

Баканова М.В. Вопросы машинного перевода и обучение лингвистов-переводчиков работе с системами автоматизированного перевода. // Проблемы информатики в образовании, управлении, экономике и технике: Сб. статей XV Междунар. научно-техн. конф. – Пенза: ПДЗ, 2015. – С. 154-157.

УДК 81'33

ВОПРОСЫ МАШИННОГО ПЕРЕВОДА И ОБУЧЕНИЕ ЛИНГВИСТОВ-ПЕРЕВОДЧИКОВ РАБОТЕ С СИСТЕМАМИ АВТОМАТИЗИРОВАННОГО ПЕРЕВОДА

М.В. Баканова

THE QUESTIONS OF MACHINE TRANSLATION AND TRAINING LINGUISTS-INTERPRETERS WORKING WITH SYSTEMS OF AUTOMATED TRANSLATION

M. V. Bakanova

Аннотация. Раскрыта сущность, достоинства и виды систем машинного перевода. Определены этапы машинного перевода. Приведены примеры заданий по работе с системами автоматизированного перевода.

Ключевые слова: машинный перевод, качество перевода, on-line системы автоматизированного перевода.

Abstract. The essence, advantages and types of machine translation are revealed. The stages of machine translation are determined. The tasks for working with on-line systems of automated translation are given.

Keywords: machine translation, quality of translation, on-line systems of automated translation.

Машинный перевод (МП) составляет одну из центральных областей использования информационных технологий в лингвистике. Это обусловлено, с одной стороны, тем, что в машинном переводе сфокусированы все проблемы компьютерной лингвистики, а с другой стороны, тем, что современное общество диктует необходимость перевода текстов различной функциональной и профессиональной направленности.

Машинный перевод – выполняемое компьютером действие по преобразованию текста на одном естественном языке в эквивалентный по содержанию текст на другом языке, а также результат такого действия [3, с. 15].

Преимущества автоматизированного перевода очевидны:

- скорость машинного перевода гораздо выше, чем ручного, что экономит время при переводе больших объемов информации. Так, нормой научно-технического перевода считается время 10 дней на авторский лист. Система машинного перевода позволяет получить перевод сотен авторских листов за 1 час;

- стоимость машинного перевода гораздо ниже, чем перевод вручную;

- используя системы машинного перевода, можно перевести содержание любого веб-ресурса или запрос поисковой системы.

Это свидетельствует об актуальности обращения к проблеме машинного перевода, который, несмотря на то, что уступает по качеству перевода, выполненному человеком, всё равно помогает преодолевать языковые проблемы и оставаться интересной проблемой компьютерной лингвистики в целом.

Человек, как правило, участвует в корректировке машинного перевода и доведении его до удобочитаемого вида.

С точки зрения роли человека в процессе выполнения машинного перевода различают следующие его виды [1, с. 54]:

МАНТ (machine-assisted human translation) – перевод, осуществляемый человеком с использованием компьютера. Человек использует компьютерные программы для упрощения процесса перевода (электронные словари, терминологические базы данных и т.д.), но сам перевод выполняет человек.

НАМТ (human-assisted machine translation) – машинный перевод при участии человека. Здесь одинаково важно и участие человека. Роль человека сводится к выбору предлагаемых машиной рений и редактированию текста перевода. Сюда относятся системы переводческой памяти – Translation Memory. Основная идея таких систем состоит в сборе и хранении уже однажды переведенных переводчиком текстов, для того, чтобы в процессе перевода предлагать уже готовый перевод фразы или фрагмента текста. Такие программы избавляют переводчика от рутинной работы.

ФАМТ (fully-automated machine translation) – полностью автоматизированный машинный перевод. Данный тип перевода является наиболее сложным, так как основывается на учёте объективных законов функционирования языка и мышления. Основная проблема состоит в том, что эти законы еще не до конца изучены.

Схема машинного перевода включает следующие этапы [2, с. 80–81]:

1. Ввод в компьютер текста на ИЯ.
2. Морфологический анализ текста, который включает определение части речи и морфологических характеристик каждого слова.
3. Синтаксический анализ каждого предложения текста ИЯ, включающий поиск основных членов предложения и определение типов синтаксических связей между ними, выражаемых в виде дерева зависимостей или дерева непосредственных составляющих.
4. Семантический анализ каждого предложения ИЯ, в результате которого создаётся семантическое представление этого предложения, независимо от типа языка (общее и для ИЯ, и для ПЯ).
5. Синтаксический синтез предложений ПЯ, включающий создание предложений правильной синтаксической структуры, соответствующей правилам ПЯ и типу синтаксической структуры предложения на ИЯ.
6. Морфологический синтез каждого слова в составе отдельных предложений текста ПЯ, представляющий собой постановку слов ИЯ в нужных морфологических формах.
7. Вывод текста на ИЯ.

Для функционирования систем машинного перевода требуется лингвистическое, программное и информационное обеспечение систем МП.

Лингвистическим обеспечением выступают словари слов и словосочетаний с соответствующими признаками для ИЯ и ПЯ, морфологические таблицы суффиксов и окончаний, базы грамматических правил и т.д.

К программному обеспечению относятся программы выполнения перевода, ведения словарей, формирования базы правил и т.д.

Информационное обеспечение представляет база экстралингвистических знаний о предметной области.

К наиболее распространенным в России системам машинного перевода относятся: Stylus, Universal Translator, Socrat, Polyglossum, Promt, WebTranSite. Все они являются мультязычными системами машинного перевода, содержащими множество словарей по различным предметным областям. Программа WebTranSite представляет собой систему для перевода веб-страниц и сообщений компьютерных программ.

В рамках дисциплины "Компьютерное обеспечение перевода" для студентов направления "Лингвистика" практическая работа с различными системами машинного перевода проводится по следующим направлениям:

1. Выполнение автоматического перевода одного и того же текста с помощью двух или трех on-line систем машинного перевода (например, www.translate.ru, <https://translate.google.ru/>, <https://translate.yandex.ru/>) и представление результатов в виде таблицы.

2. Характеристика протестированных on-line систем машинного перевода по следующим параметрам:

- время, затраченное на перевод;
- качество перевода;
- необходимость постредктирования.

3. Редактирование полученных вариантов перевода с последующим анализом выполненной работы по критериям:

- количество лексических замен отдельных слов;
- количество лексических замен словосочетаний;
- количество удалений вариантов перевода слов;
- количество исправлений неверного согласования;
- количество исправлений неверного управления;
- частота вставки дополнительных слов;
- частота вставки дополнительных словосочетаний;
- изменение структуры предложения и т.д.

4. Перевод и сравнение переводов текстов различной функциональной направленности, выполненные в любой on-line системе машинного перевода.

Обучение использованию систем машинного перевода с последующим анализом проведенной работы представляет профессиональный интерес для будущих лингвистов-переводчиков, так как использование таких систем является частью их профессиональной компетенции, а также представляет собой актуальное направление компьютерной лингвистики.

Библиографический список

1. Всеволодова А.В. Компьютерная обработка лингвистических данных : учеб. пособие для студ., аспирантов, преподавателей-филологов. – М.: Флинта; Наука, 2007. – 96 с.
2. Зубов А.В., Зубова И.И. Информационные технологии в лингвистике : учеб. пособие. – М.: Академия, 2004. – 208 с.
3. Марчук Ю.Н. Проблемы машинного перевода. – М., 1983. – 112 с.

Баканова Марина Викторовна
Пензенский государственный
университет, г. Пенза, Россия
E-mail: bakanova_marina@list.ru

Bakanova M.V.
Penza state university,
Penza, Russia