

Русско-белорусский и белорусско-русский машинный перевод глаголов ФСК восприятия

*Красковский Н. И. асп. БГУ
науч. рук. проф. Ровдо И. С., д-р филол. наук*

Эксперты в области создания систем автоматической обработки естественно-языковых текстов, в частности разработчики онлайн-сервисов статистического машинного перевода (Statistical-based machine translation), отмечают, что для подобных программ перевод одного слова двумя и более, а также выявление и перевод неэквивалентных по составу лексики идиоматических выражений — непосильная задача, поскольку данные операции подразумевают как минимум синтаксический анализ входного текста, в некоторых же случаях — и семантический анализ (а это уже элементы машинного перевода, основанного на правилах). Однако зачастую специалисты забывают об опции использования машинных переводчиков в качестве **электронных словарей**, предполагающих возможность постоянного пополнения параллельных лексических баз.

Очевидно, что, например, в базы Google-переводчика заложены достаточно полные русско-белорусский и белорусско-русский словари. Подтверждением их корректной работы служат как раз примеры точного перевода однословных конструкций неоднословными. Так, при двустороннем переводе безэквивалентной лексики белорусского и русского языков Google-переводчик во многих случаях выдает неоднословные соответствия: *чыгунка — железная дорога, железная дорога — чыгунка, адвеку — испокон веков, испокон веков — спрадвеку, адвеку* [1]. Стоит отметить, что «пословная стратегия» корректно работает при переводе коротких предложений. Предложения типа *мы ехали по железной дороге* переводятся преимущественно правильно, предложение же *мы ехали по железной дороге из Минска в Москву* оказывается уже слишком сложным, и *по железной дороге* переводится как *на железной дороге*. Электронные онлайн-словари, в свою очередь, используются именно для перевода коротких текстовых фрагментов.

На протяжении нескольких лет нами исследуются ЛСГ и ФСК глаголов восприятия в русском и белорусском языках [2; 3; 4]. Глаголы восприятия, являясь базовыми и незаменимыми, регулярно становятся объектом перевода. Редкий текст обходится без глагольных слов и перифрастических замен, выражающих семантику зрительного, слухового, обонятельного, осязательного, вкусового восприятия. Анализ возможностей Google-переводчика как электронного словаря относительно обработки глагольной лексики указанных групп позволил выявить ряд проблем в работе сервиса. Например, для

бел. глагола *засекчы* во фразе *Яны асцерагаліся, што тут могуць іх засекчы фаішыцкія дазорнікі* (А. Кулакоўскі) программа подбирает формально точный перевод *засечь*: *Они опасались, что здесь могут их засечь...* Однако среди вариантов перевода Google выдает и глаголы *зарезать, зарубить*, а это уже говорит о том, что машина относит периферийный глагол восприятия к иному синонимическому ряду — «путает» с единицей группы деструктивных глаголов.

Еще большие проблемы возникают при переводе неоднословных перифрастических конструкций с русскими и белорусскими функционально-текстовыми глаголами восприятия, относящимися по своему основному значению к ЛСГ глаголов перемещения в пространстве. Русскоязычную фразу *Тут у меня глаза разбрелись так, что я не мог сладить с моими мыслями* (И. Долгоруков) программа переводит как *Тут у мяне вочы разбрыліся так, што я не мог справіцца з маімі думкамі*. Белорусскоязычную фразу *Ягоны позірк адразу, хоць і без спеху, перасунуўся з Дж. Дж. О'Молая на твар Стывэна* как *Взгляд его сразу, хотя и не спеша, передвинулся...* В первом случае оптимальным объектом выдачи должен был стать вариант *разышился*, во втором — *переместился*. Машина не отождествляет синтагму *глаза разбрелись* с глаголами *смотреть / посматривать*, синтагму *позірк перасунуўся* с глаголами *паглядзець / паглядваць*, из-за чего становится невозможным подбор нужного синонима. Следовательно, иноязычный пользователь, изучающий глагольную лексику при помощи данного ПО, останется на базовом уровне знания языка.

Наши наработки в области составления параллельных ФСК русских и белорусских глаголов восприятия могли бы существенно улучшить статистический машинный переводчик (в данном случае сервис Google): возможности статического переводчика полностью зависят от объема и качества параллельных текстов, мы же в наших исследованиях составляем и соотносим параллельные парадигмы русских и белорусских глаголов (создаем фрагмент словаря соответствий), выводим уточненные синонимические ряды, в которые включаем функционально-текстовые глаголы восприятия, относящиеся по своему основному значению к ЛСГ глаголов перемещения в пространстве и употребляющиеся в перифрастических заменах [2; 3; 4].

Литература

1. Google-переводчик [Электронный ресурс]. — Режим доступа: <http://translate.google.ru>. — Дата доступа: 15.05.2016.
2. Красковский, Н. И. Лексико-семантическая группа русских и белорусских глаголов восприятия (связь с другими ЛСГ) / Н. И. Красковский // Сборник работ 69-й научной конференции студентов и аспирантов БГУ, 14–17 мая 2012 г.,

Минск. В 3 ч. Ч. 3 / Белорус.гос. ун-т. — Минск : Изд. центр БГУ, 2013. — С. 241–245.

3. Красковский, Н. И. Лексико-семантическая группа глаголов восприятия в русском и белорусском языках: особенности структуры / Н. И. Красковский // Мова і літаратура : матэрыялы 70-й навук. канф. студэнтаў і аспірантаў філал. ф-та БДУ, Мінск, 24 крас. 2013 г. / пад рэд. К.А. Тананушкі. —Мінск :РІВШ, 2013. — С. 36–40.

4. Красковский, Н. И. К вопросу о составе словаря синонимов русского и белорусского языков / Н.И. Красковский // Карповские научные чтения: сб. науч. ст. Вып. 9: в 2 ч. Ч. 2 / редкол.: А.И. Головня (отв. ред.) [и др.] — Минск: «Белорусский Дом печати», 2015. — С. 26–29.