

# ОСОБЕННОСТИ ОНЛАЙН-ПЕРЕВОДА С ФРАНЦУЗСКОГО ЯЗЫКА НА РУССКИЙ (НА ПРИМЕРЕ РАБОТЫ СЕРВИСА ЯНДЕКС.ПЕРЕВОДЧИК)

Л. Э. Щекина

Воронежский государственный университет

**Аннотация:** Целью данного исследования является анализ процесса онлайн-перевода текстов различных функциональных стилей на французском языке для выявления и классификации основных ошибок веб-сервиса Яндекс.Переводчик. В этой связи были изучены вопросы определения, истории и типологии машинного перевода, оценки качества перевода, была представлена система возможной классификации переводческих ошибок. На основе изученных примеров были сделаны выводы о возможностях и недостатках веб-сервиса Яндекс.Переводчик.

**Ключевые слова:** машинный перевод; оценка качества перевода; переводческие ошибки.

В динамично развивающемся XXI веке, во время автоматизации и ускорения процессов человеческой деятельности вопрос о необходимости использования средств машинного перевода встает особенно остро. В связи с этим нам кажется актуальной задача развития данного направления. Благодаря распространению всемирной паутины, увеличению количества интернет-магазинов и в целом времени, проводимому в сети, люди все чаще обращаются к онлайн-помощникам для получения перевода, объяснения, толкования того, что они нашли в интернете. Очевидно, что пока не представляется возможным заказать у веб-сервиса нотариально заверенный перевод высокого качества. Но перевести отзыв о продукте для его заказа, инструкцию для купленной техники, скрипт сериала, изучаемого по вечерам — это стало возможным? В нашем исследовании мы задались целью проанализировать процесс онлайн-перевода текстов различных функциональных стилей на французском языке для выявления и классификации основных ошибок веб-сервиса Яндекс.Переводчик.

Нам кажется логичным начать исследование с определения термина «*машинный перевод*». Изначально машинный перевод (далее: МП) рассматривался как процесс чисто автоматической обработки текстов без участия человека. Позже в ходе работы с данным инструментом стало понятно, что для получения качественного перевода без вмешательства человека не обойтись. Большая российская энциклопедия определяет МП как «*перевод текстов с одного естественного языка на другой при помощи компьютера*» [1]. В. В. Воронович говорит о *процессе перевода, реализуемом компьютером полностью или почти полностью* [2, с. 3]. Он также отмечает, что это определение является узким пониманием значения термина. В широком смысле МП — это «*область научных исследований, находящаяся на стыке лингвистики, математики, кибернетики*», целью которой является «*построение систем, реализующих машинный перевод в узком смысле*» [там же]. В нашем исследовании мы будем опираться на узкое значение термина.

Направление МП получило развитие благодаря появлению ЭВМ в 40-е годы XX века. Значимым событием в истории перевода стал Джорджтаунский эксперимент 1954 года — первый в мире публичный эксперимент по машинному переводу, прошедший на базе Джорджтаунского университета в США [6, с. 7-8]. В ходе эксперимента компьютеру предлагалось перевести несколько предложений с русского на английский язык. По тем меркам эксперимент прошел успешно, в большей степени благодаря тому, что его условия были довольно примитивны (был произведен перевод 49 тщательно подобранных простых предложений на базе грамматики из шести правил с использованием словаря на 150 единиц). Ученые были вдохновлены этими результатами, но специально созданный комитет для оценки эффективности исследований в машинном переводе (ALPAC) в 1966 году представил отчет о нерентабельности и низкой эффективности проделанных работ, что привело к уменьшению финансирования данной области. Позже технический прогресс и появление ПК привели к тому, что в восьмидесятых годах стали появляться коммерческие системы МП (*Ariane, EUROTRA, Distributed Language Translation system*), началась активная работа над созданием автоматизированных инструментов перевода - *словарей, грамматических редакторов и памяти перевода* [10, с. 5]. Популяризация интернета в 90-е годы и высокий уровень возможностей персональных компьютеров обеспечили реальный спрос на МП. Сейчас основные исследования ведутся в области МП, основанного на правилах, статистического МП, а также гибридного МП, сочетающего в себе лучшие характеристики вышеперечисленных систем и пытающегося минимизировать их недостатки [6, с. 8-10]. В настоящее время наиболее известными и распространёнными в мире системами МП являются системы компании Systran и компании Google. В России на текущий момент конкурентоспособными аналогами остаются компании «ПРОМТ» и «Яндекс.Переводчик».

Классифицировать МП можно по следующим параметрам. По степени участия человека в переводе выделяют:

- ФАМТ (Fully-automated Machine Translation) - полностью автоматический машинный перевод;
- НАМТ (Human-assisted Machine Translation) – машинный перевод с участием человека (человек при этом выступает в качестве редактора ИТ и/или ПТ);
- МАНТ (Machine-assisted Human Translation) - перевод, выполненный человеком при помощи компьютера (с использованием электронных словарей, автоматических редакторов текста, системы памяти переводов и т. д.) [7, с. 32].

В зависимости от назначения перевода применяют *программы, выполняющие «грубый перевод»* [1], так называемый «джистинг» [11], для общего ознакомления с содержанием текста или же программы, ориентированные на более качественный перевод (часто с вмешательством человека).

Также различают системы МП *в зависимости от характера текста* (письменный или устный) на вводе.

По особенностям архитектуры системы выделяют: *перевод, основанный на правилах (Rule-based machine translation, RBMT), статистический (корпусный) перевод и гибридный перевод* [8, с. 4-6; 11].

Требования, предъявляемые к переводчику и веб-сервису, осуществляющему перевод, несколько отличаются. Если при оценке текста, переведенного человеком, анализируются, например, стилистические особенности ПТ, культурная и историческая соотнесенность текстов, достижение прагматической цели, то от компьютера подобных характеристик мы требовать не можем, поскольку система не владеет метаинформацией и не может ее осознать.

Для анализа работы систем МП используют методы *экспертной оценки и автоматические методы*. В первом случае исследователи проводят субъективную оценку результатов работы системы по двум параметрам: *адекватность* (правильная передача смысла ИТ) и *гладкость* текста (соответствие ПТ нормам ПЯ, грамматическая корректность). В качестве автоматической оценки перевода чаще всего используются показатель *Word Error Rate (WER)*, метрики *BLEU* и *NIST*. Данные инструменты сравнивают перевод, осуществленный системой МП, с эталонным переводом и вычисляют, насколько они похожи по показателям *точности и полноты перевода* [10].

Единой классификации переводческих ошибок на данный момент не существует. Опираясь на классификации ошибок, предлагаемые разными авторами [3-5], мы создали систему основных ошибок в переводе, которую нам будет удобно использовать при анализе функционирования веб-сервиса Яндекс.Переводчик. Она представлена на рисунке:



Также необходимо отметить, что представленные ошибки взаимосвязаны между собой и ошибка одного типа может быть причиной появления ошибки другого.

В качестве базы для исследования нами были выбраны тексты различных функциональных стилей, являющиеся наиболее популярными для перевода с помощью веб-сервисов: онлайн-шопинг (отзывы, описание товара/сайта), новостные ленты, техническая документация (инструкция по эксплуатации), юридическая документация, учебный перевод [12].

В ходе анализа были получены следующие результаты. Чаще всего веб-сервис совершает лексико-семантические ошибки (52 примера). Из них:

- большая часть приходится на ошибки, связанные с неверным подбором эквивалента лексической единицы

Исходный текст	Текст перевода
<i>Tout au long de l'année, faites-vous plaisir grâce à BrandAlley qui vous invite à <b>craquer</b> sur des vêtements de marque...</i>	<i>В течение всего года побалуйте себя благодаря BrandAlley, который приглашает вас <b>взломать</b> дизайнерскую одежду...</i>

Глагол «*craquer*» может переводиться как «взламывать что-либо», «распаковывать», «открывать для себя».

Исходный текст	Текст перевода
<i>Pour mettre à jour le système d'exploitation de votre téléphone, il est nécessaire de <b>passer par</b> le logiciel intégré au téléphone.</i>	<i>Чтобы обновить операционную систему телефона, необходимо <b>пройти через</b> встроенное в телефон программное обеспечение.</i>

В данном случае л.е. переведена буквально. Имеется в виду, что обновление следует осуществить через настройки программного обеспечения.

Исходный текст	Текст перевода
<i>En étudiant les différences de <b>contamination des matériaux</b> arrivés à la surface, on peut <b>redessiner</b> la <b>plomberie</b> sous-jacente.</i>	<i>Изучая различия в <b>загрязнении материалов</b>, поступающих на поверхность, можно <b>перерисовать базовую сантехнику</b>.</i>

Термины «контаминация пород» и «подземное строение/водопровод» относятся к геологии и были переданы неверно.

- Ошибки передачи имен собственных

Исходный текст	Текст перевода

<i>Dans l'Oise, important foyer de contamination, les écoles sont fermées</i>	<i>В УАЗе, крупном очаге заражения, школы закрыты</i>
-------------------------------------------------------------------------------	-------------------------------------------------------

Название французского департамента Уаза было заменено на название автомобильного завода УАЗ в России.

### ● Ошибки в переводе игры слов

Исходный текст	Текст перевода
<i>Si vous rencontrez les problèmes suivants:</i> 1. <i>Maux de tête</i> 2. <i>Mauvaise vue</i> 3. <i>Mauvaise Humeur</i> 4. <i>Insomnie</i> 5. <i>Nervosité</i> 6. <i>Vertige.</i> <i>Ce n'est pas COVID 19</i> <i>ce sont les symptômes de manque d'argent</i> <i>!!! Ca s'appelle <b>POCHVID 20</b></i>	<i>При возникновении следующих проблем:</i> 1. <i>Головные боли</i> 2. <i>Плохое зрение</i> 3. <i>плохое настроение</i> 4. <i>Бессонница</i> 5. <i>Нервозность</i> 6. <i>Головокружение.</i> <i>Это не COVID 19, это симптомы</i> <i>отсутствия денег !!! Это называется</i> <i><b>ПОЧВИД 20</b></i>

Веб-сервис не смог распознать игру слов, которую можно было бы перевести как «карманный вирус» или «вирус пустых карманов».

### ● Отсутствие перевода, или пропуск лексической единицы

Исходный текст	Текст перевода
<i>La République <b>en marche</b> doit chercher un</i> <i>nouveau candidat pour le remplacer : le</i> <i>nom d'<b>Agnès Buzyn circule</b></i>	<i>Республика должна искать нового</i> <i>кандидата на его замену-имя <b>Агнес</b></i> <i><b>Бужин</b></i>

Две л.е. были пропущены, а имя бывшего министра здравоохранения Франции (Аньес Бузен) передано неверно.

### ● Перевод, смысл которого трудно распознать

Исходный текст	Текст перевода
<i>En effet, s'il n'y a qu'une seule chambre</i> <i>magmatique, celle-ci est donc de belle</i> <i>taille.</i>	<i>Действительно, если есть только одна</i> <i>магматическая камера, поэтому она</i> <i>красива по размеру.</i>

Здесь можно было бы предложить следующий перевод: «если речь идет только об одной магматической камере, то она должна быть действительно крупного размера».

Далее по частотности следуют морфологические ошибки (16 примеров). В том числе:

### ● Ошибки в роде л.е.

Исходный текст	Текст перевода
<i>... et j'en suis très contente</i>	<i>...и я очень доволен им</i>

### ● Ошибки в употреблении времен глаголов

Исходный текст	Текст перевода
<i>Les magmas présents <b>seraient</b> très riches en silice...</i>	<i>Присутствующие магмы <b>будут</b> очень богаты кремнеземом...</i>

В данном случае время глагола выражает предположение, высказанное одним из ученых.

### ● Ошибки в употреблении падежей

Исходный текст	Текст перевода
<i>L'hiver, c'est aussi le moment idéal pour changer votre intérieur</i>	<i><b>Зимой</b> это также идеальное время, чтобы изменить свой интерьер</i>

Также были выявлены синтаксические ошибки в построении предложений (10 примеров):

Исходный текст	Текст перевода
<i>Une odeur un peu citronné et la fraîcheur qu'il dégage en fond un très bon déodorant.</i>	<i>Немного лимонный запах и свежесть он источает в основном очень хороший дезодорант.</i>

Здесь ошибки дополнительно обусловлены отсутствием редактирования ИТ.

Исходный текст	Текст перевода
<i>C'est la première fois <b>que</b> j'utilise ce produit...</i>	<i><b>Это первый раз,</b> когда я использую этот продукт...</i>

Неверный перевод ограничительной конструкции.

Пунктуационных ошибок в проанализированных примерах было выявлено меньше всего (2 примера):

Исходный текст	Текст перевода
<i>Tout le reste, les notifications, le calendrier, le réveil, le minuteur, les appels, la météo, etc. peut être exécuté sans encombre et révèle un design</i>	<i>Все остальное, уведомления, календарь, будильник, таймер, звонки, погода и т. д. Можно запустить без помех и выявить визуально привлекательный</i>

*visuellement attrayant.*

*дизайн.*

В данном случае было отмечено неправильное оформление обобщающего слова.

В целом результаты исследования могут быть представлены в форме следующей таблицы:

<b>Типы ошибок</b>	<b>Количество ошибок</b>
<b>Лексико-семантические:</b>	<b>52</b>
л.е. переведена пeverпо:	42
ошибка в выборе эквивалента	29
имя собственное	10
сочетаемость слов	2
игра слов	1
л.е. не переведена (пропущена)	4
смысл фразы в целом не понятен	6
<b>Морфологические:</b>	<b>16</b>
род	6
предлог	4
время	3
падеж	2
число	1
<b>Синтаксические:</b>	<b>12</b>
построение предложения	10
согласование членов предложения	2
<b>Пунктуационные</b>	<b>2</b>

Взяв за основу результаты исследования, мы пришли к следующим выводам.

Веб-сервис Яндекс.Переводчик чаще всего совершает ошибки по следующим причинам:

- недостаточный анализ контекста и метаинформации,
- отсутствие предредактирования ИТ,
- используемые правила перевода и примеры в корпусе в полной мере не удовлетворяют потребностям перевода.

Если для устранения первой причины необходимо создание и внедрение искусственного интеллекта, то для решения двух последующих достаточно увеличивать имеющийся корпус системы и использовать онлайн-переводчик хотя бы с минимальным знанием исходного языка. Приведенные примеры ошибок в переводах французских текстов могут быть использованы в дальнейшем при создании и усовершенствовании систем машинного перевода.

Веб-сервис Яндекс.Переводчик может применяться пользователями для перевода формализованных (технических, юридических) текстов с французского языка на русский – здесь было найдено наименьшее количество ошибок, при этом основные были связаны с передачей терминов. Также программа может помочь с пониманием общего смысла прочитанного в новостной ленте или в интернет-магазине. Однако осуществить перевод высокого класса в настоящий момент под силу только переводчику-человеку.

### Литература

1. Богуславский И. М. Автоматический перевод / И. М. Богуславский. // Большая российская энциклопедия. Том 1. - Москва, 2005. - 156 с.
2. Воронович В. В. Машинный перевод: конспект лекций по специальности Компьютерная лингвистика, Компьютерное обучение языкам / В. В. Воронович. – Минск : Белорусский Государственный университет, 2013. - 39 с.
3. Гарбовский Н. К. Теория перевода : учебник / Н. К. Гарбовский. – Москва: Издательство Московского университета, 2004. – 544 с.
4. Комиссаров В.Н. Теория перевода (лингвистические аспекты): учебник для институтов и фак. иностр. яз. / В. Н. Комиссаров. - Москва.: Высш. шк., 1990.
5. Кочеткова Н. С. Особенности машинного перевода / Н. С. Кочеткова, Е. В. Ревина // Филологические науки. Вопросы теории и практики. – Тамбов : Грамота, 2017. - №6. – С. 106-109.
6. Кулагина О. С. О современном состоянии машинного перевода / О. С. Кулагина // Математические вопросы кибернетики. - Москва, 1991. – Вып.3. - 50 с.
7. Шилихина К. М. Основы прикладной лингвистики : учебное пособие по специальности 021800 (031301) - Теоретическая и прикладная лингвистика / К. М. Шилихина ; Воронеж. гос. ун-т .— Воронеж : ЛОП ВГУ, 2006. - 51 с.
8. Chéragni Mohamed Amine. Theoretical overview of machine translation. Proceedings ICWIT, 2012. - 160 с.
9. Koehn P. Statistical machine translation. – Cambridge University Press, 2009. - 433 с.
10. Kübler N. La traduction automatique: traduction machine // Université Paris-Diderot (Paris 7). – 2007. - 14 с.
11. Молчанов А. Статистические и гибридные методы перевода в технологиях компании PROMT / А. Молчанов – URL: [https://www.promt.ru/company/technology/white\\_papers/molchanov-2013.pdf](https://www.promt.ru/company/technology/white_papers/molchanov-2013.pdf) (дата обращения : 18.04.2020).
12. Трудности перевода : 25 лет спустя. – URL: <https://ponedelnikmag.com/post/trudnosti-perevoda-25-let-spustya> (дата обращения : 20.04.2020).