

УДК 81.33

**М.Б. Раренко,**

кандидат филологических наук, Институт научной информации по общественным наукам (ИНИОН) Российской академии наук (РАН).  
117418, Москва, Нахимовский проспект, д. 51/21;  
e-mail: rarenco@rambler.ru

## МАШИННЫЙ ПЕРЕВОД КАК ВЫЗОВ

Историю машинного перевода с самого его зарождения и вплоть до настоящего времени можно рассматривать как вызов человеческой мысли и интеллекту. В настоящее время основными видами машинного перевода признаны четыре: во-первых, машинный перевод «по правилам» (англ. Rule-Based Machine Translation, RBMT), во-вторых, «статистический» машинный перевод (англ. Statistical Machine Translation, SMT), в-третьих, «гибридный» (англ. Hybrid Machine Translation, HMT), возникший на стыке перевода «по правилам» и «статистического» перевода», и, в-четвёртых, нейронный машинный перевод (англ. Neural Machine Translation, NMT).

История машинного перевода развивалась стремительно, знала взлёты и падения, периоды вдохновения и разочарования. Однако, по справедливому высказыванию М.Н. Марчука, тема машинного перевода будет всё время актуальной в силу ряда причин, одна из которых — желание человека познать границы невозможного.

**Ключевые слова:** перевод, машинный перевод, история машинного перевода, машинный перевод «по правилам», «статистический» машинный перевод, гибридный машинный перевод, нейронный машинный перевод.

Являясь древнейшей человеческой практикой, переводческая практика чутко реагировала на запросы общества и в каждый конкретный период времени разрабатывала наиболее востребованные формы перевода. На смену первым устным переводчикам — толмачам, задача которых заключалась в преодолении языкового барьера при общении между представителями разных племён и народностей, пришёл перевод письменных текстов. В начале XX в. человечество стало грезить о «прекрасном будущем», когда человека-переводчика полностью заменит машина, а перевод станет быстрым и недорогим, обладая при этом довольно высоким качеством. Сегодня можно с уверенностью сказать, что мечта сбылась (или почти сбылась).

В начале XX в. общественно-политическая ситуация в мире подвергается кардинальным изменениям, что, в частности, приводит к тому, что отношение к переводу (в настоящее время под переводом понимают, с одной стороны, процесс, в результате которого появляется вторичное речевое произведение (в отличие от первичного оно создаётся средствами иного (переводящего) языка, а с другой, — непосредственно сам результат переводного процесса, т.е. само вторичное речевое произведение на переводящем языке (более подробно см.: [Раренко, 2010: 115–117]), и именно так мы будем понимать перевод в нашей статье) меняется. Если до XX в. активнее всего развивается перевод литературный, т.е. перевод художественной литературы (в том числе и Священных книг, в первую очередь, Библии), и в целом основным критерием и требованием хорошего (успешного) перевода считается умение переводчика воссоздавать во вторичном тексте особенности стиля текста оригинала, то уже в самом начале XX в. переводу начинает подвергаться разные деловые документы — коммерческие, научно-технические материалы, политические, военные тексты, соответственно, и общие требования к переводу — качество, сроки и т.д. претерпевают значительные изменения. Более того, поскольку значительно возрастает объём переводимых документов, появляются идеи, с одной стороны, разработать программы интенсивной подготовки переводчиков, с другой — «механизировать» процесс перевода.

Мысли о возможности перевода без непосредственного участия человека (или с его минимальной вовлеченностью в процесс) в первые десятилетия XX столетия материализовались в середине столетия. А 4 марта 1947 г., день, которым датировано письмо американского математика и специалиста по дешифровке Уоррена Уивера<sup>1</sup> другому известному американскому математику и философу Норберту Винеру<sup>2</sup>, ставшее впоследствии достоянием общественности, поскольку в нём было сформулировано предложение, ставшее основным правилом, которое легло в основу машинного перевода, а именно — рассматривать задачу перевода как дешифровку текста, вошёл в историю человечества как день рождения машинного перевода (под которым в самом общем виде понимают как автоматизированный (полностью, т.е. без какого бы то ни было

---

<sup>1</sup> Уоррен Уивер (Warren Weaver, 1894–1978) — американский математик, администратор.

<sup>2</sup> Норберт Винер (Norbert Wiener, 1894–1964) — американский математик, один из основоположников кибернетики и теории искусственного интеллекта, автор книги «Кибернетика, или управление и связь в животном и машине» (Cybernetics: Or Control and Communication in the Animal and the Machine, 1947).

вмешательства человека, или частично, т.е. с участием человека) процесс по преобразованию текста, первоначально созданного на одном естественном (неискусственном) языке во вторичный текст на ином естественном языке, который бы по содержанию и по форме соответствовал первому, так и результат такого действия (более подробно см.: [Раренко, 2010: 94–96]).

Написанию знаменитого письма предшествовал ряд не менее важных для развития машинного перевода событий. Так, ещё в 1933 г. изобретатели из СССР и Франции, не зная о разработках друг друга, почти одновременно, с разницей всего в два с половиной месяца независимо друг от друга во Франции и в СССР запатентовали два «переводных» устройства (отметим, что ни первая, ни вторая машина не способна была выполнить какой-либо перевод), но запатентованные машины, будучи представленными как переводящие механизмы, определили вектор развития машинного перевода. Запатентованный 22 июля 1933 г. «Механический мозг» (изобретатель — французский гражданин армянского происхождения Жорж Арцруни), в сущности, представлял собой довольно большой по объёму механизированный словарь, воспроизведённый на бумажной ленте. Советский учёный П.П. Троянский запатентовал «Машину для подбора и печатания слов при переводе с одного языка на другой или на несколько других одновременно» (патент СССР № 40995 от 5 сентября 1933 г.), но и его переводящее устройство было не совсем автоматизированным, поскольку для его функционирования необходимо было наличие как минимум двух помощников со знанием языка, с которого осуществлялся перевод, и языка, на который перевод осуществлялся. Сам процесс «перевода» происходил следующим образом: 1) первый помощник вводил в машину последовательно начальную форму слова предназначенного для перевода предложения (фрагмента), указывая грамматические (морфологические) категории слова и его синтаксическую роль в предложении (изобретатель ставил перед собой задачу в ближайшее время механизировать этот этап при помощи лингвистов); 2) затем следовал этап «машинного перевода» (в результате которого слова предложения вторичного текста вновь оказывались в начальной форме); 3) второй помощник придавал «переведённому» машиной фрагменту приемлемый вид (т.е. выступал в качестве литературного редактора). Таким образом, «Машина...» П.П. Троянского фактически представляла собой двуязычный словарь.

Важно подчеркнуть тот факт, что в этот же период во всём мире шли активные работы над созданием электронно-вычислительных машин (ЭВМ). Особенно значительными были достижения в этой

области молодого немецкого инженера Конрада Цузе<sup>3</sup>, который в 1938 г. собрал первую работающую электронно-вычислительную машину в мире, названную им Z1, вслед за ней появились две её усовершенствованные версии — Z2 и Z3, и модель Z3, собранная изобретателем на основе телефонных реле, обладала всеми характеристики современного компьютера. К. Цузе, приступая к работе над следующей версией своей ЭВМ — Z4, сам же разработал первый в мире высокоуровневый язык программирования Планкалькуль (букв. «исчисление планов»). Изобретение молодого немца стало настоящим прорывом в компьютеростроении. Следующим важным шагом для развития машинного перевода стал запуск в конце 1943 г. в Великобритании компьютера *Colossus*, созданного в первую очередь для расшифровки кодов фашистской Германии.

Таким образом, к моменту окончания Второй мировой войны в мировом сообществе сложились все предпосылки создания компьютерных переводчиков — уже существовали, пусть весьма простые, даже отчасти примитивные для современного человека, механические системы перевода, были разработаны первые ЭВМ. И своеобразным «толчком» в «объединении» механических систем перевода и компьютеров выступило письмо Уоррена Уивера к Норберту Винеру. И вновь человечество оказалось перед вызовом интеллектуальной мысли.

7 января 1954 г. — ещё одна важная дата в истории машинного перевода. В этот день американские учёные продемонстрировали всему миру первую действующую программу машинного перевода. Она стала совместной разработкой, с одной стороны, представителей компании IBM и, с другой, — сотрудников Джорджтаунского университета. Демонстрация чудо-машины вошла в историю машинного перевода под названием «Джорджтаунского эксперимента». Переводящие «способности» машины были продемонстрированы на примере пары языков, принадлежащих разным языковым группам, — русского и английского. В ходе эксперимента машина осуществила переводы заранее подготовленных 49 русских предложений на английский язык. Скептики отметили, что при переводе использовался словарь, состоящий всего из 250 слов, а грамматика включала всего шесть синтаксических конструкций, сам ввод предназначенных для перевода русских предложений осуществлялся на перфокартах, а полученный результат — перевод предложений на английский язык — распечатывался на принтере, однако и они

---

<sup>3</sup> Конрад Эрнст Отто Цузе (Konrad Ernst Otto Zuse, 1910–1995) — инженер, специалист в области компьютеростроения.

должны были признать, что прорыв в разработках машинного перевода произошёл.

«Джорджтаунский эксперимент» стимулировал разработчиков машинного перевода, заставил их поверить в свои возможности, и всего через год в СССР, в Институте точной механики и вычислительной техники Академии наук, был представлен первый советский компьютерный переводчик. Словарь переводчика по сравнению со словарём, использовавшимся при демонстрации «Джорджтаунского эксперимента» был увеличен почти в 10 раз и составлял 2 300 слов. В то же самое время в стране, в Институте прикладной математики под руководством А.А. Ляпунова (его создателями стали аспирантка О.С. Кулагина и студент филологического факультета МГУ Игорь Мельчук<sup>4</sup>), шли работы над созданием ещё одного машинного переводчика.

Как своего рода подтверждение, что работы над машинным переводом были признаны перспективным научным направлением, в декабре 1956 г. в Москве появилось Объединение по машинному переводу, ставшее основным неформальным центром общения математиков и лингвистов (его создателем был В.Ю. Розенцвейг, заведующий кафедрой перевода в Московском государственном педагогическом институте иностранных языков) (более подробно см.: [Митренина, 2017: 244–300]).

История машинного перевода развивалась стремительно, знала взлёты и падения, периоды вдохновения и разочарования. Особенно сложным (прежде всего в моральном плане) оказалось десятилетие с 1956 по 1966 гг., известное как «период разочарования», поскольку на смену восторженным ожиданиям, как машина полностью заменит переводчика-человека, пришло понимание дисбаланса в соотношении «цена — качество»: при высоких затратах на разработку систем машинного перевода качество перевода, выполняемого машиной, оставляло желать лучшего. Зато в последние два десятилетия в области машинного перевода произошли значительные подвижки, а машинный перевод считается реалией времени.

Анализируя современные достижения в области машинного перевода, многие исследователи крайне положительно оценивают его перспективы: «...принимая во внимание общую ситуацию с переводами в мире, можно смело утверждать, что альтернативы

---

<sup>4</sup> Игорь Александрович Мельчук (1932) — советский и канадский лингвист, создатель лингвистической теории «Смысл ↔ Текст», профессор Монреальского университета (на пенсии с 2009 года), почётный доктор Института языкознания РАН (2020).

машинному переводу нет» [Марчук, 2007: 253], а одной из наиболее актуальных и перспективных задач машинного перевода в настоящее время видится разработка и усовершенствование программы перевода устного текста с помощью компьютерных программ, при этом основной трудностью, с которой встречаются её разработчики, признается «индивидуальная окраска звучания сегмента речи» [Алексеева, 2004: 23].

Поскольку в современном обществе наблюдается большой обмен информацией на разных языках, а новые технологии, в том числе машинного перевода, постоянно появляются и совершенствуются, машинный перевод стремительно развивается. На смену устаревающим технологиям машинного перевода приходят более современные, позволяющие достичь более высоких результатов в более сжатые сроки.

В настоящее время основными видами машинного перевода признаны четыре: во-первых, машинный перевод «по правилам» (англ. Rule-Based Machine Translation, RBMT), во-вторых, «статистический» машинный перевод (англ. Statistical Machine Translation, SMT), в-третьих, «гибридный» (англ. Hybrid Machine Translation, HMT), возникший на стыке перевода «по правилам» и «статистического» перевода, и, в-четвёртых, нейронный машинный перевод (англ. Neural Machine Translation, NMT). Развитие машинного перевода в целом следует логике развития теории перевода в целом, отражая её основные этапы. Так, машинный перевод «по правилам» разрабатывался на основе лингвистической информации об исходном и переводном языках в то время, когда в общей теории перевода лингвистический подход был главенствующим.

Каждый из четырёх подходов к машинному переводу обладает своими преимуществами и недостатками, однако последний — нейронный машинный перевод, в основе работы которого словари с фиксированными словами, а перевод осуществляется в основном на базе открытого словаря, включающем в себя имена, даты, числа и т.д.), признаётся на данный момент самым перспективным и многообещающим на том основании, что потенциально способен устранить многие недостатки предыдущих систем машинного перевода, к которым относят, прежде всего, 1) трудоёмкость и длительность разработки системы, 2) необходимость поддерживать и постоянно обновлять лингвистические базы данных, 3) так называемый «машинный акцент» при переводе, а также 4) часто непредсказуемый результат перевода.

О.В. Митренина, поясняя, что нейронный машинный перевод «моделирует на основе статистики работу нейронных сетей

человеческого мозга», уточняет, что «первый такой переводчик был запущен компанией Google в ноябре 2016 г.» и «сразу показал значительное улучшение качества переведённых текстов... и это направление, как и другие способы компьютерной обработки языка с помощью нейронных сетей, сейчас развивается наиболее активно» [Митренина, 2017: 9]. Специалисты в области машинного перевода видят ключевое преимущество NMT, прежде всего, в его способности анализировать и изучать непосредственно процесс перевода исходного текста в связанный текст на переводном языке. Отмечается, что «нейронный машинный перевод приближен к идеальному человеческому» [Сафина, Камаев, 2017: 71].

Если ещё тридцать лет назад критики машинного перевода сомневались в целесообразности дальнейших разработок в этой области, то сейчас в возможностях машины осуществлять перевод на весьма высоком уровне нет сомнений. Тем не менее нельзя однозначно утверждать преимущество машинного перевода перед переводом, выполненным человеком, и наоборот. Прежде всего, следует отметить, что перед этими двумя видами перевода по-прежнему стоят совершенно разные задачи. И основное преимущество машинного перевода состоит, безусловно, в том, что он осуществляется максимально оперативно. Также, говоря о преимуществе машинного перевода перед переводом, осуществляемым человеком, отмечается, что при выполнении заданий на перевод текста с одного языка на другой «компьютер опирается на адресные ссылки и осуществляет поиск по содержанию, что позволяет говорить о более надёжной и оперативной памяти таких программ», однако преимуществом «человеческого» перевода признаётся то, что «человек обладает гораздо большей свободой выбора, а также восприятием контекста, что на данный момент позволяет ему показывать более эффективные и корректные результаты, чем машина» [Котенко, 2020: 226]. Также замечено, что в основе наиболее современного подхода в области машинного перевода лежат сети долгой краткосрочной памяти. Они представляют такой тип рекуррентных (т.е. букв. «возвращающихся», от лат. *recurrens*) нейронных сетей, которые в памяти хранят некий контекст выполненных ранее переводов. Ключевым преимуществом рекуррентного типа сети признается сегодня её относительная невосприимчивость (игнорирование) к временным разрывам (в том числе довольно длительным) при совершении машиной операций (т.е. способность машины в течение длительного времени хранить в памяти фрагменты текста).

В заключение отметим, что, несмотря на то что в разные периоды истории машинного перевода отмечается разная интенсивность работ в области машинного перевода, проблемы автоматизации перевода, как отмечает Ю.Н. Марчук, не теряют своей актуальности (и не потеряют) в силу нескольких причин. Во-первых, машинный перевод значительно облегчает задачу обработки информации, в том числе на разных языках, которая сегодня стоит как никогда остро (среди других альтернатив преодоления языковых барьеров — создание универсального языка, изучение языков и т.д., — перевод по-прежнему является наиболее эффективным и совершенным средством). Во-вторых, происходит постоянное усовершенствование старых технологий и появление новых, поэтому у человека всегда есть интерес автоматизировать некоторые процессы, перепоручить их машине (т.е. речь идёт о так называемом интеллектуальном вызове). В-третьих, имеет место быть «научная привлекательность проблемы машинного перевода» [Марчук, 2007: 245], а машинный перевод, являясь центральной проблемой искусственного интеллекта на современном этапе (по мнению Ю.Н. Марчука [Марчук, 2007: 249, 269–300]), представляет собой область междисциплинарных исследований, где востребованы знания, умения и навыки специалистов самых разных научных направлений.

### **Список литературы**

*Алексеева И.С.* Введение в переводоведение. СПб.: Филологический факультет СПбГУ; М.: Издательский центр «Академия», 2004. 352 с.

*Котенко В.В.* Перспективы развития нейронного машинного перевода в контексте концепции открытого образования // Учёные записки университета имени П.Ф. Лесгафта, 2020. № 4 (182). С. 225–230.

*Марчук Ю.Н.* Компьютерная лингвистика. М.: АСТ: Восток — Запад, 2007. 317 с.

*Митренина О.В.* Назад, в 47-й: к 70-летию машинного перевода как научного направления // Вестн. Новосиб. гос. ун-та. Серия: Лингвистика и межкультурная коммуникация, 2017. Т. 15, № 3. С. 5–12.

*Раренко М.Б.* Машинный перевод // Основные понятия переводоведения (Отечественный опыт). Терминологический словарь-справочник / Отв. редактор канд. филол. наук Раренко М.Б. М., 2010а. С. 94–96.

*Раренко М.Б.* Перевод // Основные понятия переводоведения (Отечественный опыт). Терминологический словарь-справочник / Отв. редактор канд. филол. наук Раренко М.Б. М., 2010б. С. 115–117.

*Сафина Д.Р., Камаев Р.Р.* Использование искусственных нейронных сетей в современном машинном переводе // Информационные технологии в исследовательском пространстве разноструктурных языков: сборник трудов конференции. Казань, 2017. С. 70–72.

***Maria B. Rarenko,***

Cand. Sc. (Philology), Lecturer at the Institute of Scientific Information for Social Sciences (INION) of the Russian Academy of Sciences (RAS), Senior Research Fellow, Russia. 51/21 Nakhimovskiy Prospekt, Moscow 117418, Russia;  
e-mail: rarenco@rambler.ru

## **MACHINE TRANSLATION AS A CHALLENGE**

The history of machine translation can be seen as a challenge to human thought and intellect from its inception up to the present day. Currently, four main types of machine translation are recognized: first, Rule-Based Machine Translation (RBMT); second, Statistical Machine Translation (SMT); third, Hybrid Machine Translation (HMT), which arose at the intersection of the two aforementioned types; fourth, Neural Machine Translation (NMT).

The history of machine translation has developed rapidly, with ups and downs, periods of inspiration and disappointment. However, according to M.N. Marchuk, the topic of machine translation will always be relevant for a number of reasons, one of which is people's desire to know the boundaries of the impossible.

**Key words:** translation, machine translation, history of machine translation, machine translation “by rules,” “statistical” machine translation, hybrid machine translation, neural machine translation.

### ***References***

*Alekseeva I.S.* Vvedenie v perevodovedenie [Introduction to Translation Studies]. Saint Petersburg: Filologicheskii fakul'tet SPBGU; Moscow: Izdatel'skii tsentr “Akademiya”, 2004. 352 p. (In Russian).

*Kotenko V.V.* Perspektivy razvitiya nejronnogo mashinnogo perevoda v kontekste koncepcii otkrytogo obrazovaniya [Prospects for the Development of Neural Machine Translation in the Context of the Concept of Open Education]. Uchenye zapiski universiteta imeni P.F. Lesgafta, 2020. No. 4 (182), pp. 225–230 (In Russian).

*Ju. N.* Komp'yuternaja lingvistika [Computational Linguistics]. Moscow: AST: Vostok — Zapad, 2007. 317 p. (In Russian).

*Mitrenina O.V.* Nazad, v 47-j: k 70-letiju mashinnogo perevoda kak nauchnogo napravlenija [Back to the 47th: to the 70th Anniversary of Machine Translation as a Scientific Direction]. *Vestn. Novosib. gos. un-ta. Serija: Lingvistika i mezhekul'turnaja kommunikacija*. 2017. Vol. 15, No. 3, pp. 5–12 (In Russian).

*Rarenko M.B.* Mashinnyj perevod [Machine translation]. Osnovnye ponjatija perevodovedeniya (Otechestvennyj opyt). Terminologicheskij slovar'-spravochnik. Otv. redaktor kand. filol. nauk Rarenko M.B. Moscow, 2010a, pp. 94–96 (In Russian).

*Rarenko M.B.* Perevod [Translation]. Osnovnye ponjatija perevodovedenija (Otechestvennyj opyt). Terminologicheskij slovar'-spravochnik. Otv. redaktor kand. filol. nauk Rarenko M.B. Moscow, 2010b, pp. 115–117 (In Russian).

*Safina D.R., Kamaev R.R.* Ispol'zovanie iskusstvennyh nejronnyh setej v sovremennom mashinnom perevode [The Use of Artificial Neural Networks in Modern Machine Translation]. Informacionnye tehnologii v issledovatel'skom prostranstve raznostrukturnyh jazykov: sbornik trudov konferencii. Kazan', 2017, pp. 70–72 (In Russian).

Поступила в редакцию 12.05.2021

После доработки 16.05.2021

Принята к публикации 06.06.2021