

УДК 81'322.4

Бездорожев С.В.*(г. Москва)*НОВЫЙ ВЗГЛЯД НА РОЛЬ ПЕРЕВОДЧИКА:
ПРОБЛЕМА ИСПОЛЬЗОВАНИЯ МАШИННОГО ПЕРЕВОДА

Аннотация. Данная статья посвящена некоторым актуальным проблемам лингвистики и переводоведения, таким, как: роль переводчика как творческой личности в переводческом процессе, системы машинного перевода (МП), критерии оценки качества МП, совместная работа переводчика и системы машинного перевода, а также стратегии эффективного постредактирования перевода. Автор данной работы убежден в возможности использования МП (даже в его современном состоянии) во многих отраслях науки и бизнеса. В тексте статьи приводятся многочисленные примеры как минимум удовлетворительного качества машинного перевода в различных сферах жизни, и особенно в Интернете. Автор статьи пытается определить, какие ниши в сфере перевода можно поручить МП, а какие – переводчику, чья работа и по сей день представляет собой творческий процесс.

Ключевые слова: машинный перевод, автоматизированный перевод, постредактирование перевода, переводчик.

S. Bezdorozhev*(Moscow)*A NEW LOOK AT THE TRANSLATOR'S ROLE:
THE PROBLEM OF USING A MACHINE TRANSLATION

Abstract. This article is focused on some vital problems of linguistics and translation, such as: the role of a translator as a creative person in the translation process, machine translation systems, criteria for assessing the quality of machine translation (MT), the cooperative work of a translator and a machine translation system, as well as strategies for effective post-editing of the translation. The author states it is quite possible to use the MT (even in its current state) in many areas of science and business. The article presents numerous examples of machine translation of texts from various spheres of life, especially the Internet, which have a rather satisfactory quality. The author makes an attempt to define which niches in the field of translation can be entrusted to MT, and which – to a translator whose work remains a creative process up till now.

Key words: Machine translation, computer-aided translation, translation postediting, interpreter.

Рассуждая о качестве машинного перевода (МП), мне вспомнились несколько рассказов зарубежных писателей-фантастов, которые можно свести к одной общей теме. Они повествуют о некотором устройстве, отвечающем только правду. У данного устройства есть одна особенность – оно способно отвечать только на правильно поставленные вопросы, однако ни один человек так и не смог задать ни одного корректного, с точки зрения данного механизма, вопроса. Принцип работы этого устройства напоминает интерфейс между современным компьютером и человеком. Проблемы взаимодействия человек-машина являются актуальными в современной лингвистике. То же самое и в машинном переводе, компьютер переводит все, что видит, «как есть», если не указать в правилах иное.

Пока одни ученые полностью отрицают пользу машинного перевода и говорят о его неприменимости в реальных условиях [6, с. 55], другие активно используют возможности компьютера.

Наше общество на современном этапе развития является информационным, и тому есть много свидетельств: потоки информации во всех сферах деятельности человека стали интенсивнее, появилось множество технических средств обработки и передачи данных, таких, как компьютер и специализированное программное обеспечение. Но общение между различными народами все так же затруднено, и основная причина этого – языковой барьер. К традиционным методам преодоления языковых барьеров относят изучение иностранных языков и переводческую деятельность, однако они в настоящее время являются недостаточными. С увеличением политических, экономических, культурных связей и ростом информационных потоков между странами и материками возникает потребность для быстрого перевода огромных массивов текста. Продолжают совершенствоваться различные системы для автоматизированного перевода текста – СМП, переводческие накопители; огромные по объему и наполнению базы знаний в поисковых системах. В этом большая заслуга прикладной лингвистики, чьей задачей также является оценка качества перевода различных интеллектуальных систем. Благодаря прикладной лингвистике выполняется разработка интеллектуальных лингвистических систем – информационно-поисковых, систем машинного перевода и всевозможных других систем для автоматизированной обработки текста. Оценка качества работы данных продуктов, а также их сравнительные характеристики представляют собой первостепенно сложную задачу.

Роль переводчика в современном мире

Устанавливая языковое соответствие между исходным языком и языком перевода, переводоведение двигалось путем понимания переводческого процесса как многогранного явления, при котором соотносятся не только лексические и грамматические формы, но и языковое представление мира и обстоятельства общения вместе с обширным кругом внеязыковых факторов, которые определяются общим понятием культуры. Снова в этом процессе основную миссию выполняет творческая личность переводчика.

Размышляя о процессе перевода как таковом, мы подразумеваем не только межязыковое преобразование текста, которое заключается в его трансформации с одного языка на другой язык, но и психофизиологические процессы, связанные с интеллектуальным характером данного вида деятельности. Творческий характер деятельности переводчика заключается как раз в способности находить личные, единичные, «не предусмотренные» ни одной теорией перевода аналогии, однако если переводчик не постиг теорию перевода, то ни способность к установлению контактов, ни адекватность перевода обеспечены не будут.

Основная задача переводчика – наиболее полная передача содержания оригинала, и, большей частью, фактическое единство содержания оригинала и переведенного текста весьма значительно. Необходимо различать **достижимую эквивалентность**, которую следует понимать как максимальное единство сущности двух текстов на разных языках, допускаемое отличиями этих языков, на которых они созданы, и **переводческую эквивалентность**, которую В.Н. Комиссаров определяет как смысловую близость оригинального текста и его перевода, которая достигается переводчиком непосредственно в процессе перевода [1, с. 51].

Характерной чертой профессии переводчика является ее весьма индивидуальный характер. Непосредственно в этом заключаются сложности профессии и особенный интерес к ней, поскольку в процессе работы неизменно появляются многочисленные интеллектуальные проблемы, которые требуют моментального решения, иногда повергающие переводчика в состояние стресса. В таких ситуациях помочь переводчику никто не сможет, все зависит исключительно от его характера, знаний и индивидуальных психофизических качеств, отличающих людей друг от друга.

Профессионал, владеющий нужными сведениями о переводе, а также о его теории и практике, может верно определить сложность появившихся задач и отыскать единственно верный путь их решения.

Принимая во внимание все вышеизложенное, можно сказать, что даже поведение переводчика в момент работы имеет специфический характер. Он обязан ощущать, как аудитория воспринимает его речь. Для настоящего профессионального переводчика должен существовать лишь смысл текста, который он переводит.

В итоге работа переводчика может быть рассмотрена как коммуникативный акт (трансформация структур и объектов исходного языка в структуры и объекты другого языка на основании определенных правил), где переводчик играет роль посредника, обладающего широким кругозором, обширными знаниями в сфере лингвистики (как иностранной, так и родной), культуры, национальных обычаев страны-носительницы языка.

Совместное использование МП и труда переводчика

Следует отметить, что машинный перевод в совокупности с пред- или постредактированием позволяет достичь очень хороших результатов. Однако даже он один по качеству обычно соответствует уровню «рабочего перевода» [6, с. 51]. Многие ученые, и отечественные, и зарубежные, например, Ю.Н. Марчук и Йорик Уилкс, сходятся во мнении, что машинный перевод находится сейчас в стадии эволюции и улучшения качества. По словам последнего, большинство клиентов, приобретающих продукты для МП, довольны качеством автоматического перевода, которое сейчас составляет 65-70% правильно переведенных предложений [8, с. 225].

Процесс перевода, состоящий из нескольких этапов с использованием МП (по крайней мере, в одном из них), называется **автоматизированным переводом**. От чистого машинного перевода он отличается тем, что процесс перевода осуществляется человеком, а машина лишь помогает произвести готовый текст либо быстрее, либо качественнее.

Типичный пример – использование памяти переводов (translation memory, TM) для работы в системах автоматизированного перевода, таких, как SDL Trados или IBM Translation Manager. Концепция TM предполагает обнаружение в переводимом тексте сегментов, переводы которых уже существуют в базе данных (БД) переводов, и в итоге за счет этого происходит уменьшение объема работы переводчика. Это явление называется «выравнивание» или «сопоставление» (alignment). Фрагменты, которые остались непереуведенными после данной процедуры, передаются для ручной обработки переводчикам или переводятся с помощью системы машинного перевода. Переводчик на данном этапе

может обособить вновь переведенные сегменты, а также сохранить новые пары двуязычных параллельных текстов в базе данных. Наилучшее применение такой схемы – на однотипных текстах, где повторяемость слов и словосочетаний достаточно высока: например, технических описаниях, инструкциях для пользователей и т. п.

Многие компании, например, такие, как Adobe, PayPal, Cisco и др. используют машинный перевод для перевода пользовательских интерфейсов и другой локализационной информации. При этом технологическая цепочка в отношении машинного перевода следующая: **настройка перевода** на собственную документацию с использованием разнообразных ресурсов (словари, правила перевода и т. д.), машинный перевод, и затем – **постредактирование**, т. е. качество машинного перевода можно оценить как **«перевод, пригодный для эффективного пост-редактирования»**. Здесь речь идет о переводах больших объемов текста с соблюдением единого стиля и терминологии, поэтому для компаний имеет смысл делать предварительную настройку перевода. Оценить объем постредактирования можно исходя из данных, взятых из презентации Е.А. Лайкова, сотрудника компании Промт [2, с. 18]:

Исходный текст:	Машинный перевод:	Перевод переводчика:	Комментарий:
To this end, final analysis results are not quite obvious	С этой целью окончательные результаты анализа не совсем очевидны	Заключение, полученное в результате анализа, не дало окончательного ответа на этот вопрос	(При этом в глоссарии: analysis results = результаты анализа) Объем постредактирования предложения: 100 %
To this end, final analysis results are not quite obvious	С этой целью окончательные результаты анализа не совсем очевидны	Окончательные результаты анализа, тем не менее, не дали ответа на этот вопрос	(При этом в глоссарии: analysis results = результаты анализа) Объем постредактирования предложения: 82 %
SWIFT also brings the financial community together to work collaboratively to shape market practice and define standards	SWIFT также объединяет финансовое сообщество для работы совместно, чтобы сформировать практику рынка и определить стандарты	SWIFT также объединяет финансовое сообщество для совместной работы в целях формирования рыночной практики и определения стандартов	Объем постредактирования предложения: 65 %

Исходный текст:	Машинный перевод:	Перевод переводчика:	Комментарий:
Distributor may not take legal action against any infringer without SDA's prior written consent	Дистрибьютор может не подать в суд против никакого нарушителя без предварительного письменного согласия SDA	Дистрибьютор не имеет права подавать в суд на какого-либо нарушителя без предварительного письменного согласия SDA	Объем постредактирования предложения: 34 %
Our study focuses on emergent visual rhetorical practices that can inspire advertisers	Наше исследование сосредотачивается на визуальных риторических методах на стадии становления, которые могут вдохновить рекламодателей	Наше исследование сосредотачивается на визуальных риторических приемах, находящихся на стадии становления, которые могут вдохновить рекламодателей	Объем постредактирования предложения: 14 %

Стратегии постредактирования

- Полное редактирование перевода с проверкой всех компонентов текста по исходному фрагменту
- Использование МП исключительно как подсказки по терминологии
- Использование МП в качестве лексического или грамматического подстрочника с подбором синонимичных выражений
- Чтение МП и корректировка вызывающих сомнения компонентов с обращением к исходному тексту
- Чтение МП и исправление явных погрешностей и ошибок без обращения к исходному тексту.

Для других целей постредактирование может и не понадобиться. Например, машинный перевод (в том числе предварительно настроенный) используется для перевода новостей или комментариев в блогах. В частности, новостной сайт SMI2¹ использует технологию компании Promt для перевода новостей. На данном ресурсе ежедневно генерируется большой объем текстов, и редактировать их перевод вручную было бы нерационально. Кроме того, знаменитый сайт для путешественников TripAdvisor² переводит комментарии пользователей с помощью МП. В этом случае особая точность и стиль не нужен, пользователям достаточно понять, как в целом оценивается тот или иной отель.

¹ <http://smi2.ru/>

² <http://www.tripadvisor.com/>

Оценка качества работы систем машинного перевода

Улучшение качества современного машинного перевода – это очень трудоёмкая задача. Как верно замечает Ю.Н. Марчук, моделирование перевода «требует, по существу, разработки искусственного интеллекта в довольно объёмной области умственного усилия» [6, с. 49].

За многие годы активных разработок в области машинного перевода текста предоставлялось множество формальных решений для создания качественных автоматизированных интеллектуальных систем в данной отрасли. Для осуществления идеального машинного перевода необходимы мощные интеллектуальные устройства, воссоздающие, как минимум, три модели человеческой деятельности: человеческое мышление, речевую деятельность и мир, в котором он обитает. К сожалению, человечеством пока еще не изобретены настолько мощные машины. Однако в настоящее время существуют такие системы для компьютерного перевода текста:

1. СМП (системы машинного перевода), выполняющие полностью автоматизированный перевод;
2. системы с функцией Translation Memory, имеющие возможность предоставить средства для Machine Assisted Human Translation (МАНТ) – перевода, выполняемого человеком при помощи ПК;
3. переводчики онлайн, осуществляющие перевод текста прямо в окне Вашего браузера;
4. словари онлайн;
5. машинный перевод с контролем языка на основе мощных баз знаний;
6. услуги переводчиков в сети Internet.

Современные СМП предоставляют возможность переводить большие объемы однородного текста, имея в своем арсенале небольшой узкоспециализированный словарный запас и упрощенную грамматику. Основная проблема любой СМП как интеллектуальной системы заключается в недостаточной формализуемости материала, с которым они работают. Перевод – процесс творческий и довольно сложный, для выполнения которого требуется не только хорошая лингвистическая подготовка, но и знание области, к которой переводимый текст относится. В данном процессе большую роль играет смысл высказывания, а не точное значение каждого слова (например, игра слов, речевые обороты, сленговые высказывания, характерные для определенной местности и т. д.), чего трудно достичь при переводе с помощью компьютера. Оценка СМП выполняется по разным критериям и на всех стадиях

разработки, начиная от проектирования и заканчивая готовым коммерческим продуктом. Оценка готового продукта выполняется экспертом-переводчиком, в некоторых случаях – конечным потребителем. Имеется множество типологий ошибок машинного перевода, и окончательный вердикт для продукта зависит от оценочных шкал самого эксперта. В настоящее время единой шкалы для оценки качества перевода СМП не существует.

Выводы

Ситуация на сегодняшний день такова, что развитие многочисленных интеллектуальных систем значительно опережает развитие систем проверки эффективности их работы. Существуют разрозненные, часто субъективные критерии качества их деятельности, например, подход, основанный на ранжировании заданий по уровням сложности или набору параметров, описывающих элементы некой предметной области. Нельзя с точностью сказать, что какой-либо метод оценки СМП полон и адекватен, а сами измеряемые параметры в совокупности полны, не искажены и не противоречивы. Все это в общей сложности не позволяет использовать для выполнения плохо формализованных задач традиционные методы (методы оптимизации, методы статистики и аппроксимации). К сожалению, единого универсального метода для оценки интеллектуальных систем различных типов – СМП, информационно-поисковых, любых других методов для автоматизированной обработки текстов – не существует.

Так какую же роль будет играть среднестатистический переводчик дня завтрашнего? Возможно, он как минимум в пять раз увеличит скорость работы за счет развития СМП, памяти переводов, а также эффективных стратегий постредактирования и превратится из творческой личности практически в программиста, отлично владеющего ПК и знающего языки разметки, которые используются в системах автоматизированного перевода.

Однако останутся и мастера литературного слова, которые займут свою небольшую нишу. Художественные произведения, журналистские очерки, рекламные тексты, насыщенные эмоциональными образами, игрой слов, авторским употреблением слов, словотворчеством и др. будут переводиться хуже, чем инструкции, интерфейсы, новости и т. п.

На другом полюсе находятся, например, тексты из IT-сферы. Многие из них специально создаются в расчете на машинный перевод и на быстрое понимание (т. н. «**Контролируемый язык**», [3, с. 1]). Иногда та-

кие тексты изначально составляет компьютер из готовых «кирпичиков» (скажем, подставляет в готовый шаблон имя пользователя и дату). Такие тексты переводятся намного лучше.

Автор данной работы решил отказаться от идей утопичности машинного перевода (МП) и принимает тот факт, что МП стал технологической реальностью. В Интернете имеется достаточное количество различных практических систем машинного перевода даже для самых экзотических языковых пар (русско-арабский и арабско-русский, русско-китайский перевод и т. п.), которые являются достаточно успешными для узкого круга задач (например, перевод инструкций, метеосводок, юридических документов и т. д.). И в настоящее время происходит постоянное расширение данного круга задач.

Следовательно, можно сделать вывод о перспективности и экономической выгоде машинного перевода. Ю.Н. Марчук отмечает, что профессиональный перевод, выполненный специалистами, требует до 10 дней на авторский лист (24 страницы печатного текста) и стоит довольно дорого, в то время как перевод машинный осуществляется практически мгновенно и стоит гораздо дешевле [5, с. 58].

Если попытаться заглянуть в будущее, то можно предположить дальнейшее направление развития МП: применение искусственного интеллекта (по словам Ю.В. Рождественского) [8, с. 133] или комбинация ИИ и статистического метода (убежден Йорик Уилкс).

Литература:

1. *Комиссаров В.Н.* Теория перевода (Лингвистические аспекты). – М.: Высшая школа, 1990. – 250 с.
2. *Лайков Е.А.* Постредактирование машинного перевода: [Электронный ресурс]. – URL: <http://www.promt.ru/images/postediting.ppt> (дата обращения: 30.09.2012).
3. *Малышева И.А.* Выбирайте выражения, или Контролируемый язык: [сайт]. [2011]. URL: <http://www.promt.ru/press/news/36554/> (дата обращения: 30.09.2012).
4. *Марчук Ю.Н.* Машинный перевод и типология текстов // Журнал «Теория и практика перевода». – 2010. – № 1 (8).
5. *Марчук Ю.Н.* Модели перевода. – М.: Изд-во Академия, 2010. – 174 с.
6. *Орел М.А.* Стоп машина! // Журнал «Мосты». – 2009. – № 4 (24). – 80 с.
7. *Рождественский Ю.В.* Философия языка. Культуроведение и дидактика. – М.: ГрантЪ, 2003. – 239 с.
8. *Wilks Y.* Machine Translation: Its Scope and Limits. Springer, 2009. – 252 p.