Например, во время обрядов, освященной бобровой струей, водой и сѣсом очищают постройки от нечистых духов, от порчи и т. п. Помимо охранительных функций эти предметы обладают целительными свойствами. Их используют в родильной обрядности.

В качестве оберегов манси использовали и кости животных, например щучью челюсть, медвежьи лапы которые также широко употреблялись и в медицинских целях. До сих пор в мансийских домах можно встретить хранящиеся над дверями щучьи челюсти. Считается, они ограждают дом от злых людей, от напастей. Применение костей животных может иметь отпугивающее значение, и связано с их поражающими свойствами (когти, зубы).

Анализ представлений о социальных ролях показывает, что для разных групп общества в поздней традиции в нормативных и аксиологических целях предназначен особый круг мифологических сюжетов. Конечно, он очерчивается достаточно условно, потому что большинство поверий общеизвестны.

В обрядах и повседневной жизни манси используют много вариантов сообщения с иным миром и защиты себя от лесных пугающих существ. Некоторые из них имеют ограниченный характер распространения.

Среди рассмотренных примеров доминируют мотивы защиты от персонажей иного мира, им уступают способы вызывания духов, сообщения им информации и задабривания, что вполне закономерно. Разнообразие форм и средств, их сохранение указывают на то, что в мировоззрении народа манси, в его повседневном быту представления о потусторонних силах имели большое значение и до сих пор актуальны.

В целом, основной набор вещей и действий с предметами совпадает у манси с таковыми и у других народов. Кроме этого, важна семантика предмета в народном быту, то есть и его очистительные свойства.

1. Голева Т. Г. Мифологические персонажи в системе мировоззрения коми-пермяков. СПб.: Изд-во «Мама-тов», 2010. 80 с.
2. Толстой Н. И. Магическая сила голоса в охранительных ритуалах // Очерки славянского язычества. М. Индрик, 2003б. С. 428–430.
3. Панченко Л. Н. Персонажная структура мансийского фольклора: лесная женщина Миснэ // Вестник угроведения. 2019. Т. 9. № 4. С. 692–701.
4. Панченко Л. Н. Лесной великан Мѣнkv - персонаж мансийского фольклора // Вестник угроведения. 2018. Т 8. № 4. С. 658–671.
5. Ромбандеева Е.И. История народа манси (вогулов) и его духовная культура (по данным фольклора и обрядов). Сургут: АИИК «Северный дом» и Северо-Сибирское региональное книжное издательство, 1993. 208 с.
6. Зеленин Д. К. Избранные труды. Статьи по духовной культуре 1917–1934. М.: Индрик, 1999, 352 с.

Полякова Н.Е.

О критериях оценивания качества машинного перевода

*Пермский государственный национальный исследовательский университет
(Россия, Пермь)*

doi: 10.18411/lj-05-2020-173

idsp: ljjournal-05-2020-173

Аннотация

В статье рассматривается понятие машинного перевода и его отличие от автоматизированного, дается краткий обзор существующих на сегодняшний день онлайн-сервисов и систем машинного перевода, а так же методов оценивания их работы. Кроме того, предлагается анализ преимуществ и недостатков каждого метода, в

результате которого определяются системы, выполняющие наиболее качественный машинный перевод.

Ключевые слова: перевод, машинный перевод, автоматизированный перевод, оценка перевода

Abstract

The article considers the concept of machine translation and draws a distinction between machine and automated translation. It also provides a short overview of currently used online-services and systems of machine translation as well as the analyses of the evaluation methods for the translations made by these services. In addition, as a result of this examination the systems providing the most qualitative translation are identified.

Key words: translation, machine translation, automated translation, evaluation of translation

На сегодняшний день неуклонно увеличивается количество людей, использующих электронные ресурсы для осуществления переводов текстов с одного языка на другой. В связи с этим увеличивается и количество программ, выполняющих данную функцию (сервис перевода текстов Google Translate, online-переводчик компании ПРОМТ, online-переводчик InterTran и др.) быстро, доступно, а главное – бесплатно. Однако, одна из актуальных проблем, связанных с машинным переводом – это проблема оценки его качества.

Прежде всего, следует различать два понятия: машинный перевод и автоматизированный перевод. Автоматизированным считается перевод текстов с одного языка на другой, осуществляемый человеком с использованием специализированного программного обеспечения, за счет чего обеспечивается единообразие терминов и сокращается время перевода. Машинный же перевод это процесс перевода текстов с одного языка на другой машиной посредством специальной компьютерной программы. В отличие от автоматизированного перевода, здесь перевод производит программа, а человеку лишь остается откорректировать полученный результат. В данной статье будет рассматриваться именно этот тип перевода, поскольку обращение к нему становится все более популярным, так же увеличивается количество программ, выполняющих данную функцию. В настоящее время существуют и развиваются высококачественные системы машинного перевода, на помощь которым приходят новые информационные технологии и корпусная лингвистика [Дроздова: 2015]. Сегодня по типу осуществления перевода выделяют:

- системы прямого перевода: производится пословный перевод, отсутствуют модули полного семантического и синтаксического анализа;
- трансферные системы: используются переводные соответствия: эквивалентные, вариативные, трансферные, то есть преобразующие текст для правильной передачи;
- системы семантического перевода: принцип действия основан на применении семантических баз данных.

При этом машинный перевод без редактирования, выполненного человеком, не представляется возможным, так как программное обеспечение не дает возможности для подбора адекватного эквивалента некоторым семантическим конструкциям. Наряду с этим существует ряд проблем для осуществления задач машинного перевода. Например, одной из главных сложностей машинного перевода является принципиальная неоднозначность языка. Вместе с развитием программ для осуществления машинного перевода и количества пользователей данных программ возникает проблема оценки понятности и качества такого типа перевода.

Чтобы оценить качество машинного перевода в настоящее время используются следующие методы:

- Экспертная оценка
- Автоматические методы
- Оценка с точки зрения конкретной задачи

Если машинный перевод оценивается с точки зрения конкретной задачи, то в этом случае учитываются такие факторы, как время, потраченное на редактирование или точность передачи информации

Экспертная оценка осуществляется переводчиками-экспертами, которые оценивают перевод на адекватность и гладкость текста. Адекватность в данном случае означает правильную передачу смысла исходного текста, а гладкость текста демонстрирует соответствие перевода нормам принимающего языка, правильность с точки зрения грамматики [Koehn, 2010]. Традиционно шкала оценивания включает в себя значения от 1 до 5, которые присваиваются тексту субъективно. Основная проблема данного типа оценивания – субъективность экспертов, кроме того, это достаточно дорогостоящий и трудоемкий процесс.

Автоматические методы оценки качества машинного перевода представляют собой технологии оценки качества без эталона. На сегодняшний день существует несколько таких методов, основанных на различных техниках - сравнении строк, информационном поиске и др. Для автоматической оценки работы машинного перевода используются показатель Word Error Rate (WER), метрики BLEU и NIST. Эти инструменты позволяют успешно сравнивать работу разных систем МП и оценивать улучшения в работе конкретной системы. Используются также метрики точность, полнота и F-мера [Koehn, 2010].

Каждый из упомянутых выше инструментов оценивания имеет свой принцип работы.

Word Error Rate основывается на измерении расстояния между машинным и образцовым переводом. По сути WER измеряет минимальное количество изменений, которые необходимо сделать, чтобы из результата работы машинного переводчика получить эталонный перевод [Koehn, 2010]. Инструмент BLEU (Bilingual Evaluation Understudy) позволяет учитывать не только точность перевода отдельных слов, но и цепочек слов и является самым современным методом оценки машинного перевода на сегодняшний день. Алгоритм BLEU сравнивает последовательные фразы автоматического перевода с последовательными фразами, которые он находит в эталонном переводе, и взвешенно подсчитывает количество совпадений. Для получения достоверного результата для данного инструмента важен правильный порядок слов. По мнению специалистов, эта метрика лучше всего реализуется на уровне целого текста, поскольку на уровне предложения метрика обнуляется и может работать некорректно. Третьим популярным инструментом является разработанная на основе BLEU метрика NIST, которая позволяет оценить правильный выбор лексики. Для использования метрик BLEU и NIST требуется корпус предложений на исходном языке и различные эталонные переводы этих предложений, выполненные человеком.

На сегодняшний день существует ряд исследований, в которых анализируются и сравниваются результаты работы различных машинных переводчиков. Наиболее популярными среди пользователей являются Google Translate, онлайн-переводчик PROMT, Dictier, Free Language Translator, Яндекс переводчик, TranslateIt!, DeepL. Все эти системы обладают неоспоримым набором преимуществ: они бесплатные, предоставляют моментальный перевод, обеспечивают перевод в рамках нескольких языков. Большинство таких исследователей выделяют Google как обеспечивающий наиболее адекватный перевод.

Альтернативой Google на текущий момент является машинный переводчик DeepL (deep learning), который владеет 42 языковыми комбинациями между немецким, английским, французским, испанским, итальянским, польским и голландским языками.

Данная технология основывается на архитектуре нейронных сетей, что, по утверждению создателей, впервые позволило переводить большие фрагменты текста практически без ошибок. Одной из особенностей системы является то, что если при редактировании эксперт выявил неточность, система сама предлагает несколько альтернативных вариантов. Учитывая все возрастающую популярность данного ресурса, нами было проведено исследование качества предоставляемого DeepL перевода на материале нескольких трех различных стилей текста: научного, публицистического, официально-делового. Перевод осуществлялся с английского языка на русский и наоборот. Полученные результаты свидетельствуют о действительно высоком качестве данного переводчика в отношении всех трех стилей, особенно при переводе с русского языка на английский. Следует отметить адекватное построение предложений и выбор лексических единиц, можно даже говорить о сохранении жанрово-стилистических особенностей. Однако при более детальном разборе и сопоставлении текстов на языке оригинала и языка перевода были выявлены следующие недочеты. Иногда машинный переводчик «пренебрегал» тематическим распределением, что отвечает за логическую связность текста. Данное явление было более очевидно при оценке качества перевода с английского языка на русский. При подробном анализе адекватности перевода отрывков из публицистических текстов, оказалось, что программа не всегда выбирает адекватную контекстуально лексическую единицу. И то, и другое было ожидаемо, учитывая, что подобные промахи встречаются и в переводах, выполненных человеком.

Безусловно, перевод, выполняемый машинными сервисами, с каждым годом становится все более качественным, однако, на сегодняшний день все равно требуется проверка и коррекция живым переводчиком.

1. Дроздова К. А. Машинный перевод: история, классификация, методы [Текст] // Филологические науки в России и за рубежом: материалы III Междунар. науч. конф. (г. Санкт-Петербург, июль 2015 г.). — СПб.: Свое издательство, 2015. — С. 139-141. — URL <https://moluch.ru/conf/phil/archive/138/8497/> (дата обращения: 07.03.2020).
2. Переходько И.В., Мячин Д.А. Оценка качества компьютерного перевода // Вестник Оренбургского государственного университета. 2017. № 2 (202). С. 92-99.
3. Скворцова О. В., Тихонова Е. В. Проблемы и преимущества автоматизированного и машинного переводов // Молодой ученый. — 2016. — №9. — С. 1287-1289. — URL <https://moluch.ru/archive/113/29477/> (дата обращения: 09.03.2020).
4. Улиткин И.А. Автоматическая оценка качества машинного перевода // Перевод и когнитология в XXI веке: материалы V международной научной теоретической конференции (27-30 апреля 2012). М.: Изд-во МГОУ, 2012. С. 61–77.
5. Улиткин И.А. Автоматическая оценка качества машинного перевода научно-технического текста // Вестник МГОУ. Серия: Лингвистика. 2016. № 4. С. 174-182.
6. Koehn P. Statistical Machine Translation. Cambridge: Cambridge University Press, 2010.
7. White J.S. How to evaluate machine translation // Computers and Translation: A Translator's Guide / Somers, H. Amsterdam/Philadelphia: John Benjamins, 2003. pp. 211–244.

Саньярова Р.Р., Салихова З.М.

Лексемы Кара/черный и ак/белый в башкирском языке (лексико-семантический анализ)

*Башкирский государственный университет
(Россия, Уфа)*

doi: 10.18411/lj-05-2020-174

idsp: ljjournal-05-2020-174

Аннотация

В данной статье рассматриваются цветообозначения кара/черный и ак/белый в башкирском языке, также дается лексико-семантический анализ этих лексем.