



УДК 81'322.4
ББК 81.184

СРАВНИТЕЛЬНЫЙ АНАЛИЗ СИСТЕМ МАШИННОГО ПЕРЕВОДА

Арестова Анастасия Александровна

Студент кафедры теории и практики перевода,
Волгоградский государственный университет
n-arestova@mail.ru
просп. Университетский, 100, 400062 г. Волгоград, Российская Федерация

Аннотация. Настоящая статья посвящена исследованию онлайн-переводчиков Google-Translate, PROMT и Яндекс.Перевод. На примере фрагментов текстов различных жанров предлагается лингвистический анализ переводов в разных языковых направлениях. В ходе анализа выявляются типичные ошибки, допускаемые в машинном переводе, и определяются возможности и сферы использования каждой из переводческих систем. Устанавливается, что машинные переводчики пригодны для перевода текстов с более простым синтаксисом. Делается вывод о том, что профессиональные переводчики могут использовать эти системы для перевода определенных типов текстов, чтобы сократить время и оптимизировать процесс перевода.

Ключевые слова: информационные технологии, электронные словари, автоматизация перевода, машинный перевод, качество перевода.

Сегодня в нашу жизнь все больше внедряются информационные технологии. Компьютеры, ноутбуки и телефоны стали неотъемлемой частью нашей жизни. Мы используем эти устройства как в быту, так и на учебе или работе. Безусловно, переводчику необходимо не только владеть компьютером на уровне продвинутого пользователя, но также выполнять быстрый и качественный перевод. Несколько сотен тысяч переводов в год осуществляют переводчики и бюро переводов, и некоторые из них из-за большой загруженности не всегда могут справиться с постоянно возрастающим потоком заказов на перевод. Единственным способом увеличения скорости перевода является использование в переводческой деятельности компьютера, так как компьютер может переводить в миллиарды раз быстрее человека. Человек-переводчик тратит 20 % своего времени непосредственно на перевод, 40 % – на поиск незнакомых слов по словарю и/или параллельным текстам

и 40 % – на перепечатку и оформление перевода; компьютер в свою очередь при переводе тратит 95 % времени на перевод и 5 % на пополнение словаря [1, с. 78–79]. Если человек (при условии максимальной загрузки) в среднем переводит 4–5 авторских листов в месяц, то такая система машинного перевода как SYSTRAN переводит до 120 авторских листов в час. Эти количественные характеристики свидетельствуют о преимуществе машинного перевода. Однако качество такого перевода значительно уступает переводу, выполненному человеком [1, с. 79].

На сегодняшний день существует множество различных средств, которые способны оптимизировать процесс перевода. К таким средствам относятся электронные переводные словари, системы переводческой памяти, системы машинного перевода и многие другие [4, с. 68]. Многие профессиональные переводчики высказывают аргументы за и против использования машинного перевода, но

все они единогласны в том, что с помощью машинного перевода больших объемов текстов можно значительно сэкономить время, затрачиваемое на перевод.

Кристиан Буатэ выделяет три вида современного машинного перевода:

1) «информативный» машинный перевод. Это грубый машинный перевод, пословный, достаточный, чтобы ознакомиться с текстом на незнакомом языке;

2) «профессиональный» машинный 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» машинный перевод. Авторы текстов заранее избавляют их от неоднозначностей и работают в режиме диалога с компьютером. Таким образом можно значительно повысить качество перевода [3, с. 260].

Существует ряд различных систем машинного перевода, и некоторые из них справляются с переводом более качественно, а некоторые менее качественно. Но для того чтобы определить, какой из переводчиков справляется с переводом в той или иной языковой паре или с переводом определенного типа текстов, необходимо провести анализ качества перевода на основе отдельных фрагментов различных типов текстов.

Среди наиболее известных машинных переводчиков можно выделить следующие: «Google-Translate», «PROMT» и «Яндекс.Перевод». Работа переводчика «Google-Translate» основана на статистическом анализе: система подбирает эквивалент перевода, основываясь на частоте употреблений, и в итоге подставляет вариант, имеющий наиболее высокий процент совпадений. Также следует отметить, что изначально переводчик «Google-Translate» создавался для перевода текстов с английского языка на русский язык и с русского языка на английский язык, и до сих пор данная система при переводе на другие языки использует английский язык в качестве языка-посредника, за счет чего качество перевода может заметно пострадать, так как зачастую переводчик «Google-Translate» при переводе на другие языки сохраняет син-

таксис исходного языка. Автоматический переводчик «PROMT» действует по принципу «перевода по правилам», то есть работает по алгоритму, в соответствии с которым система анализирует текст на исходном языке и на основе проведенного анализа синтезирует перевод [4, с. 3]. Переводчик «Яндекс.Перевод» так же, как «Google-Translate», основан на статистическом переводе. Система самостоятельно строит свои словари на основе анализа миллионов переведенных текстов. При переводе из словарей, с учетом контекста, выбираются наиболее подходящие, по мнению программы, варианты [2].

Сопоставим примеры переводов, выполненные системами «Google-Translate», «PROMT» и «Яндекс.Перевод» с немецкого языка на русский. При переводе фрагмента текста технического характера – инструкции по эксплуатации электронного устройства (прибора для измерения кровяного давления) – были получены следующие результаты:

Оригинал:

Wenn Sie mehrere Messungen an einer Person durchführen möchten, warten Sie zwischen den einzelnen Messungen jeweils 5 Minuten. Mindestens 30 Minuten vor der Messung sollten Sie nicht essen, trinken, rauchen oder sich körperlich betätigen. Wiederholen Sie die Messung im Falle zweifelhaft gemessener Werte.

«Google-Translate»:

Если вы хотите выполнить несколько измерений на человека, ждать между измерениями через каждые 5 минут. По крайней мере, 30 минут перед измерениями, вы не должны есть, пить, курить или выполнять физические нагрузки. Повторите тест в случае сомнительно результатов.

«PROMT»:

Если бы Вы хотели проводить несколько измерений в лице, ждите между отдельными измерениями соответственно 5 минут. За минимум 30 минут до измерения Вы не должны были есть, пить, курить или заниматься физически. Повторите измерение в случае сомнительно измеренной стоимости.

«Яндекс.Перевод»:

Если несколько измерений для одной личности, которую вы хотите выполнить,

ждать между отдельными измерений каждые 5 минут. По крайней мере, 30 минут перед измерением, вы не должны есть, пить, курить или физически манипулировать. Повторите измерения в случае сомнительное измеренного значения.

Все три системы при переводе данного отрывка неправильно перевели местоимение *Sie*, которое в оригинале пишется с большой буквы и, соответственно, должно писаться с большой буквы и в русском языке. Также все три системы допустили стилистическую ошибку при переводе лексемы *mindestens*. Системы «Google-Translate» и «PROMT» допустили лексическую ошибку при переводе фразы *an einer Person*, но смысл остается понятен. В переводе онлайн-переводчика «Яндекс.Перевод» мы видим нарушенную синтаксическую конструкцию первого предложения. Переводчики «Google-Translate» и «Яндекс.Перевод» не смогли распознать во втором предложении повелительное наклонение при переводе слова *warten*. Помимо этого всеми онлайн-переводчиками была неверно передана модальность глагола *sollten*. В тексте на немецком языке имелось в виду не прошедшее время, а сослагательное наклонение. Также все переводчики допустили лексическую неточность при переводе словосочетания *sich körperlich betätigen*. В последнем предложении всеми переводчиками была допущена пунктуационная ошибка: отсутствует запятая перед словосочетанием *в случае*. Помимо этого все системы дословно перевели последнее предложение. Также слово *zweifelhaft* было переведено как наречие, а не как прилагательное.

Для сравнения качества перевода возьмем отрывок инструкции по эксплуатации на английском языке (инструкция по эксплуатации для духового шкафа):

Оригинал:

Electrical products are dangerous to children. Keep children away from the product when it is operating and do not allow them to play with the product.

«Google-Translate»:

Электрические продукты опасны для детей. Держите детей подальше от продукта, когда он работает, и не позволяйте им играть с устройством.

«PROMT»:

Электрические продукты опасны для детей. Держите детей отдельно от продукта, когда он работает, и не позволяйте им играть с продуктом.

«Яндекс.Перевод»:

Электротехническая продукция опасна для детей. Держите детей подальше от продукта при его эксплуатации и не позволяйте им играть с изделием.

Все три переводчика допустили лексическую ошибку при переводе слова *products*, в данном случае имелось в виду *оборудование*. В переводе системы «Яндекс.Перевод» допущена грамматическая ошибка: несогласование подлежащего и сказуемого в числе.

Усложним задачу, взяв фрагмент текста из рекламного проспекта с техническими характеристиками (машина каблирования шинного корда CableCorder CC3) на английском языке:

Оригинал:

Fewer thread breaks

The outer yarn delivery system on CableCorder CC4 releases caught yarns thus ensuring trouble-free unwinding.

«Google-Translate»:

Меньше перерывы резьбы

Внешняя пряжи система доставки на CableCorder cc4 релизов поймали нити, таким образом, обеспечивая бесперебойную раскручивание.

«PROMT»:

Меньше разрывов нити

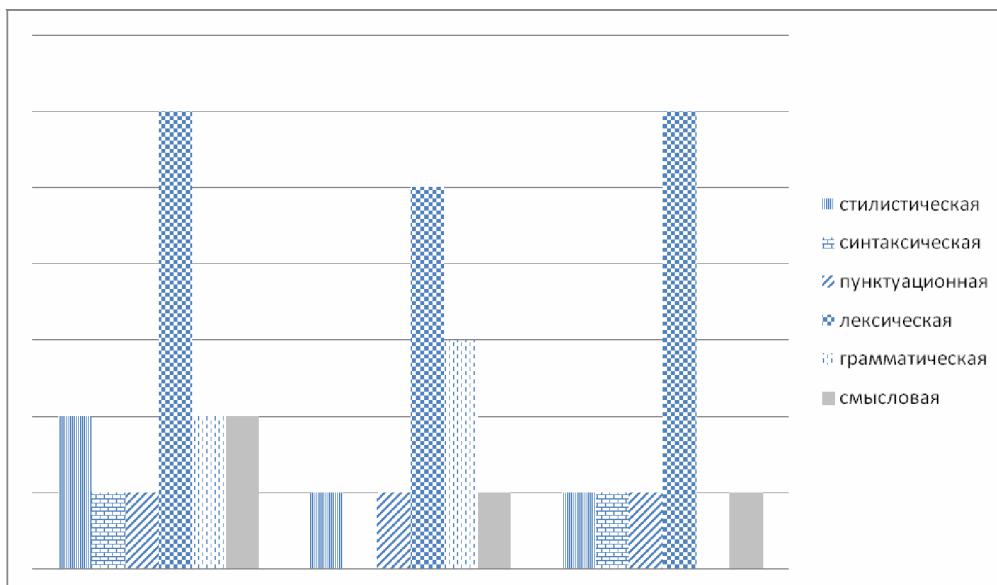
Внешняя система доставки пряжи на CableCorder CC4 выпускает пойманную пряжу, таким образом гарантирующую безаварийное раскручивание.

«Яндекс.Перевод»:

Меньшее количество обрывов нити

Внешняя система нитеподачи на CableCorder CT4-релизы поймал пряжи обеспечивая тем самым бесперебойную разматывать.

На данном примере мы видим, что с текстами, где присутствуют технические термины, системы машинного перевода справляются хуже. Переводчик «Google-Translate» перевел термин *thread* как *резьба*, что является смысловой ошибкой. *Outer yarn delivery*



Схема

system было переведено некорректно во всех трех случаях.

Глагол *releases* был переведен системами «Google-Translate» и «Яндекс.Перевод» как *релиз*. Переводчики «Google-Translate» и «Яндекс.Перевод» неправильно распознали часть речи при переводе слова *caught*, в данном случае это не глагол в прошедшем времени, а прилагательное. В переводе переводчика «Google-Translate» словосочетание *бесперебойную раскручивание* не согласуется в роде и падеже. Переводчик «PROMT» допустил пунктуационную ошибку: после конструкции *таким образом* отсутствует запятая. Также переводчик «Яндекс.Перевод» перевел существительное *unwinding* как глагол.

Результаты данного анализа представлены на схеме, где мы видим, что все системы допустили больше всего лексических ошибок. Переводчик «PROMT» не допустил ни одной синтаксической ошибки, а «Яндекс.Перевод» ни одной грамматической.

На основе проведенного исследования можно сделать вывод о том, что с помощью автоматических переводчиков можно переводить более простые тексты, где имеются устойчивые конструкции и словосочетания.

Если переводить тексты, предложения в которых осложнены придаточными конструкциями, то это существенно усложнит задачу

для систем машинного перевода и отнимет много времени у человека, так как придется редактировать большее количество текста.

В заключение стоит отметить, что машинные переводчики созданы для того, чтобы сократить время, затрачиваемое на перевод, но не сделать перевод вместо человека.

Практически во всех случаях после использования переводчика необходима редакция, выполненная человеком.

СПИСОК ЛИТЕРАТУРЫ

1. Зубов, А. В. Информационные технологии в лингвистике / А. В. Зубов, И. И. Зубова. – М. : Академия, 2004. – 208 с.
2. Манджиков, О. Яндекс запустил сервис онлайн-перевод / О. Манджиков, Д. Литвинова // ПРЕСС-РЕЛИЗЫ. – Электрон. текстовые дан. – Режим доступа: https://company.yandex.ru/press_releases/2011/0316/index.xml. – Загл. с экрана.
3. Марчук, Ю. Н. Компьютерная лингвистика : учеб. пособие / Ю. Н. Марчук. – М. : АСТ : Восток–Запад, 2007. – 317 с.
4. Новожилова, А. А. Машинные системы перевода: качество и возможности использования / А. А. Новожилова // Вестник Волгоградского государственного университета. Серия 2, Языковедение. – 2014. – № 3 (22). – С. 67–73.

COMPARATIVE ANALYSIS OF MACHINE TRANSLATION SYSTEMS

Arestova Anastasia Alexandrovna

Student, Department of Translation Theory and Practice,
Volgograd State University
n-arestova@mail.ru
Prosp. Universitetsky, 100, 400062 Volgograd, Russian Federation

Abstract. The article considers three machine systems, Google-Translate, Promt and Яндекс.Перевод. The article presents the linguistic analysis of the translations quality of texts fragments of various genres made in different language combinations by these systems. The most typical mistakes made by the systems are revealed, the possibilities and spheres of use are determined. The author reveals that the systems are not useful for translating texts which are full of compound and complex sentences.

Key words: information technologies, electronic dictionaries, computerized translation, machine translation, quality of translation.