

УДК 802.0:61

М.В. Вилькевич, И.Л. Небогатикова (3 курс, ФМедФиБ), М.С. Коган, к.т.н., доц.

ОЦЕНКА МАШИННОГО ПЕРЕВОДА ТЕКСТОВ МЕДИЦИНСКОЙ ТЕМАТИКИ

Основным языком науки является английский, значит большинство научных статей написано именно на нем. У специалиста, которому необходимо получить какие-либо сведения из англоязычных источников, возникает вопрос: «Тратить ли массу времени на ручной перевод или же воспользоваться средствами машинного перевода (МП)?». В данной работе мы попытались ответить на этот вопрос относительно текстов медицинской тематики. Для исследований были использованы два типа наиболее доступных переводческих программ: PROMT Office Giant (разработчик ПроМТ), ранее рассматривалось его использование для перевода международной корреспонденции [1], и Pragma 1.0 (разработчик Trident Software).

В работе рассматривались 4 типа медицинских текстов: тексты из учебного пособия «Английский для медиков» [2] – образец самого простого уровня медицинского английского, «Key Words in Science and Technology. Unit 4. Medicine» [3] – пример более сложного языка, тексты из электронных журналов [4,5]: научно-популярная статья – как самый востребованный и распространенный вариант текстов, и научная статья – как пример наиболее сложного специализированного текста.

Мы разобрали 32 предложения (по 8 из каждого источника). Оценка качества машинного перевода проводилась по следующим параметрам [6].

Информативность. Оценивалась степень точности перевода оригинала, т.е. насколько верно передается смысл. В рамках данного эксперимента она определялась способностью читателя ответить на вопросы к тексту, т.е. правильно понять оригинальный текст. Чтобы оценить информативность, мы опросили 20 студентов третьего курса факультета медицинской физики и биоинженерии, т.е. людей, ориентирующихся в медицинских вопросах. Респондентам было предложено рассмотреть по 2 предложения из каждой группы текстов и ответить на три вопроса к каждому из них.

Правильный перевод слов и терминов. Достоинством переводчика является предложение нескольких значений слова (термина) на выбор. Чтобы оценить этот пункт, мы считали число неверно переведенных или вовсе непереведенных слов.

Грамматика. Определяется следующими критериями: согласование по времени, роду и падежу, использование времен, порядок слов. Для оценки подсчитывалось число грамматических ошибок в тексте.

Оценка машинного перевода на примере одного предложения приведена ниже.

Исходный текст.

Man's first real insight into the nature of infectious disease is known to have come only a hundred years ago.

Ручной перевод.

Первое реальное представление человека о природе инфекционных заболеваний, как известно, появилось лишь сто лет назад.

МТ (Pragma).

Первая настоящая интуиция мужчины в натуре инфекционной болезни известна, чтобы прибыть только сотня годы тому назад.

Анализ МТ.

Информативность. Зададим вопросы к тексту машинного перевода. Что является подлежащим в предложении? МТ дает ответ «интуиция» – это неверный ответ.

Представление о чем появилось? МТ не позволяет ответить на этот вопрос. Когда появилось представление о природе инфекционного заболевания? Нет ответа.

Правильный перевод слов и терминов. Неверно переведенные слова выделены курсивом. Всего 4 неверно переведенных слова.

Грамматика. Одна ошибка «сотня годы».

Наиболее информативными оказались переводы пособия «Английский язык для медиков», что объясняется простотой его текстов. Наибольшее количество ошибок в переводе терминов было допущено при работе со статьями. Однако при работе с научными статьями медицинской тематики в большинстве рассмотренных случаев электронные переводчики обеспечивали достаточную информативность, хоть и ошибались в переводе слов и терминов. Это указывает на возможность использования их в помощь людям, плохо владеющим языком, но имеющим достаточное представление о теме. В работе с научно-популярными статьями электронные переводчики проявили себя хуже из-за наличия в них сленга, разговорных слов и т.п. В переводах учебной литературы также было достаточное число ошибок перевода терминов. Количество ошибок согласования было одинаковым для всех рассмотренных текстов.

ЛИТЕРАТУРА:

1. К.Н.Васильев Автоматизация перевода международной корреспонденции организации// XXX неделя науки СПбГТУ. Ч.ХІ: Материалы межвуз. науч. конф.– СПб.: Изд-во СПбГТУ, 2002.

2. М.С.Муравейская, Л.К.Орлова Английский язык для медиков (учебное пособие). Москва: Изд-во «Флинта» 2000.

3. Mascull B. Unit 4. Medicine// Key Words in Science and Technology. UK. 1997. С. 80-105.

4. Helicobacter pilory// [http:// kidshealth.org](http://kidshealth.org).

5. G.Wang, Y.Li, X.Li Correlation of Three-dimensional Structures with the Antibacterial Activity of a Group of Peptides Designed Based on a Nontoxic Bacterial Membrane Anchor// www.pubmedcentral.gov 2003.

6. Л.Н.Беляева Лингвистические автоматы в современных информационных технологиях СПб 2001.