

ошибочно перевёл слово *immutable*, что в результате исказило изначальный смысл, заложенный в фразу. В свою очередь, переводчик от Bing, несмотря на непопулярность среди русскоязычной аудитории, показал себя лучше Яндекс, исказив лишь слово *sent*. Перевод же от популярного в России и Европе переводчика PROMT оказался хуже всех, потерялся смысл всей исходной фразы.

Как мы видим, методы онлайн-переводчиков довольно схожи, но результат перевода различен. Исходя из вышеуказанного, можно прийти к выводу, что для текстов технического содержания более адаптированы переводчики от Google, и в меньшей степени от Bing. В свою очередь мы не рекомендуем использовать переводчики российских компаний, такие как PROMT и Yandex, так как переводы этих сервисов были ошибочными.

В свою очередь отметим, что оценка перевода в данных онлайн сервисах затруднена тем фактом, что сервисы не распространяют информацию о методах и технологиях, которые они используют для операций перевода текстов. Однако считаем, что объём аудитории сервиса в оценке является немаловажным фактором, и, учитывая высокую посещаемость русскоязычных пользователей на сайт «Google Translate», можно сказать, что этот сервис располагает достаточно большим количеством правил для перевода с русского языка на английский и наоборот.

Таким образом, системы машинного перевода еще далеки от совершенства. Профессиональные переводчики утверждают, что машина контекст не понимает и не поймёт. В свою очередь, новые возможности программирования и вычислительной техники будут вносить свой вклад в совершенствование и дальнейшее развитие теории и практики машинного перевода в жизни людей.

Список использованных источников:

1. Электронные ресурсы для изучения иностранных языков [Электронный ресурс]. – Режим доступа: https://knowledge.allbest.ru/pedagogics/2c0b65635a2bc79a5d43b88521306d37_0.html. – Дата доступа 21.02.2019.
2. Машинный перевод [Электронный ресурс]. – Режим доступа: <http://study-english.info/article065.php>. – Дата доступа 21.02.2019.
3. Проблемы машинного перевода [Электронный ресурс]. – Режим доступа: http://moiperevod.ru/info/articles/article_detail.php?id=112. – Дата доступа 21.02.2019.

СОВРЕМЕННАЯ ЛЕКСИКОГРАФИЯ: ВОЗМОЖНОСТИ ЭЛЕКТРОННЫХ СЛОВАРЕЙ

Литвиненко И.А.

*Белорусский государственный университет информатики и радиоэлектроники
г. Минск, Республика Беларусь*

Петрова Н.Е. – к.филол.н., доцент

Статья посвящена современной лексикографии, в частности, компьютерной лексикографии. Рассмотрено понятие электронного словаря, определены его виды. Обращается внимание на сложность технологии составления электронных словарей, приводятся примеры наиболее распространённых в сети Интернет словарей. Основная цель работы – показать преимущества электронных словарей в сравнении с бумажными лексикографическими источниками.

За последние десятилетия сильно ускорилось развитие науки и технологического прогресса, расширилось научно-техническое влияние на все сферы жизни. Параллельно развивалась и терминология, как бытовая, так и профессиональная. Именно с этим связан современный интерес к языку профессиональной коммуникации. Каждый год приток новых терминов увеличивается, и их необходимо систематизировать и описать. Этим и занимается лексикография – раздел языкознания, занимающийся вопросами составления словарей и их изучения.

Компьютерная лексикография – прикладная научная дисциплина в языкознании, которая изучает методы использования компьютерной техники для составления словарей [1]. Актуальность темы нашей работы связана с тем, что компьютеризация и информатизация проникли во все сферы жизни современного общества, и доступ к интернету для обычного человека весьма прост. Как следствие, переход от бумажных на электронные словари обусловлен повышением потребности в подобных ресурсах, а также вызван существенными преимуществами электронных словарей.

В общем, лексикография бывает теоретическая и практическая. Практическая лексикография выполняет общественно важные функции, обеспечивая обучение языку, описание и нормализацию языка, межъязыковое общение, научное изучение языка. Лексикография стремится найти наиболее оптимальные и допустимые для восприятия способы словарного представления всей совокупности знаний о языке [2].

Теоретическая лексикография охватывает комплекс проблем, связанных с разработкой макроструктуры (отбор лексики, объём и характер словника, принципы расположения материала) и микроструктуры словаря (структура словарной статьи, типы словарных определений, соотношение разных видов информации о слове, типы языковых иллюстраций и т. п.), с созданием типологии словарей, с историей лексикографии [3].

Именно теоретическая лексикография связана с таким объектом, как электронный или автоматический словарь. Он представляет собой любой упорядоченный, относительно конечный массив лингвистической информации, представленный в виде списка, таблицы или перечня, удобного для размещения в памяти ЭВМ и снабженного программами автоматической обработки и пополнения. Различаются автоматические словари конечного пользователя (АСКП) и автоматические словари для программ обработки текста (АСПОТ) [4].

Автоматические словари, предназначенные для конечного пользователя, чаще всего являются компьютерными версиями хорошо известных обычных словарей, таких, как словарь Ушакова онлайн (<http://ushakovdictionary.ru>) или толковый словарь Даля онлайн (<http://slovardalja.net/>) [5]. Они представляют собой одни из самых популярных толковых словарей русского языка. В словари включена лексика художественной литературы, распространенная научная, производственная и общественно-политическая лексика, а также включены слова, встречающиеся во многих говорах. Словарь Даля – один из крупнейших словарей русского языка. Содержит около 200 000 слов и 30 000 пословиц, поговорок, загадок и присловий, служащих для пояснения смысла приводимых слов. Пагинация оригинального издания в электронных вариантах словарей не воспроизведена.

Автоматические словари для программ обработки текста (информационно-поисковые тезаурусы, частотные словари, рубрикаторы, классификаторы, словари морфологического анализа, словари для машинного перевода) включают подробную информацию о морфологических, синтаксических и семантических особенностях функционирования слова. Количество зон словарных статей в таких словарях варьируется от 1 до 100. Каждая зона словарной статьи содержит особый тип словарной информации: лемма, грамматическая информация или стилистические пометы, зоны значения и толкования. Примером частотного словаря может служить «Частотный словарь современного русского языка», авторов О. Н. Ляшевской и С. А. Шарова или его электронный вариант – «Новый частотный словарь русской лексики» (<http://dict.ruslang.ru/freq.php>). Словарь основан на коллекции текстов Национального корпуса русского языка, представляющей современный русский язык периода 1950–2007 годов. Объем выборки, на которой строится большинство разделов словаря, составляет 92 млн. словоупотреблений [6].

В машинном переводе разработчики используют собственные словари. Как правило, в систему загружается два словаря для перевода с одного языка на другой. Такие системы могут производить переводы на множество языков мира, но качество такого перевода зависит от работы нейросети, которая наполняет внутрисистемные словари. Системы машинного перевода могут быть как бесплатными (Google Переводчик), так и платными (PROMT, ABBYY Lingvo). Также стоит заметить, что некоторые из них могут поставлять свои словари отдельно от полноценной системы.

Рассмотрим основные преимущества электронных словарей над бумажными. Явным преимуществом электронных словарей является мультимедийность и гипертекстуальность [7]. Мультимедийность, или совместное использование в электронных словарях таких мультимедийных средств, как звук, изображение, видео, даёт возможность добавления изображений в словарную статью с целью лучшего иллюстрирования того или иного понятия. Отметим, что иллюстрации должны не только и не столько установить жёсткую ассоциацию у пользователя словаря, а скорее помочь ему создать определённый образ в своём сознании, который поможет запомнить то или иное слово. Помимо изображений, в электронных словарях используются также аудио- и видеоматериалы, которые расширяют возможности презентации лексикографического материала [8].

Гипертекстуальность заключается в использовании электронного гипертекста – модели организации электронного текста, характеризующейся специфической структурированностью и разветвленной системой программно поддерживаемых внутритекстовых и межтекстовых переходов, предполагающей возможность читательского интерактивного воздействия на последовательность воспроизведения композиционных единиц [9]. Это позволяет читателю незамедлительно перейти к другому тексту в процессе чтения, а после вернуться обратно на то же место. Главным плюсом данной технологии является возможность получить доступ к непересекающимся информационным ресурсам, между которыми могут быть установлены однонаправленные и перекрестные ссылки [10].

Одним из преимуществ электронного словаря назовём и то, что электронный словарь не привязан к бумажным ресурсам и не имеет ограничений по объёму, вследствие чего становится возможным не просто увеличить количество включенных в словарь единиц, но и расширить описание грамматических характеристик, состава слова, фразеологии и т.д. Такой словарь можно

называть объёмным [11]. В тоже время электронный формат позволяет добавлять дополнительные иллюстрации, схемы, чертежи, а также видеoinформацию. Внедрение технологии синтеза речи в компьютерную лексикографию позволяет прослушать произношение искомым слов в аудио формате. Это особо важно для тех, кто изучает иностранные языки.

Ещё одним важным преимуществом электронных словарей является динамичность. К авторам бумажных словарей постоянно выдвигаются новые требования, необходимые правки накапливаются, и к моменту выпуска нового издания словаря, они уже являются неявными и устаревшими. Электронный словарь позволяет в любой момент времени добавлять новую или актуализировать старую информацию.

Сравним несколько популярных электронных лексикографических ресурсов, а именно: «Викисловарь» (<https://ru.wiktionary.org>), источник «Грамота.ру» (<http://gramota.ru/slovari/>), «СловОнлайн» (<http://slovonline.ru/>), «Gufo.me» (<https://gufo.me/>). Каждый из данных ресурсов предоставляет доступ к большому количеству словарей. Преимуществом «Грамота.ру» мы считаем то, что можно искать слова с ошибками (выполнять проверку правописания), где предлагается список похожих слов, среди которых будет и искомое слово, но уже с исправленной ошибкой. Если «Gufo.me» и «Викисловарь» на подобный запрос выдаст набор статей, где фигурируют части введенного неправильного слова, но исправлять и предлагать ничего не будет, то «СловОнлайн» просто сообщит, что ничего не найдено. Во всех выбранных ресурсах, за исключением «СловОнлайн», поиск слова осуществляется сразу по всем словарям на сайте. «СловОнлайн» сначала требует выбрать, в каком конкретном словаре нужно искать слово, и только после этого предложит список словарей, где оно ещё встречается. «Викисловарь» и «Gufo.me» дают возможность посмотреть перевод слова на различные языки. Единственным из перечисленных словарей, где используется аудио- и видеoinформация, а также большое количество иллюстрации является «Викисловарь».

Подытоживая сравнение, можно прийти к выводу, что «Викисловарь» на текущий момент является самым объёмным и интерактивным словарем в русскоязычном сегменте электронных лексикографических ресурсов. Визуальное оформление самого ресурса не является таким удобным как на «Gufo.me», отсутствует возможность проверки правописания и поиска слов с ошибками, как на «Грамота.ру», но, благодаря большой мультимедийной составляющей, наличию морфологических, семантических и синтаксических свойств, этимологии, антонимов, синонимов и устойчивых сочетаний, «Викисловарь» можно назвать наиболее удобным и универсальным электронным лексикографическим ресурсом.

Таким образом, рассмотренная в статье функциональность современных электронных словарей уже сейчас является ключевым фактором при выборе источника информации. Но, с возрастанием требованиями к данной функциональности, возрастает и сложность её реализации. Развитие сферы информационных технологий рано или поздно заставит человечество перейти на использование исключительно электронных ресурсов. Поэтому именно сейчас крайне важно усовершенствовать и вывести на новый уровень технологии, связанные с составлением электронных словарей.

Список использованных источников:

1. Компьютерная лексикография [Электронный ресурс]. – Режим доступа: https://ru.wikipedia.org/wiki/Компьютерная_лексикография
2. Лексикография [Электронный ресурс]. – Режим доступа: <https://ru.wikipedia.org/wiki/Лексикография>
3. Лексикография [Электронный ресурс]. – Режим доступа: <https://ru.wikipedia.org/wiki/Лексикография>
4. Баранов, А. Н. Введение в прикладную лингвистику: учеб. пособие / А. Н. Баранов. – 5-е изд. – Москва : Моск. Гос. ун-т им. М. В. Ломоносова, 2017. – 368 с.
5. Чепик, Е.Ю. Компьютерная лексикография как одно из направлений современной прикладной лингвистики / Е. Ю. Чепик // Ученые записки Таврического национального университета им В.И. Вернадского. – 2006. – Т. 19 – № 3-4. – С. 274-279.
6. Ляшевская, О.Н. Частотный словарь современного русского языка / О.Н. Ляшевская, С.А. Шаров. – Москва, 2009. – С. 5-12.
7. Кантышева, Н. Г. Гипермедиаальные структурные элементы представления специальных знаний в компьютерной лексикографии / Н. Г. Кантышева // Вестник Челябинского государственного университета. Сер. Филология. Искусствоведение. – 2012. – Вып. 71. – № 32 (286). – С. 47-50.
8. Маевский, С.С. Особенности создания учебного электронного толкового словаря русского языка : дис. ... канд. филол. наук : 10.02.02 / С.С. Маевский. – Минск, 2015 – 66 с.
9. Дёдова, О. В. Лингвосомиотический анализ электронного гипертекста (на материале русскоязычного Интернета) : автореф. дис. ... д-ра филол. наук : 10.02.01, 10.02.19 / О. В. Дёдова. – Москва, 2006. – 253 с.
10. Стройникова, С.А. Электронный лексикографический гипертекст как объект компьютерной лингвистики и современной лексикографии / С. А. Стройникова // Высшее гуманитарное образование XXI века: проблемы и перспективы. В 2-х т. Т.2. Филология и другие науки. – Самара, 2009. – С. 364–366.
11. Кантышева, Н. Г. Гипермедиаальные структурные элементы представления специальных знаний в компьютерной лексикографии / Н. Г. Кантышева // Вестник Челябинского государственного университета. Сер. Филология. Искусствоведение. – 2012. – Вып. 71. – № 32 (286). – С. 47-50.