

УДК 378.147

DOI: 10.34670/AR.2021.78.84.017

Обучение переводу в условиях распространения систем машинного перевода

Гришина Наталья Юрьевна

Кандидат педагогических наук, доцент,
Санкт-Петербургский политехнический университет Петра Великого,
195251, Российская Федерация, Санкт-Петербург,
ул. Политехническая, 29;
e-mail: grishinan@list.ru

Аннотация

Фокус статьи обращен к электронным машинным переводческим сервисам социальных сетей и интернет-ресурсов. Сравнение различных вариантов машинного перевода Facebook, Google translator, Yandex translator, Reverso помогает более осознанному пониманию о природе и структуре перевода. Кроме того, обсуждается вопрос, какие стратегии и тактики наиболее эффективны при переводе определенных типов текста, влияние социальных сетей на обучение переводу. Исследование проводилось при подготовке и проведения семинарских занятий по анализу переводческих ошибок, встречающихся в разных типах переводного дискурса. Рассматривались переводы, которые были сделаны в Facebook, Google и Yandex переводчиках, Reverso. Студентам самостоятельно анализировали содержание оригинального текста основываясь на тексте перевода. После этого предъявлялся оригинальный текст и сравнивался с вариантами, которые были предложены электронными переводчиками. Машинный перевод может быть недостаточно эквивалентен, особенно при переводе законченных высказываний, фраз. В статье рассматриваются причины потери эквивалентности на уровне текста, фразы, фразеологических единиц и отдельных слов. При выполнении таких упражнений повышается осознанность о качестве перевода и переводческом инструментарии для выполнения качественного перевода. Социальные сети также могут повышать осознанность и влиять на обучение переводу и профессиональную переводческую деятельность.

Для цитирования в научных исследованиях

Гришина Н.Ю. Обучение переводу в условиях распространения систем машинного перевода // Педагогический журнал. 2021. Т. 11. № 6А. С. 142-149. DOI: 10.34670/AR.2021.78.84.017

Ключевые слова

Обучение переводу, машинный перевод, электронные сервисы перевода, переводной дискурс, переводческий анализ, электронные сервисы.

Введение

В работе рассматриваются переводные материалы из социальных сетей, Интернет-ресурсов. Сетевые ресурсы Facebook, Google translator, Yandex translator, Reverso были выбраны для переводческого анализа в силу своей общедоступности. Сравнение различных вариантов машинного перевода помогает более осознанному пониманию сущности перевода, какие стратегии и тактики могут быть наиболее эффективны при переводе определенных типов текста, из чего складывается лингвистическая и экстралингвистическая составляющая переводного дискурса, влияние социальных сетей на обучение переводу. Исследование проводилось при подготовке и проведении семинарских занятий по анализу переводческих ошибок, встречающихся в разных типах дискурса.

Под машинным переводом в данной работе понимается перевод, который сделан автоматически специальной программой. Рассматривались переводы, которые были сделаны в электронных переводчиках Facebook, Google и Yandex переводчиках и корпусном переводчиках, таких как Reverso, Lexico Oxford, Corpus of Contemporary American English и др.

Основная часть

Коммерческий директор PROMT Никита Шаблыков в интервью TAdviser утверждает, что «по данным Global Market Insights, Inc. рынок машинного перевода к 2024 году должен составить более 1.5 млрд долларов США. Годовой прирост оценивается в 14,6%. При этом, объем рынка всех лингвистических услуг в 2019 году составит \$53.5 млрд. Сюда входят устный и письменный перевод, локализация программного обеспечения, веб-сайтов и мультимедиа, включая игры, фильмы и сериалы и многие более мелкие услуги. Машинный перевод – один из главных инструментов по оптимизации процесса перевода, локализации и работы с текстовыми большими данными».¹ Приведенные цифры говорят об увеличении услуг оказания перевода с помощью машинных программ практически в 2 раза, начиная с 2016 года.

Машинный перевод в настоящее время является доступным переводческим инструментом. Электронные сервисы постоянно улучшаются и в функциональном плане становятся эффективнее, что позволяет их использовать как в профессиональной, так и в обучающей деятельности. Студенты используют машинный переводческий сервис на занятиях для быстрого перевода научной, художественной литературы и публицистики, необходимой для осуществления образовательной деятельности. В связи с этим возникла потребность проанализировать качество машинного переводческого сервиса.

Машинный перевод делается автоматически в специальной программе. При этом он может быть сделан как с помощью различных интернет-сайтов, так и с помощью стационарных переводческих программ, предлагаемых различными коммерческими компаниями. Facebook, Google и Yandex переводчики, Reverso, Lexico Oxford и другие переводческие сервисы в настоящий момент гибридные. При построении алгоритмов машинного перевода они используют различные сочетания правил, статистики и нейронные сети. Сервисы перевода компаний Google и Yandex используют нейронный перевод. Google использует нейронный

¹ Рынок технологий машинного перевода: интервью с Никитой Шаблыковым, PROMT.

машинный перевод Google (GNMT)² вместо ранее используемых статистических методов.³ Яндекс.Переводчик также имеет гибридную модель, где изначально свой вариант перевода предлагает и статистическая модель и нейросеть, а после этого технология CatBoost, в основе которой лежит машинное обучение, выбирает лучший из полученных результатов.⁴ В основе машинного перевода заложены алгоритмы, которые формируются на основе семантического и семиотического анализа слова и корпусного анализа текстов, где данные слова проявляются в определенных контекстуальных значениях.

Модели нейронного машинного перевода используют глубинное обучение и обучение признакам. Для их работы требуется лишь малая часть памяти по сравнению с традиционными системами статистического машинного перевода (SMT). Кроме того, в отличие от традиционных систем перевода, все части модели нейронного перевода обучаются совместно от начала до конца, чтобы максимизировать эффективность перевода [Kalchbrenner, Blunsom, 2013; Sutskever, Vinyals, Quoc, 2014; Kyunghyun et al., 2014].

Двунаправленная рекуррентная нейронная сеть (RNN), также известная как кодировщик, используется нейронной сетью для кодирования исходного предложения для второй рекуррентной сети, также известной как декодировщик, которая используется для предсказания слов в конечном языке [Bahdanau et al., 2014]. В отличие от традиционного статистического машинного перевода, нейронный машинный перевод направлен на создание единой нейронной сети, которую можно совместно настраивать для максимальной производительности перевода. В основном все модели машинного нейронного перевода состоят из кодировщика, который встраивает исходное предложение в вектор фиксированной длины, и потом декодер генерирует перевод.

Для анализа студентам предлагались упражнения, которые помогали понять, как работают машинные переводчики, по какому принципу машина ищет части исходного предложения по целевым словам. Мы предположили, что более детальное изучение алгоритмов работы машинных сервисов поможет в формировании более осознанного подхода к практической переводческой деятельности. Кроме того, сравнительный анализ машинных переводчиков показывает, что гибридная модель позволяет мягко выравнять перевод, что комфортно согласуется с внутренней интуицией переводчика. Результаты, представленные ниже говорят о более осознанном понимании семиотики термина и тезаурусному моделированию [Горбунов, 2007].

Осознанное понимание механизмов работы автоматизированных переводчиков приводит к появлению устойчивого знания не только о структуре словаря родного и иностранного языка, но и грамотном структурировании предложений в зависимости от типов переводных материалов.

Методика обучения

Студентам предлагалось проанализировать разные варианты переводов, которые непосредственно использовались во время переводческих семинаров. В качестве примеров

² Lewis-Kraus, Gideon. The Great A.I. Awakening, The New York Times.

³ Google's Large Scale Deep Neural Networks Project - Greg Corrado, Senior Research Scientist.

⁴ В «Яндекс.Переводчик» встроили нейросеть с фантазией - CNews.

были взяты варианты слов и предложений, который встречались студентам на семинарских занятиях с разной тематической составляющей и которые требовали быстрого перевода. Для анализа использовались такие онлайн-переводчики как Facebook, Google и Yandex переводчики, Reverso, Lexico Oxford и другие.

Во время обсуждения рекламного дискурса студентам предлагалось оценить качество перевода встроенного переводчика Facebook, а именно объявления о вакансиях для научных работников. На начальном этапе студентам демонстрировался вариант перевода объявления о вакансии, сделанный всторонним автоматизированным переводчиком, и предлагалось догадаться об оригинальном тексте. После этого демонстрировался исходный вариант и предлагалось проанализировать варианты перевода одной той же должности на языке оригинала. Необходимо было догадаться в чем причина различия в переводе на русский язык должности "Assistant Professor". Один из вариантов перевода – «Доцент». В других вариантов "assistant professor" переводится как «Домохозяин профессор» и «Ассистент профессор». Во время обсуждения студенты пытались понять каким образом у машины сформировался алгоритм закрепления разных вариантов перевода для определенного слова или фразы. После предъявления примеров обсуждались лексикографические, семантические и семиотические теории, которые могут быть заложены в алгоритмы машинных переводчиков, а также что может способствовать созданию качественного перевода.

Во время переводческого семинара посвященному теме переводческих ошибок, с помощью Reverso сравнивались варианты перевода слова «kerran», «луч света» с иврита. Студентам предлагалось оценить работу машинного корпусного сервиса перевода Reverso с точки зрения точности представленных вариантов. Кроме того, предлагалось подумать о возможности допущения ошибки при переводе в описании Моисея, который стал изображаться с рогами, что произошло вполне возможно после неудачного перевода.

При анализе перевода книги Лоуренса Венути [Venuti, 1998] также предлагалось сравнить варианты перевода на русский язык предложения «the mere identification of a translation scandal is the act of judgment» сделанные с помощью Google и Yandex переводчиков. Google translator переводит данное предложение как «простая идентификация скандала с переводом это акт судебного решения». Yandex переводчик предлагает следующий вариант «простое выявление скандала с переводом – это акт осуждения». Участникам предлагалось сравнить два варианта перевода одного и того же предложения и проанализировать лингвистическую составляющую разделив каждое слово на означаемое и означающее. Для понимания составляющих частей морфологической и синтаксической структуры слова, оказалось полезным обратиться к теории порождающей грамматики Ноама Хомского [Chomsky, 1957]. Мы предположили, что данное упражнение поможет лучше понять структурный состав лексической единицы и алгоритмы кодировки слов в машинных переводческих сервисах.

На занятиях по академическому письму магистранты столкнулись с проблемой перевода понятия "Likert scale". При случайном ошибочном написании слова «scale» как «Likert sKale» переводчик Google предлагает перевести его как «Ликерт скале». При другом ошибочном написании «skale» переводит как «скале». При правильном написании предлагает вариант «шкала Лайкерта». Студентам предлагается сделать выводы о причинах разницы в переводном дискурсе. Кроме того, в процессе анализа различных вариантов перевода фразы обнаружилось, что изначально понятие «Likert scale» переводилось как «шкала Ликерта», так как имя ученого, в честь которого была названа шкала, транскрибировалось, а впоследствии возможно стало транслитерироваться.

Все примеры, которые использовались на занятии, были различной тематики и из различных предметных областей, сделанные с помощью машинных электронных переводчиков.

Обсуждение результатов

Во время анализа приведенных выше примеров, участники пришли к выводу, что машинный перевод не всегда эквивалентен, особенно при переводе законченных высказываний, фраз, устойчивых выражений. Смысл на уровне целого текста, фразового уровня, фразеологических единиц теряется часто, что может быть связано со многими факторами. Если слово неверно набрано, то варианты машинного перевода могут варьироваться. Анализ теории позволяет предположить, что машинные переводчики используют алгоритмы, которые придерживаются правил произношения, чтения и других грамматических правил, существующих в языке оригинала. Интересно заметить, что при прослушивании слов с ошибочной орфографией электронный переводчик читает их в соответствии с правилами и вполне возможно, именно это является причиной появления варианта произношения слова «Likert» как «Лайкерт», вместо предположительно оригинального варианта произношения транскрипцией фамилии автора.

Детальный анализ лингвистической и экстралингвистической составляющей позволяет лучше понять из чего состоит переводной дискурс. При выполнении таких упражнений повышается осознанность о практической составляющей переводческой деятельности, его качестве, инструментах для осуществления качественного перевода.

Опросы, которые проводились после занятия подтверждают наше предположение, что осознанность в понимании механизмов функционирования языков в целом в частности позволит сформировать устойчивое знание о лексическом и грамматическом составе языков оригинала и родного языка.

Осознанность о практической составляющей переводческой деятельности, его качестве, инструментах для осуществления качественного перевода во время выполнения упражнений, озвученных выше повышается [Гришина, 2016]. Детальный анализ лингвистической и экстралингвистической составляющей позволяет понять из чего состоит переводной дискурс.

Использование социальных сетей может быть очень полезной практикой во время занятий по переводу, так как появляется значительное количество примеров для обсуждения теоретических и практических аспектов машинных переводческих сервисов. Такая возможность возникает благодаря общедоступности электронных переводчиков и возможности их использования на занятии.

При детальном изучении предложенных вариантов переводов можно говорить о том, что качество перевода во многом зависит от качества предъявляемого текста.

Осознанное понимание механизмов работы автоматизированных переводчиков приводит к появлению устойчивого знания не только о структуре словаря, родного и иностранного языка, но и грамотном структурировании предложений в зависимости от типов переводных материалов.

При работе с примерами переводов приходилось обращаться к различным современным теориям, что позволило их детальному изучению и усвоению. Большинство студентов, которые участвовали в обсуждении машинных вариантов переводов утверждают, что теоретические аспекты [Алексеева, 2020, 7-12] помогли лучше понять механизмы работы электронных переводческих сервисов.

Осознанность не являлась основным фокусом наблюдения для данного исследования, но некоторые техники, такие как целеполагание на здесь и сейчас, сосредоточение на инструкции

при выполнении задания и др., использовались на занятии. Можно предположить, что они также могли повлиять на качество усвоения материала.

Если рассматривать осознанность в широком смысле, как особый целенаправленный способ обращения внимания [Azevedo, Shane, 2019], то можно говорить, что эффективность усвоения материала увеличивается и повышается мотивация к обучению. В большей степени осознанность в данной работе упоминается как физиологический когнитивный процесс, происходящий в текущий момент времени, который направляет внимание участвующего на текущий момент и задания. Кроме того, важно заострить внимание на безоценочность при выполнении любого задания. Для концентрации внимания на текущем моменте, необходимо создавать ситуацию, когда вы подмечаете, что происходит в данный момент. Например, задаете вопросы, которые помогают откровенно говорить, что получилось, а что не удалось в процессе симуляционных действий [Гришина, 2017, 58], что вызвало негативные реакции и почему. Это помогает проанализировать как индивидуальные сложности, с которыми столкнулся студент во время симуляции, так и сложности, с которыми столкнулись другие участники. Данное упражнение позволяет максимально сымитировать будущую профессиональную деятельность. Вовремя обсуждения большинство участников признают и осознают все те сложности, с которыми они могут столкнуться во время профессиональной деятельности, и понимают необходимость работы с ними.

Заключение

Машинные сервисы перевода могут помогать и улучшать качество переводного дискурса. Электронные сервисы легкодоступны и могут быть эффективны при обучении теоретическим аспектам лингвистической деятельности и при их практическом применении.

Детальное обсуждение алгоритмов работы машинных переводчиков помогает лучше понять структуру как иностранного языка, так и родного. Кроме того, повышается понимание о содержательных компонентах переводных текстов.

Осознанный подход к выполнению упражнений помогает эффективному восприятию материала и закреплению навыков. Кроме того, увеличивается мотивация к выполнению определенного переводческого действия.

Социальные сети могут повышать осознанность, а также влиять на обучение переводу.

В дальнейшем при обсуждении алгоритмов работы машинных переводческих сервисов хотелось бы включить в обучение переводчиков практикум по алгоритмам написания электронных словарей, что значительно повысит понимание в механизмах функционирования языка и переводческой практике.

Библиография

1. Алексеева И.С. Проблемы обучения переводу в современном мире // Язык и культура в эпоху глобализации. СПб., 2020. С. 7-12.
2. Горбунов Е.Ю. Тезаурусное моделирование английских грамматических терминов морфологии и морфосинтаксиса: автореф. дис. ... канд. филол. наук. СПб., 2007. 24 с.
3. Гришина Н.Ю. Влияние интернет-технологий на процесс принятия решения как новый политический тренд (на примере компании «Кембридж Аналитика») // Научно-технические ведомости Санкт-Петербургского государственного политехнического университета. Гуманитарные и общественные науки. 2019. Т. 10. № 1. С. 69-80.
4. Гришина Н.Ю. Формирование неосознанной компетентности у студентов-международников при обучении дисциплине «Межкультурная коммуникация» // Педагогический журнал. 2016. Т. 6. № 5А. С. 295-304.

5. Гришина Н.Ю. Формирование профессиональных компетенций студентов посредством методики драматизации в преподавании социальных дисциплин // Вопросы методики преподавания в вузе. 2017. Т. 6. № 21. С. 57-66.
6. Кобозева И.М. Когнитивно-семантический подход к описанию средств связи предложений (на примере коннекторов со значением непосредственного следования) // Грамматические процессы и системы в синхронии и диахронии. М., 2016. С. 17.
7. Azevedo A., Shane M.J. A new training program in developing cultural intelligence can also improve innovative work behavior and resilience: A longitudinal pilot study of graduate students and professional employees // International Journal of Management Education. 2019. 17(3). P. 1025-1036.
8. Bahdanau D., Kyunghyun C., Bengio Y. Neural Machine Translation by Jointly Learning to Align and Translate. URL: <https://iclr.cc/archive/www/lib/exe/fetch.php%3Fmedia=iclr2015:bahdanau-iclr2015.pdf>
9. Chomsky N. Syntactic Structures. The Hague: Mouton, 1957. 117 p.
10. Kalchbrenner N., Blunsom P. Recurrent Continuous Translation Models. 2013. URL: <https://aclanthology.org/www.mt-archive.info/10/EMNLP-2013-Kalchbrenner.pdf>
11. Kyunghyun C. et al. On the Properties of Neural Machine Translation: Encoder-Decoder Approaches // Proceedings of SSST-8, Eighth Workshop on Syntax, Semantics and Structure in Statistical Translation. 2014. P. 103-111.
12. Sutskever I. et al. Sequence to sequence learning with neural networks. NIPS, 2014. URL: <https://proceedings.neurips.cc/paper/2014/file/a14ac55a4f27472c5d894ec1c3c743d2-Paper.pdf>
13. Venuti L. The Scandals of Translation: Towards an Ethics of Difference. Routledge, 1998. 210 p.

Automatic translation services help to teach translation theory and practice

Natal'ya Yu. Grishina

PhD in Pedagogy, Associate Professor,
Peter the Great Saint Petersburg Polytechnic University,
195251, 29, Politekhnikeskaya str., Saint Petersburg, Russian Federation;
e-mail: grishinan@list.ru

Abstract

The article examines the texts translated with social networks and other automatic services. Comparison of different computer translation practices of Facebook, Google translator, Yandex translator, Reverso gives clearance and awareness about translating activities, which strategies and techniques may be used to make the translation process more effective. It analysis the better ways to translate different types of texts, the structure of linguistic and extralinguistic components of translating discourse, the effect of social networks in teaching translation and interpretation techniques. The research was done prior and after the classes on translating mistakes analysis in different types of discourses. Different social networks as Facebook, Instagram etc., so as such Internet engines as Google, Yandex, Reverso are used for translating along with different online dictionaries and corpus services. Students were targeted to guess about the original content basing their conclusions on the translated texts. The original text that was demonstrated after gives an opportunity to compare automatic translation and variants suggested by students. In conclusion, machinery translation does not always give equivalent translation of original texts. Interference into the original text can bring to its misunderstandings and misinterpretations. The problem is more obvious with translation of phraseological units, completed utterances and sentences. Translation analysis raises awareness about translation quality and tools that can be used to make the translation more effective. A detailed analysis of the linguistic and extra linguistic components allows us to understand about the structure of translated discourse. Social media may give a significant assistance in teaching translation and interpretation.

For citation

Grishina N.Yu. (2021) Obuchenie perevodu v usloviyakh rasprostraneniya sistem mashinnogo perevoda [Automatic translation services help to teach translation theory and practice]. *Pedagogicheskii zhurnal* [Pedagogical Journal], 11 (6A), pp. 142-149. DOI: 10.34670/AR.2021.78.84.017

Keywords

Translation teaching, automatic translation, electronic translation services, translating discourse, translation analysis.

References

1. Alekseeva I.S. (2020) Problemy obucheniya perevodu v sovremennom mire [Problems of teaching translation in the modern world]. In: *Yazyk i kul'tura v epokhu globalizatsii* [Language and culture in the era of globalization]. St. Petersburg.
2. Azevedo A., Shane M.J. (2019) A new training program in developing cultural intelligence can also improve innovative work behavior and resilience: A longitudinal pilot study of graduate students and professional employees. *International Journal of Management Education*, 17(3), pp. 1025-1036.
3. Bahdanau D., Kyunghyun C., Bengio Y. *Neural Machine Translation by Jointly Learning to Align and Translate*. Available at: <https://iclr.cc/archive/www/lib/exe/fetch.php%3Fmedia=iclr2015:bahdanau-iclr2015.pdf> [Accessed 12/12/2021]
4. Chomsky N. (1957) *Syntactic Structures*. The Hague: Mouton.
5. Gorbunov E.Yu. (2007) *Tezaurusnoe modelirovanie angliiskikh grammaticheskikh terminov morfologii i morfosintaksisa. Doct. Dis.* [Thesaurus modeling of English grammatical terms of morphology and morphosyntax. Doct. Dis.]. St. Petersburg.
6. Grishina N.Yu. (2016) Formirovanie neosoznannoi kompetentnosti u studentov-mezhdunarodnikov pri obuchenii distsipline "Mezhkul'turnaya kommunikatsiya" [Teaching the subject "Intercultural Communication": formation of unconscious competence of students of International Affairs program]. *Pedagogicheskii zhurnal* [Pedagogical Journal], 6 (5A), pp. 295-304.
7. Grishina N.Yu. (2017) Formirovanie professional'nykh kompetentsii studentov posredstvom metodiki dramatzatsii v prepodavanii sotsial'nykh distsiplin [Formation of professional competencies of students through the method of dramatization in teaching social disciplines]. *Voprosy metodiki prepodavaniya v vuze* [Questions of teaching methods at the university], 6, 21, pp. 57-66.
8. Grishina N.Yu. (2019) Vliyanie internet-tekhnologii na protsess prinyatiya resheniya kak novyi politicheskii trend (na primere kompanii «Kembridzh Analitika») [The influence of Internet technologies on the decision-making process as a new political trend (on the example of Cambridge Analytica company)]. *Nauchno-tekhnicheskie vedomosti Sankt-Peterburgskogo gosudarstvennogo politekhnicheskogo universiteta. Gumanitarnye i obshchestvennye nauki* [Scientific and technical bulletin of the St. Petersburg State Polytechnic University. Humanities and social science], 10, 1, pp. 69-80.
9. Kalchbrenner N., Blunsom P. (2013) *Recurrent Continuous Translation Models*. Available at: <https://aclanthology.org/www.mt-archive.info/10/EMNLP-2013-Kalchbrenner.pdf> [Accessed 12/12/2021]
10. Kobozeva I.M. (2016) Kognitivno-semanticheskii podkhod k opisaniyu sredstv svyazi predlozhenii (na primere konnektorov so znacheniem neposredstvennogo sledovaniya) [Cognitive-semantic approach to the description of means of communication of sentences (on the example of connectors with the meaning of direct following)]. In: *Grammaticheskie protsessy i sistemy v sinkhronii i diakhronii* [Grammatical processes and systems in synchrony and diachrony]. Moscow.
11. Kyunghyun C. et al. (2014) On the Properties of Neural Machine Translation: Encoder-Decoder Approaches. In: *Proceedings of SSST-8, Eighth Workshop on Syntax, Semantics and Structure in Statistical Translation*.
12. Sutskever I. et al. (2014) *Sequence to sequence learning with neural networks*. NIPS. Available at: <https://proceedings.neurips.cc/paper/2014/file/a14ac55a4f27472c5d894ec1c3c743d2-Paper.pdf> [Accessed 12/12/2021]
13. Venuti L. (1998) *The Scandals of Translation: Towards an Ethics of Difference*. Routledge.