в основе которого лежит текст на языке перевода. Если учитывать посредников, то можно говорить о *интерлингвистическом МП* или *трансфертном переводе*. Для качественного перевода системы должны обладать большим количеством словарей, включающих морфологическую, синтаксическую и семантическую информацию, а также полный набор используемых необходимых правил.

Все системы МП в целом делятся на три основные группы: основанные на правилах, базирующиеся на примерах и синтаксические [1; 2]. В основе систем, основанных на правилах, лежит лингвистическая информация о принятых в языке-источнике и языке-цели правилах. Этот метод называется классическим: на основе исходных данных текст переводится последовательно, минимальной единицей при этом является предложение.

Системы, базирующиеся на примерах, представляют собой совокупность систем, основанных на принципе «связь структуры входного и выходного предложения». Они подразделяются на три категории: системы пословного перевода (выдают продукт низкого качества, но используются для перевода длинных списков или подстрочника для ТМ-систем – translation memory systems), трансфертные и интерлингвистические системы. В процессе перевода имеется посредник, через которого проходит смысл исходного выражения. Разница между этими двумя способами в том, что в трансфертных системах посредник зависит от пары языков,

участвующих в процессе перевода [3]. Любая трансфертная система состоит из пяти частей: морфологического анализа; лексической категоризации; лексического трансфера; структурного трансфера; морфологической генерации [4].

Интерлингвистический машинный перевод является одним из традиционных. При этом исходный текст переводится в абстракцию, не зависящую от языковой системы, а переводной текст базируется на этой абстрактной системе. Ключевое преимущество данного подхода заключается в том, что новый язык в систему добавляется исключительно математическим способом – вносятся правила языка и на основе смысла машина интерпретирует текст с одного языка на другой. Однако, как бы хорошо это не звучало в теории, данный подход не реализован пока даже для двух языков.

Для сравнительного анализа перевода были выбраны образцы художественного текста (отрывок из новеллы Г. Хилл «Leaving Bill») и деловой документации (фрагмент из годового отчета японской компании Stanley Electric).

Одной из главных целей художественной литературы является доставление эстетического удовольствия читателю, а также донесение скрытых смыслов, заложенных автором произведения. Каждое предложение направлено на создание у читателя определенного эмоционального настроения. В данном случае обе задачи машинным переводом оказались не выполнены, напр.: *Темно-зеленая дверь с ее латунным молотом сделала дом **надежным и надежным** местом, чтобы создать дом и создать семью.*

Для того, чтобы передать идею, заложенную в предложении, используется генерализация и добавление более сложных синонимов слова *создавать*: *Медный молоток на темно-зелёной двери придавал дому солидный и надёжный вид. Безопасное место для того, чтобы **наладить быт и завести семью.***

Сравните машинный перевод: *Прижавшись к мраку, она направилась к главной улице* и традиционный перевод: *Собравшись с духом, она направилась в сторону главной улицы.*

Еще один пример ошибки машинного перевода: *Дождь запотевал воздух, и Труда поправила свой платок, не желая, чтобы сырость отменила несколько часов в бигуди, которые она перенесла этим утром.* Здесь помимо генерализации и добавления эпитетов, использован прием аллитерации. Традиционный перевод: *Морозящий дождь создавал туман, и Труда поправила платок на голове, не желая, чтобы сырость свела на нет все ее много-часовые мучения с бигуди этим утром.*

Отсутствие у машины фоновых знаний привело к неверному переводу фамилии английского писателя Ч. Диккенса: *Дом выглядел теплыми и привлекательными, как что-то из рождественской сцены в романе Дикена.* Их наличие не позволило бы допустить такую ошибку в традиционном переводе: *Здание выглядело таким теплым и манящим, будто из Рождественской сцены романа Диккенса.*

Машина выполняла перевод каждой языковой единицы, в итоге предложения превращались в набор слов, не связанных по смыслу: *Клубничные конфеты, топы большого пальца, зимняя смесь, розовые креветки, летающие тарелки, жевательные орехи и т.д.* При традиционном переводе выполняется поиск информации об английских сладостях, переводчик может посетить личные сайты и кондитерские магазины онлайн, чтобы найти русские аналоги упомянутых конфет, а при их отсутствии он применит калькирование: *Клубничные конфеты, фруктовые леденцы, набор из мятных, травяных и ментоловых сладостей, жевательные креветки, лимонные вафли с щербетом, орешки в шоколаде...*

Пример выбора не подходящего по смыслу перевода лексической единицы *indulgence*: *Труд был назван в результате его **снисхождения**.* Для придания нужного эстетического эффекта переводчик найдет в словарях приемлемое значение: *Труде дали такое имя как раз с его **разрешения**.*

Можно приводить много других примеров, где ошибками машинного перевода являются нарушение порядка слов, игнорирование контекста, допущение аллитерации, необходимость поиска эпитетов.

Несколько иначе обстоят дела с машинным переводом деловой документации, для которой характерно наличие однозначных клишированных конструкций, хотя используются сложные предложения с большим количеством причастных и деепричастных оборотов. Для отчетности в финансовом документе в деталях описывается каждое совершенное компанией действие, потеря при машинном переводе ключевого значения термина или отсутствие перевода как такового может привести к серьезным последствиям при работе с иностранными предприятиями. Пример: в одном из предложений слово *back-office* осталось непереведенным, хотя его составляющие были заложены в программе: *путем тщательной ликвидации отходов в производственном процессе и back-office*. В аутентичном переводе, данный термин передан транслитерацией и сопровождается кратким определением: *...путем тщательной утилизации отходов производства и создания **бэк-офисов (отделы, департаменты корпорации, которые выполняют рутинные, административные, обслуживающие функции управления бизнесом как таковым)**.*

Часто программа дублировала перевод различных лексических единиц в каждом предложении, что усложняло понимание текста в целом: *В таких бизнес-условиях Stanley Electric и ее **дочерние компании** прилагают усилия, в рамках которых **дочерние компании** работали над разработкой продуктов...*

*Ожидается, что в Азии внутренний спрос будет оставаться устойчивым, и **ожидается**, что экспорт продолжит расти незначительно на фоне спроса в развитых странах.*

*Что касается зарубежной экономики, то США и Европа **умеренно выздоравливали**, в то время как Азия в целом продолжала **умеренно развиваться**...*

При традиционном переводе используется прием опущения и добавления, найдены синонимы слова умеренно: *В такой экономической обстановке «Stanley Electric» и её дочерние предприятия прилагали все усилия, в рамках которых эти компании занимались разработкой продуктов...*

Предполагается, что в Азии внутренний спрос останется устойчивым, а также, что экспорт продолжит умеренно расти на фоне спроса в развитых странах. С точки зрения зарубежной экономики, в США и Европе произошёл относительно небольшой подъем, в то время как Азия в целом продолжала показывать стабильный рост.

Отметим дословный машинный перевод экономических терминов, что приводило к потере смысла лексической единицы в целом: *Ожидается, что японская экономика останется умеренной экспансионистской тенденцией, несмотря на падение цен на акции, ср.: Ожидается, что японская экономика продолжит придерживаться умеренной тенденции по расширению своих объемов, несмотря на падение цен на акции.*

В соответствии с этим бизнес-клиентом дочерние компании полны решимости добиться ощутимых результатов от своей деятельности в области инноваций в производстве... , ср.: При таком деловом климате, объединённые дочерние компании твёрдо намерены добиться ощутимых результатов от производственных инноваций... ”.

Перечислим типичные ошибки машинного перевода деловой документации:

1) общие ошибки: самые частотные – это морфологически ошибки в числе, падеже, времени, наклонении, степеней сравнения прилагательных и наречий, предлогах, имёнах собственных;

2) ошибки перевода сложных предложений с двумя и более оборотами процент потери ядерного значения очень высок;

3) синонимичные повторы, тавтология;

4) перевод каждой отдельной лексической единицы;

5) отсутствие эстетической составляющей;

6) незнание сложной терминологии.

Сегодня машинный перевод без редактуры является быстрым инструментом для создания общего представления о том, что исходный фрагмент на иностранном языке призван донести до читателя. С переводческой точки зрения финальный вариант, который выдает программа, чаще всего является некорректным. Совет – отказаться от полного машинного перевода в пользу автоматизированного, позволяющего контролировать и вносить правку в процессе перевода, либо осуществлять постредактирование. Его удобнее всего делать в окне перевода программы, так как любые предложения по улучшению перевода учитываются и прорабатываются разработчиками системы, они заносятся в базу программы. Именно так происходит улучшение моделей перевода.

Можно порекомендовать программу AfterScan Antique, а во время постредактирования – использовать переводной словарь Reverso Context. Данная программа позволяет эффективно перевести слова и выражения,

совершая глобальный поиск и используя несколько полноразмерных словарей, и затем предоставляет перевод и примеры употребления практически любой лексической единицы или выражения.

Как видно, ошибки в переводе, выполняемом машиной, достаточно многочисленны. Рекомендации по исправлению и улучшению машинного перевода сводятся к использованию различных дополнительных переводческих программ, а также «помощи» при обучении модели компьютерной программы перевода. Имея дело с важными документами и художественными текстами, перевод, осуществленный с помощью программного обеспечения, следует считать первым предварительным вариантом. Ручное корректирование текста следует осуществлять в обязательном порядке.

ЛИТЕРАТУРА

1. *Баранов, А. Н.* Введение в прикладную лингвистику / А. Н. Баранов. – М. : Эдиториал УРСС, 2001. – 360 с.

2. *Совпель, И. В.* Инженерно-лингвистические принципы, методы и алгоритмы автоматической переработки текста / И. В. Совпель. – Минск: Вышэйшая школа, 1991. – 118 с.

3. *Исламов, Р. С.* Лексическая амбивалентность естественного языка в среде систем машинного перевода: дис. ... канд. филол. наук: 10.02.19 / Р. С. Исламов. – Кемерово, 2015. – 220 с.

4. *Новиков, В. А.* Трансфер в современных системах машинного перевода: дис. ... канд. филол. наук: 10.02.21 / В. А. Новиков. – Москва, 2001. – 194 с.