

О. В. Митренина

*Санкт-Петербургский государственный университет
Университетская наб. 7/9, Санкт-Петербург, 199034, Россия*

mitrenina@gmail.com

НАЗАД, В 47-Й: К 70-ЛЕТИЮ МАШИННОГО ПЕРЕВОДА КАК НАУЧНОГО НАПРАВЛЕНИЯ

Предлагается целостная картина развития методов машинного перевода и обстоятельств появления ведущих переводных систем и технологий: от дешифровки, через перевод по правилам, к статистическим методам перевода и искусственным нейронным сетям, которые обеспечивают «дешифровку» текста на новом, более глубоком уровне. Делается вывод о том, что спустя 70 лет внимание исследователей вновь обращено на самую первую и, казалось бы, давно уже оставленную научную программу.

Ключевые слова: машинный перевод, история машинного перевода, машинный перевод на основе правил, статистический машинный перевод, нейронный машинный перевод, нейронные сети

Предшественники компьютерных переводчиков

Две первые машины для перевода были запатентованы независимо друг от друга в СССР и во Франции в 1933 году — еще до появления компьютеров. Они почти ничего не переводили, но с их появлением мечта о машинах-переводчиках, наконец, начала воплощаться: ведь это были реальные устройства, пусть пока не слишком успешные.

Первый патент на изобретение переводного устройства был выдан 22 июля 1933 года французу армянского происхождения Жоржу Арцруни. Название его машины звучало многообещающе: «Механический мозг». Но по сути это был лишь большой механизированный словарь на бумажной ленте.

Второй патент через два с половиной месяца получил российский ученый П. П. Троянский¹. Он описал «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 года). Этой машине в ее работе требовались два помощника: один должен был знать исходный язык, а второй — язык перевода. Первый помощник задавал машине базовую форму каждого слова, его грамматические категории и роль в предложении. Этот этап Троянский надеялся впоследствии механизировать. Только после этого машина осуществляла перевод. В новом тексте все слова оказывались в базовой форме. После

¹ Позже он добавил к своей фамилии фамилию жены и стал подписываться как Смирнов-Троянский.

этого за работу брался второй помощник. Он приводил переведенный машиной текст в литературную форму [Hutchins 2004].

В 1937 году на Парижской всемирной выставке Арцруни представил работающий образец своей машины. Троянский в это время искал и не находил помощи лингвистов, с помощью которых он надеялся решить проблему неоднозначности языка. До появления первого компьютера оставался всего один год.

В 1938 году молодой немецкий инженер Конрад Цузе сдвинул столы в гостиной родительского дома и собрал на них первую в мире электронно-вычислительную машину. Названная создателем Z1, она занимала 4 квадратных метра и весила почти 500 кг. Вслед за ней Цузе разработал две улучшенных модели Z2 и Z3. При этом Z3, собранная на основе телефонных реле, уже обладала всеми свойствами современного компьютера.

Когда во время Второй мировой войны Берлин начали бомбить, Цузе понял, что машины надо эвакуировать. Он успел вывезти на лошадях и спрятать в деревне новый незаконченный Z4 — для этого компьютера Цузе разработал первый в мире высокоуровневый язык программирования Планкалькюль, то есть «исчисление планов». Машины Z1, Z2 и Z3 вместе с документацией погибли при бомбардировке. Но Цузе после войны разработал много новых машин, а в конце 1980-х воссоздал и легендарный Z1. Цузе считал, что устройство Вселенной похоже на сеть взаимосвязанных компьютеров и верил, что правильно созданные машины смогут его отразить.

К середине войны в мире действовало уже несколько электронно-вычислительных машин. В конце 1943 года в Великобритании был запущен компьютер *Colossus*. Его единственной задачей стала расшифровка секретных кодов фашистской Германии. Эта сверхмощная для своего времени машина работала на основе 1500 электронных ламп, ее никогда нельзя было выключать. *Colossus* позволил сократить время расшифровки перехваченных сообщений с нескольких недель до нескольких часов.

Когда Вторая мировая война закончилась, мир был готов для создания компьютерных переводчиков: уже существовали и механические системы перевода, и компьютеры. Оставалось их лишь соединить. Толчком к такому соединению послужило письмо Уоррена Уивера к Норберту Винеру.

Перевод как дешифровка

Машинный перевод как научное направление родился 4 марта 1947 года. В этот день американский математик и специалист по дешифровке Уоррен Уивер написал математику и философу Норберту Винеру письмо, в котором предложил рассматривать задачу

перевода как дешифровку текста. Уивер считал, что Винер лучше любых других ученых подходит для решения проблемы машинного перевода [Weaver 1949:11].

Винер всю жизнь изучал устройство мира. Свою первую научную работу он написал в возрасте 7 лет, это было исследование по дарвинизму. В 18 лет он защитил в Гарварде диссертацию по математической логике. К началу Второй мировой войны Винер был профессором пяти университетов и заведующим кафедрой Массачусетского технологического института. Он ушел добровольцем в армию и разработал там математическую модель наведения зенитного огня. Эту модель он попытался распространить на все процессы, происходящие в окружающем мире. В 1947 году Винер заканчивал работу над главным своим трудом, книгой «Кибернетика, или управление и связь в животном и машине».

Уивер во время Второй мировой войны работал шифровальщиком. В своем письме Винеру он описал задачу машинного перевода как дешифровку²: «Когда я вижу текст на русском языке, я говорю себе, что на самом деле он написан по-английски и зашифрован при помощи странных знаков. И мне надо его просто расшифровать» [Weaver 1949:4].

В отличие от человека, компьютер легко считает частоту отдельных элементов текста и частоту их сочетаний. Он может запомнить, в каком окружении встречаются различные элементы. Эти и другие приемы дешифровки позволят найти ключ к преобразованию закодированного сообщения из цепочки символов в текст на понятном языке. Примерно так видел Уивер задачу автоматического перевода.

Норберт Винер, знавший около десяти иностранных языков, не вдохновился письмом Уивера, но идея начала жить своей независимой жизнью. Казалось, что скоро машина начнет «дешифровать» тексты с одного языка на другой.

От статистики к правилам

В 1952 году в США состоялась первая конференция, посвященная машинному переводу. Ее собрал математик и организатор науки Иегошуа Бар-Хиллел. Конференция получилась очень далекой от идей дешифровки. Вместо математических методов на ней

² Дешифровка очень ценилась после войны. Национальными героями Америки тогда стали 11 индейцев племени навахо, члены знаменитой команды шифровальщиков. Их называли секретным оружием США. Враги могли разгадывать зашифрованные тексты на английском, но зашифрованный язык навахо был для них непреодолим. Эти индейцы, как сказал награждавший их президент США, «спасли жизнь огромному числу людей и ускорили наступление мира на охваченных войной территориях» [Бейкер 2008:9]. После войны казалось, что методами дешифровки можно раскрыть законы устройства мира.

обсуждались способы описания правил и словарей для естественных языков, в первую очередь — для английского. Возможно, это было связано с тем, что практикующих дешифровщиков среди разработчиков было слишком мало (если они были вообще), но все разработчики когда-то учились в школе, где им рассказывали про структуру предложения. Подобным структурам они начали обучать машину, хотя ей гораздо легче проводить не анализ связей, а подсчет элементов и их сочетаний — именно то, что требуется при дешифровке.

Так машинный перевод повернул от статистики к анализу предложений по правилам. Позже такой подход назовут RBMT — Rule-based Machine Translation — машинный перевод на основе правил.

7 января 1954 года американцы продемонстрировали миру первую действующую программу машинного перевода. Ее разработали совместно фирма IBM и Джорджтаунский университет. В честь университета презентацию программы назвали Джорджтаунским экспериментом. На глазах у зрителей машина перевела 49 заранее отобранных предложений с русского на английский язык.

«Девушка, которая не понимает ни слова на языке Советов, набрала русские сообщения на перфокартах. Машинный мозг сделал их английский перевод и выдал его на автоматический принтер с бешеной скоростью — две с половиной строки в секунду», — сообщалось в пресс-релизе компании IBM.

Программа использовала словарь из 250 слов и грамматику, состоящую из шести синтаксических правил. Ввод предложений осуществлялся на перфокартах, а результат распечатывался на принтере, поскольку ни мониторов, ни клавиатуры в те годы не существовало.

В октябре того же 1954 года российский реферативный журнал ВИНТИ "Математика" опубликовал сообщение о Джорджтаунском эксперименте. Заметка за подписью Д. Ю. Панова называлась «Перевод с одного языка на другой при помощи машины: отчет о первом успешном испытании». Эта публикация дала старт развитию машинного перевода в СССР. Уже через год в Институте точной механики и вычислительной техники Академии наук был представлен первый советский компьютерный переводчик, он использовал словарь из 2300 слов. Другой переводчик разрабатывался в Институте прикладной математики под руководством А. А. Ляпунова. Главными его создателями стали аспирантка О. С. Кулагина и студент филологического факультета МГУ Игорь Мельчук [Мельчук 1998].

В декабре 1956 года в Москве появилось Объединение по машинному переводу. Оно стало основным неформальным центром общения математиков и лингвистов. Его

создателем и вдохновителем был В. Ю. Розенцвейг, заведующий кафедрой перевода в Московском государственном педагогическом институте иностранных языков. Многие ученые тех лет считают его главным организатором машинного перевода в Советском Союзе. Он умел объединять людей и знал, что нужно делать, чтобы наука могла развиваться в СССР поверх бюрократических и идеологических барьеров [Мельчук 1998, Успенский 2013]. Так, его Объединение по машинному переводу никогда не оформлялась документально, никто не определял его границы и статус, поэтому бороться с ним усилиями научной бюрократии было невозможно. Под редакцией Розенцвейга в 1957 начал выходить «Бюллетень Объединения по проблемам машинного перевода», получивший позже название «Машинный перевод и прикладная лингвистика».

15–21 мая 1958 года в Москве прошла организованная Розенцвейгом первая советская конференция по машинному переводу, оказавшаяся весьма представительной — в ней приняло участие 340 человек из 79 организаций. В числе других проблем на ней рассматривались алгоритмы перевода с индонезийского, арабского, норвежского, вьетнамского и других языков. Научный сотрудник из КНР Лю Юн-Цюань рассказал о проблемах создания русско-китайского переводчика. Насколько можно судить по подробному отчету, опубликованному в «Вопросах языкознания» за 1958 год [Николаева 1958], все обсуждавшийся подходы, даже использующие статистические данные, основывались на правилах: они предполагали, что машина должна переводить теми методами, что и человек, опираясь на грамматики и словари.

По итогам московской конференции Министерство высшего образования СССР издало приказ «О развитии научных исследований в области машинного перевода», и эта дисциплина, имеющая «большое народнохозяйственное и общекультурное значение», начала бурно развиваться в Советском Союзе.

Развитие вопреки кризису

В 1960 году серьезный удар по машинному переводу нанес Иегошуа Бар-Хиллел [Bar-Hillel 1960], за восемь лет до того созвавший первую конференцию по этой теме. Он заявил, что корректный машинный перевод в принципе невозможен: в определенных контекстах компьютер никогда не сможет верно распознать многозначные слова. Значит, машина может быть лишь помощником человека-переводчика, но в этом случае вкладывать средства в проект не имеет смысла.

В 1964 году правительство США сформировало знаменитую комиссию ALPAC (Automatic Language Processing Advisory Committee — Консультативная комиссия по

автоматической обработке языка). Она должна была решить, оправданно ли тратить государственные деньги на развитие машинного перевода. В 1966 году ALPAC объявила результаты своего расследования: машинный перевод бесперспективен [Pierce et al 1966]. Государственное финансирование этой темы в США и в Европе было прекращено [Pierce et al 1966, Koehn 2010:15]. На протяжении последующих 20 лет машинный перевод на Западе развивался только в частных компаниях за счет бизнес-структур и отдельных инвесторов³.

Решение ALPAC повлияло и на развитие машинного перевода в СССР: эта тема перестала быть актуальной для руководителей советской науки, началась борьба с «засильем структурализма». Следующие два десятка лет машинный перевод и компьютерная лингвистика могли развиваться в СССР только там, где руководство находило нестандартные решения для продолжения исследований⁴.

Лучший машинный переводчик Советского Союза был создан в скромном московском институте «Информэлектро» при Министерстве электротехнической промышленности СССР. Директор института С. Г. Малинин трудоустроивал у себя ученых, которых увольняли из других институтов «за вольнодумство». Система, которую они в 1972 году начали создавать, называлась ЭТАП («ЭлектроТехнический Автоматический Перевод»).

Разработками руководил Ю. Д. Апресян, уволенный в 1972 году из московского Института русского языка Академии наук за подписание писем в поддержку диссидентов. Спроектировал архитектуру системы ЭТАП и создал для нее формальный язык логик Л. Л. Цинман. В 1968 году его уволили из Московского государственного педагогического института из-за подписи в защиту насильственно помещенного в психиатрическую больницу диссидента и математика А. С. Есенина-Вольпина. В разработке лингвистического компонента участвовала Л. Н. Иорданская, уволенная в 1974 году из Института языкознания за политическую неблагонадёжность. В основе системы лежала формальная модель языка «Смысл-Текст», предложенная в 1960-х годах И. А. Мельчуком и до сего дня остающаяся одним из лучших формальных описаний естественного языка. Но Мельчук в «Информэлектро» не работал: изгнанный в 1976 году из Института языкознания, он уехал в Канаду, где живет до сих пор. Название «ЭТАП» придумала

³ А также, как утверждает легенда, за счет мормонов, которые хотели сделать перевод Библии на все языки мира.

⁴ Сохранился рассказ о том, как молодой и беспартийный заведующий кафедрой математической лингвистики в ЛГУ А. С. Герд спас свою кафедру от закрытия. Ему удалось попасть на прием в Смольный к первому секретарю Ленинградского Обкома КПСС Г. В. Романову. Этот высокопоставленный чиновник мог решать судьбу научных направлений. В своем письме А. С. Герд так рассказывал об этой встрече: «Надо отдать должное Г. В. Романову: он не стал читать моих заранее приготовленных объяснительных записок, а попросил просто и кратко рассказать о кафедре. Я быстро и со всей своей энергией за 5-7 минут показал ему народно-хозяйственное и стратегическое значение кафедры, ее цели, характер заказных договорных тем, место и роль кафедры в Ленинграде и в СССР в целом. Г. В. Романов ничего не сказал, поблагодарил, и мы расстались. Как мне сказали потом в парткоме, он был вполне удовлетворен беседой со мною».

программист группы Татьяна Коровина, которая погибла в 1985 году [Богуславский и Иомдин 2004].

ЭТАП стал единственной российской системой тех лет, дожившей до нашего времени. Но сегодняшняя его версия ЭТАП-3 — это многоцелевой лингвистический процессор. Строго говоря, его нельзя назвать переводчиком, хотя машинный перевод на его основе в научных целях реализован.

От правил к статистике

Самым успешным переводчиком XX века стал Systran. Одноименная американская компания начала разрабатывать его в 1968 году для ВВС США. Его первая языковая пара была русско-английская, она создавалась в условиях холодной войны с Советским Союзом. В 1973 году специально для российско-американского космического проекта «Союз-Аполлон» Systran создал англо-русскую пару. К началу 1990-х система стала необычайно популярной. В частности, на ее основе действовал тогда перевод в системах Yahoo! и Google.

В 2004 году руководство компании Google решило отказаться от Systran и создать собственный принципиально новый переводчик. Все действующие в то время системы основывались на правилах. Google принял решение разработать систему перевода на основе статистики без использования грамматики и словарей. Для ее создания требуется большой параллельный корпус, в котором каждое предложение на исходном языке связано с соответствующим ему переведенным предложением. Машина анализирует, какие фрагменты предложения (например, биграммы и триграммы) часто встречаются вместе в оригинале и в переводе, а затем, получив новое предложение, строит для него (только на основе статистики, без использования лингвистических знаний) наиболее вероятное предложение-перевод. Этот подход назвали SMT — Statistical Machine Translation, статистический машинный перевод.

В 2006 году, Google представил миру первый серьезный переводчик на основе статистики. Его создал за два года молодой немецкий инженер Франц Ох (Franz Och). Когда в 2004 он начал работать над этой системой, ему было 33 года. Через 10 лет, в 2014 году, Франц Ох оставил машинный перевод и ушел из Google в компанию Human Longevity которая исследует проблему человеческого долголетия. Там он занялся анализом ДНК. Эта область очень близка к статистической обработке цепочек букв (текстов), и многие методы SMT поначалу использовались в генетике при анализе ДНК.

Появление статистического переводчика означало возвращение к идеям Уоррена Уивера, который планировал рассматривать перевод как дешифровку. Такой перевод предполагает, что из всех возможных конечных цепочек выбирается наиболее вероятная. При этом машине важно только окружение, в котором встречается то или иное слово, и частота различных последовательностей слов. Параллельно в мире начал развиваться гибридный перевод (НМТ — Hybrid Machine Translation), интегрирующий статистику в правила или правила в статистику [Митренина 2016].

От статистики к глубинным смыслам языка

Следующим этапом развития машинного перевода стал переход к искусственным нейронным сетям. Этот подход моделирует на основе статистики работу нейронных сетей человеческого мозга. Первый такой переводчик был запущен компанией Google в ноябре 2016 года. Он сразу показал значительное улучшение качества переведенных текстов [Turovsky 2016], и это направление, как и другие способы компьютерной обработки языка с помощью нейронных сетей, сейчас развивается наиболее активно.

Современная система нейронного перевода состоит из двух базовых компонентов: энкодера и декодера. Энкодер — это искусственная нейронная сеть, которая обрабатывает цепочки слов на исходном языке и представляет их в виде многомерного вектора. Этот вектор можно рассматривать как информационный субстрат фрагмента текста. При этом векторные представления одинаковых или близких по смыслу предложений на разных языках оказались очень схожи между собой (см., например, [Sutskever et al 2014]).

Декодер также представляет собой нейронную сеть, которая декодирует векторные представления, т.е. синтезирует перевод предложения на основе имеющегося семантического представления.

Нейронные сети хорошо "схватывают" синтаксические и семантические связи в предложении, поэтому многие исследователи считают полученные векторы глубинно-семантическим представлением предложения. Такой взгляд соответствует интуитивным догадкам ученых, стоявших у истоков компьютерной лингвистики. Так, Уоррен Уивер описал в 1949 году «наиболее обещающий» подход к машинному переводу в виде следующего художественного образа: представим, что каждый человек живет в своей отдельной высокой башне, но все башни стоят на едином основании. Можно пытаться докричаться до других сквозь стены башни, но процесс общения при этом будет затруднен. А можно просто спуститься вниз и найти там большой подвал, общий для всех башен. Там можно наладить простое и полезное общение с теми, кто тоже спустился из

своих башен. Может быть, переводить с китайского на арабский или с русского на португальский удобнее не прямым путем, пытаясь докричаться из одной башни в другую. Лучше спуститься к некоему общему основанию человеческой коммуникации — существующему, но пока не открытому универсальному языку, а затем вновь подняться какой-нибудь удобной дорогой [Weaver 1949:11].

Приведенное описание Уоррена Уивера образно показывает работу энкодера и декодера современной системы нейронного перевода, который был изобретен почти 70 лет спустя.

Дальнейшее развитие искусственных нейронных сетей может привести к обнаружению тех «инвариантных свойств» [Weaver 1949:2], которые находят свое выражение в текстах естественного языка и имеют отношение к базовому устройству мира. Так обретут воплощение интуиции исследователей прошлого века: Конрада Цузе, Норберта Винера, Уоррена Уивера, В. Ю. Розенцвейга, Татьяны Коровиной и их единомышленников.

Список литературы

Бейкер М. Атомы языка: Грамматика в темном поле сознания. М.: URSS, 2008. 272 с.

Богуславский И. М., Иомдин Л. Л. Машинный перевод. Интервью Радио Свобода 21 января 2004 г. <https://www.svoboda.org/a/24196111.html>

Мельчук И. А. Как начиналась математическая лингвистика // Очерки истории информатики в России. Новосибирск: Научно-исследовательский центр ОИГТМ СО РАН, 1998. <http://lyapunov.vixpo.nsu.ru/?el=698&mmedia=PDF>

Митренина О. В. Машинный перевод // Прикладная и компьютерная лингвистика. М.: URSS, 2016.

Николаева Т. Н. Конференция по машинному переводу // Вопросы языкознания, № 5, 1958. С. 149-151.

Успенский В. А. Серебряный век структурной, прикладной и математической лингвистики в СССР: Как это начиналось (заметки очевидца) / Труды по НЕматематике, 2-е издание. Книга 3. Языкознание. М.: ОГИ, 2013. http://cshistory.nsu.ru/?int=VIEW&el=259&templ=BOOK_INTERFACE

Bar-Hillel Y. The present status of automatic translation of languages. *Advances in Computers*, Vol. 1, 1960, 91-163. <http://www.mt-archive.info/Bar-Hillel-1960.pdf>

Hutchins J. Two precursors of machine translation: Artsrouni and Trojanskij. *International Journal of Translation* 16(1), Jan-June 2004, 11-31. <http://www.mt-archive.info/00/IJT-2004-Hutchins.pdf>

Koehn Ph. Statistical Machine Translation. Cambridge, UK, 2010.

Turovsky B. Found in translation: More accurate, fluent sentences in Google Translate. Blog.Google, November 15, 2016 <https://www.blog.google/products/translate/found-translation-more-accurate-fluent-sentences-google-translate/>

Pierce J., et al. Languages and machines: computers in translation and linguistics. A report by the ALPAC, Division of Behavioral Sciences, National Academy of Sciences, National Research Council. Washington, D.C.: National Academy of Sciences, National Research Council, 1966.

Sutskever I., Vinyals O., Le Q. Sequence to Sequence Learning with Neural Networks, NIPS 2014. <https://arxiv.org/pdf/1409.3215.pdf>

Weaver W. Memorandum on Translation. 1949. <http://www.mt-archive.info/Weaver-1949.pdf>

Olga V. Mitrenina

Saint-Petersburg University, St. Petersburg State University, SPbSU, SPbU, 7/9

Universitetskaya nab., St. Petersburg, 199034 Russia.

mitrenina@gmail.com

BACK TO 1947: ON THE SEVENTIETH ANNIVERSARY OF MACHINE TRANSLATION AS A SCIENTIFIC TREND

This article proposes a comprehensive view of the development of machine translation methods and the circumstances surrounding the emergence of the leading translation systems and technologies: from decryption, through rule-based machine translation, to statistical methods of translation and neural machine translation that provides a “decryption” of the text on a new and more profound level. The conclusion is reached that, after seventy years, the research consensus has once again shifted towards the original and seemingly long-abandoned scientific program.

Keywords: machine translation, the history of machine translation, rule-bases machine translation, statistical machine translation, neural machine translation, neural networks

References

Baker M. Atomy jazyka: Grammatika v temnom pole soznaniya [The atoms of language: The Mind’s Hidden Rules of Grammar]. M.: URSS, 2008. 272 p. (In Russ.)

Boguslavskij I. M., Iomdin L. L. Mashinnyj perevod [Machine translation]. Interview to Radio Svoboda, January 21, 2004. (In Russ.) <https://www.svoboda.org/a/24196111.html> (In Russ.)

Mel'chuk I. A. Kak nechinalas' matematicheskaya lingvistika [How did mathematical linguistics begin] // *Studies of Informatics in Russia. Novosibirsk, 1998.* (In Russ.)
<http://lyapunov.vixpo.nsu.ru/?el=698&mmedia=PDF>

Mitrenina O. V. Mashinnyj perevod [Machine translation]. Applied and Computer Linguistics, Moscow, 2016 (In Russ.)

Nikolaeva T. N. Konferentsija po mashinnomu perevodu [Conference on Machine translation]. Topics in the Study of Language, no 5, 1958. p. 149-151. (In Russ.)

Uspenskij V. A. Serebrjanyj vek strukturnoj, prokladnoj i matematicheskoj lingvistike s SSSR: Kak eto nachinalos' (zametki ochevidtsa) [The Silver age of structural, applied and mathematical linguistics: How it has started (notes of an eyewitness)]. *Notes on NON-mathematics*, Second edition, Book 3. Linguistics. Moscow, 2013.

http://cshistory.nsu.ru/?int=VIEW&el=259&templ=BOOK_INTERFACE

Bar-Hillel Y. The present status of automatic translation of languages. *Advances in Computers*, Vol. 1, 1960, 91-163. <http://www.mt-archive.info/Bar-Hillel-1960.pdf>

Hutchins J. Two precursors of machine translation: Artsrouni and Trojanskij. *International Journal of Translation* 16(1), Jan-June 2004, 11-31. <http://www.mt-archive.info/00/IJT-2004-Hutchins.pdf>

Koehn Ph. *Statistical Machine Translation*. Cambridge, UK, 2010.

Turovsky B. Found in translation: More accurate, fluent sentences in Google Translate. *Blog.Google*, November 15, 2016 <https://www.blog.google/products/translate/found-translation-more-accurate-fluent-sentences-google-translate/>

Pierce J., et al. *Languages and machines: computers in translation and linguistics*. A report by the ALPAC, Division of Behavioral Sciences, National Academy of Sciences, National Research Council. Washington, D.C.: National Academy of Sciences, National Research Council, 1966.

Sutskever I., Vinyals O., Le Q. Sequence to Sequence Learning with Neural Networks, NIPS 2014. <https://arxiv.org/pdf/1409.3215.pdf>

Weaver W. Memorandum on Translation. 1949. <http://www.mt-archive.info/Weaver-1949.pdf>