

ЭКСПЕРИМЕНТАЛЬНЫЙ АНАЛИЗ ИНДИКАТОРОВ ПЕРЕВОДИМОСТИ ЭКОНОМИЧЕСКИХ ТЕКСТОВ

С.О. Шереметьева, О.И. Калининкова

Южно-Уральский государственный университет, г. Челябинск, Россия

В настоящей статье анализируется влияние лингвистических характеристик англоязычных экономических текстов на качество их автоматического перевода на русский язык. Исследование проводится на основе достаточно представительного количества экспериментальных данных, полученных в результате применения наиболее популярных систем ПРОМТ и Google Translate к текстам предметной области макро- и мезэкономии, одной из наиболее важных в международном аспекте областей экономики, информационные источники которой требуют оперативного представления контента на различных национальных языках. В процессе эксперимента выявлены количественные характеристики переводческих ошибок, типичных для данной предметной области, каждой из указанных систем, и англо-русского направления перевода. Определены и сопоставлены характерные для каждой из систем индикаторы переводимости, вызывающие выявленные типы ошибок. На основе анализа полученных данных выработаны ориентированные на пользователя правила предредактирования исходного текста с целью устранения индикаторов переводимости. Показано, что применение сформулированных правил к исходному английскому тексту до введения его в систему машинного перевода дает пользователю возможность значительно повысить качество продукции этих систем без обращения к разработчикам.

Ключевые слова: машинный перевод, экономические тексты, индикаторы переводимости, предредактирование.

Введение

Проблемы перевода текстов по экономике как способа международного обмена и распространения жизненно значимой информации достаточно широко освещаются в лингвистической и переводческой литературе. Исследования в этой области ведутся во многих странах на материале различных языков и направлений перевода, чаще всего с английского на другой национальный язык (испанский, французский, немецкий, арабский, казахский, русский и т. д., и, наоборот, с неанглийского языка на английский) [1, 7, 10, 11]. Проблемы перевода экономических текстов связаны прежде всего со спецификой предметной области, которая отражается в терминологии, синтаксисе, лексико-семантических сдвигах и стиле изложения.

Все работы, анализирующие проблемы перевода экономических текстов, можно разделить на две группы. К первой группе относятся исследования, рассматривающие специфику экономических текстов, с целью обучить человека-переводчика наиболее эквивалентно выражать содержание оригинальных текстов на иностранном языке [7, 10, 11]. Ко второй группе работ принадлежат исследования, вызванные широким распространением и доступностью онлайн-систем машинного перевода (МП), популярность которых вызвана все возрастающим требованием к повышению оперативности выполнения переводческих работ. Результатом большей части этой группы является констатация несовершенства систем МП, перечисление частных, выделенных на очень небольшом материале ошибок и рекомендации обязательного постредактирования [1, 3], что, конечно не вызы-

вает никаких возражений. В отечественной литературе, практически не отмечается, что качество машинного перевода зависит не только от уровня разработки конкретной системы МП, например Google Translate [2] или ПРОМТ [5], но и от конкретных пар языков, направления перевода и лингвистических особенностей переводимого текста. Последние могут выступать в качестве так называемых индикаторов переводимости (translatability indicators) [12]. Этот термин возник в связи с исследованиями по машинному переводу и обозначает графические, лексические и синтаксические элементы текста, провоцирующие ошибки при машинном переводе. Для пользователя систем МП умение распознавать индикаторы переводимости (ИП) в тексте имеет неоспоримые преимущества. Во-первых, наличие ИП в тексте прогнозирует уровень качества машинного перевода, и при большом их количестве можно либо отказаться от машинного перевода текста, зная, что перевод будет заведомо ошибочным, либо, что намного более перспективно, устранить индикаторы переводимости из исходного текста путем предредактирования «плохих» для машинного перевода фрагментов текста. Выделяют два типа ИП: универсальные и специфические. Универсальные индикаторы переводимости вызывают проблемы при машинном переводе любых видов текстов на любых языках. Специфические индикаторы переводимости характерны для конкретных языков, направлений перевода, предметных областей и систем МП. Например, при оценке переводимости патентных текстов с английского языка на датский в системе PaTrans [8] выявлено, что наиболее проблематичными для

этой системы являются английские предложные группы, наречия и придаточные предложения, стоящие в начале предложения, а также длинные предложения с наречиями. Для повышения качества МП разрабатываются специальные методики устранения индикаторов переводимости из исходного текста, ориентированные на конкретные системы машинного перевода. К наиболее популярным из таких методик относится разработка правил, ограничивающих использование лексико-синтаксического инвентаря, которые часто называют «контролируемым языком» [9]. К сожалению, в отечественных исследованиях этому вопросу уделяется недостаточно внимания.

В настоящей статье делается попытка восполнить этот пробел. Описываются методика и результаты анализа влияния лингвистических характеристик англоязычных текстов по макро- и мегаэкономике на качество их машинного перевода на русский язык с целью выявления индикаторов переводимости и выработки правил контролируемого языка для повышения качества машинного перевода.

1. Специфика англоязычных текстов по макро- и мегаэкономике

Макроэкономика изучает совокупности связей между отдельными элементами национальной экономики, в то время как мегаэкономика исследует связи между различными национальными хозяйствами [4], что естественным образом определяет специфику используемых в текстах этой предметной области языковых средств, которые мы далее будем называть термином «подъязык». С целью выявления особенностей подъязыка макро- и мегаэкономики был составлен и проанализирован корпус научных статей этой предметной области из журнала *Economics: The Open-Access, Open-Assessment E-Journal* за 2015–2018 годы объемом 202 450 словоупотреблений. По результатам анализа были выявлены следующие характеристики подъязыка:

- насыщенность как собственной макро- и мегаэкономической терминологией, так и терминологией более узких областей экономики (финансов, менеджмента и т. д.), и как следствие – наличие большого количества многозначных терминов, например, термин «sales» может быть переведен «товарооборот» или как «объем продаж» в маркетинге и «выручка от реализации» в бухгалтерском учете, а термин «interest» может быть переведен термином «процентная ставка» или нетерминологической лексемой «интерес»;
- преобладание многокомпонентных терминов, например, «competitive abilities control», «import demand elasticity», «nominal exchange rate movements»;
- большое количество аббревиатур и сокращений: «CPI (Consumer Price Index)» – индекс потребительских цен; «NPV (Net Present Value)» – чистая приведенная стоимость; «WTO (World

Trade Organization)» – Всемирная торговая организация; «tn (trillion)» – триллион (трлн.);

- большая длина предложения (среднее предложение состоит из 26 слов);
- широкое распространение пассивных, безличных и неопределенно-личных конструкций, например, «As indicated in above equation, an increase of REER is associated with a depreciation of the RMB against foreign currency»;
- преобладание сложносочиненных и сложноподчиненных предложений с длинными рядами однородных членов, в которых преобладают существительные, прилагательные и неличные формы глагола: «Other sources of heterogeneity, such as extensive margins, import demand elasticity, different destinations, U.S. dollar peg, and the liberalization of trading rights is also important to the effects of exchange rate changes»;
- обилие предложных фраз в одном предложении: «Changes in consumer behavior can be connected with descent of shopping ability (for example amid crisis), with appearance of more cheap and/or qualitative substitutes, competitor's new advertising campaign etc.»;
- наличие многокомпонентных сочетаний, что вызывает синтаксическую многозначность: «The responses of exporters to exchange rate movements may be different across goods with different import demand elasticity»;
- широкое использование эллиптических конструкций: «Complex modern societies, characterised by specialisation, division of labour and social coordination, emerged through adaptation in different environments, based on the capacity and disposition of people to cooperate when influenced by political instruments».

2. Постановка и результаты эксперимента

С целью оценки эффективности автоматического перевода текстов предметной области макро- и мегаэкономики с английского языка на русский был проведен эксперимент с использованием двух онлайн-переводчиков Google Translate [2] и ПРОМТ [5]. Выбор именно этих систем обусловлен, во-первых, тем, что ПРОМТ и Google Translate являются наиболее популярными и легкодоступными онлайн-системами машинного перевода, и во-вторых, тем, что эти системы построены в рамках диаметрально различных подходов и по-разному осуществляют процесс перевода. Машинный переводчик ПРОМТ является лингвистической системой и выполняет перевод на основе заранее построенной базы знаний, которая включает двуязычные словари и лингвистические правила. В то время как Google Translate – это статистическая система с включением элементов технологии нейросетей, которая осуществляет перевод путем учета наиболее частотных вариантов межъязыковых эквивалентов на основе различных статистических характеристик предварительно построенных двуязычных корпусов текстов.

Анализ машинного перевода с помощью системы ПРОМТ

Тексты на входе МП	<i>(which)(1)</i> 1. Growth in China's manufacturing sector slowed in December as (2) a punishing crack-down on air pollution and a cooling (3) property market start to weigh on... 2. This study adopts (4) a discrete-time survival model (5) in our empirical investigation and further executes several extensions and robustness checks to the baseline results (6) (7)
Переводы на выходе МП	1. Рост в промышленном секторе Китая, *который замедляют (1) в декабре *как (2) ужесточение наказания в отношении загрязнения воздуха и *охлаждающегося (3) рынка недвижимости, начинает оказывать ... 2. Это исследование *принимает (4) *модель выживания дискретного времени (5) в нашем эмпирическом расследовании и далее выполняет несколько расширений и проверок надежности **результатам основания(6) (7)
Типы ошибок МП	(1) – синтаксическая ошибка; (2),(4) – семантические ошибки; (3) – стилистическая ошибка; (5) – терминологическая ошибка; (6) (7) – синтаксические ошибки
Английские ИП	(1) – бессоюзное присоединение придаточного предложения: пропущен союз «which»; (2) – многозначность союза «as»; (3), (4) – многозначность слов «cooling» и «adopts»; (5) – термин; (6) – многозначность предлог «to»; (7) – цепочка существительных

В задачи эксперимента входила проверка следующих гипотез:

1. Различные технологии МП в общем случае дают различное качество перевода одних и тех же лингвистических явлений в тексте, т. е. характеризуются различным набором индикаторов переводимости и/или различной релевантностью одних и тех же ИП для качества перевода.

2. На основе выделенных индикаторов переводимости можно сформулировать правила контролируемого языка, при использовании которых для предредактирования качество машинного перевода текста значительно повышается.

Для проверки гипотез из указанного в предыдущем разделе англоязычного корпуса был составлен подкорпус из 100 предложений, содержащих перечисленные специфические для подязыка макро- и мегаэкономики лингвистические явления.

На первом этапе эксперимента все предложения подкорпуса были переведены на русский язык системами ПРОМТ и Google Translate и полученные результаты проанализированы. При этом анализу подвергались как русские переводы (учитывались лексическая и грамматическая правильность перевода, соблюдение стилистических норм и лексическая покрываемость), так и лингвистические явления исходных английских предложений, послужившие причиной ошибок в переводе, т. е. собственно англоязычные индикаторы переводимости. В табл. 1 и 2 даны примеры переводов¹ и анализа двух одинаковых фрагментов английского предложения подкорпуса с помощью систем ПРОМТ и Google Translate. Ошибки в переводах и индикаторы переводимости, вызвавшие указанные

ошибки, выделены подчеркиванием и помечены соответствующими номерами. Для большей наглядности английские индикаторы переводимости выделены жирным шрифтом.

На втором этапе эксперимента все полученные ранее данные об ошибках в англо-русских переводах каждой из систем ПРОМТ и Google Translate, а также о вызвавших эти ошибки индикаторах переводимости были проанализированы и систематизированы. В табл. 3 перечислены основные типы и частота ошибок систем ПРОМТ и Google Translate при англо-русском переводе подкорпуса из 100 предложений подязыка макро- и мегаэкономики. Стилистические ошибки не учитывались.

Табл. 3 дает представление о ранжировании типов ошибок в автоматических англо-русских переводах, выполненных системами ПРОМТ и Google Translate, а примеры, приведенные в табл. 1 и 2, достаточно четко показывают типы индикаторов переводимости, которые вызывают переводческие ошибки данных систем.

Из табл. 3 видно, что наиболее частотными ошибками в переводах систем ПРОМТ и Google Translate являются терминологические. Например, термин «discrete-time survival model» («модель выживания с дискретным временем») (см. табл. 1 и 2) переведен неправильно обеими системами. Терминология, особенно многокомпонентная, – это очень опасный индикатор переводимости и, поэтому пользователь обязательно должен проверить правильность автоматически переведенных терминов либо по терминологическим словарям, либо с помощью интернет-поисковиков по методике, описанной в [6]. Последнее намного предпочтительнее, поскольку терминологические «бумажные» словари и их оцифрованные версии

¹ Переводы выполнены 20.02.2019.

Таблица 2

Анализ машинного перевода с помощью системы Google Translate

Тексты на входе МП	1. Growth in China's manufacturing sector (1) slowed in December as a punishing crackdown on (3) air pollution and a cooling (3) property market start to weigh on 2. This study adopts (4) a discrete-time survival model (5) in our empirical investigation and further executes several extensions (6) and (7) robustness checks (8) to the baseline results
Переводы на выходе МП	1. В декабре рост в <i>*производственном секторе</i> (1) Китая замедлился из-за <i>*жесточкого подавления</i> (2) загрязнения воздуха и <i>*охлаждения</i> (3) рынка недвижимости, что начинает сказываться на ... 2. Это исследование <i>*принимает</i> (4) <i>*модель выживания в дискретном времени</i> (5) в нашем эмпирическом исследовании и дополнительно выполняет несколько <i>*проверок расширений</i> (6) <i>*и</i> (7) <i>*устойчивости</i> (8) к базовым результатам.
Типы ошибок МП	(1), (5) – терминологические ошибки; (2), (4), (6), (8) – семантические ошибки; (3) – стилистическая ошибка; (7) – синтаксическая ошибка
Английские ИП	(1), (5) – термины; (2), (4) – многозначность компонентов словосочетания «punishing crackdown on» и слова «adopts», (3) – многозначность слова «cooling» (6), (7), (8) – двухкомпонентные сочетания и союз «and»

Таблица 3

Сравнительная статистика типов ошибок при англо-русском МП текстов по макро- и мегаэкономике

Система	Тип ошибки (частота)		
	Терминологическая	Лексико-семантическая	Синтаксическая
ПРОМТ	104	61	99
GOOGLE	42	27	50

имеют ограниченную покрываемость и просто не содержат достаточного количества терминов, особенно многокомпонентных. Что касается электронных открытых онлайн-словарей, то они выдают слишком большое количество вариантов переводов, в которых трудно ориентироваться, и могут содержать переводческие ошибки, т. к. пользователи, вносящие в эти словари новые термины, не всегда являются лексикографами или квалифицированными переводчиками. В помощь переводчику в процессе данного исследования параллельно с выделением индикаторов переводимости и выявлением неправильно переведенных терминов нами составлен глоссарий многокомпонентных терминов макро- и мегаэкономики.

Следующими по значимости (но не по частоте) для анализируемого подязыка и систем МП являются лексико-семантические ошибки, искажающие смысл текста, что, естественно, недопустимо. Такие ошибки, как и терминологические, должны предотвращаться в первую очередь. Индикаторами переводимости лексико-семантических ошибок в русском переводе в основном являются: а) английские многозначные лексемы в рамках одной или различных частей речи и б) многозначность синтаксиса, что ведет, например, к неправильному выбору опорного слова в словосочетании или к неправильному определению однородных членов предложения. Последнее,

как правило, вызывается такими индикаторами переводимости, как координирующие союзы. Особенно опасен индикатор переводимости «and» при его многократном употреблении в рамках одного английского предложения и дистантном расположении зависимых слов.

Очень частотными являются синтаксические ошибки, вызванные целым рядом индикаторов переводимости, к основным из которых относятся: длина и сложная структура предложения, дистантное расположение зависимых членов, союз «and», безличные предложения, предложения в пассивном залоге, бессоюзное присоединение придаточных предложений, цепочки существительных, а также предлоги, которые, как правило, многозначны и могут требовать после себя существительных в различных падежах (см. примеры в табл. 1 и 2).

В табл. 3 не включены стилистические ошибки, поскольку они достаточно часто являются следствием остальных типов ошибок и в общем не препятствуют пониманию переведенного текста. Стилистические ошибки также могут быть вызваны такими индикаторами переводимости английского текста, как отличный от принятого в русском языке английский порядок слов и выбор из всех синонимов слов таких, которые у русскоязычного пользователя вызывают ощущение, что «так не говорят». Например, для перевода слова «cooling» (см. табл. 1) лучше подходит эквивалент «сокращающегося».

Примеры влияния предредактирования на МП систем Google Translate и ПРОМТ

Оригинальный текст	Перевод Google Translate	Перевод ПРОМТ
<i>Firm age is measured by the difference between firm establishing year and the current year</i>	<i>* <u>Твердый</u> возраст измеряется разницей между годом <u>установления</u> фирмы и текущим годом</i>	<i>*<u>Устойчивый</u> возраст *изменен различием между *<u>устойчивым</u> годом <u>установления</u> и *<u>в</u> текущем году</i>
Предредакция (правило 4)	Перевод Google Translate	Перевод ПРОМТ
<i>The age of a firm is measured by the difference between the year of firm establishment and the current year</i>	<i>Возраст фирмы измеряется разницей между годом основания фирмы и текущим годом</i>	<i>Возраст фирмы *изменен различием между годом *<u>устойчивого учреждения</u> и в текущем году.</i>

По результатам второго этапа эксперимента с целью устранения индикаторов переводимости исходного английского текста посредством перефразирования (предредактирования) проблематичных фрагментов были сформулированы правила контролируемого языка. Ниже приведен фрагмент таких правил для англоязычных текстов по макро- и мегаэкономике:

1. Проверять перевод терминологии по методике, описанной в [6].
2. Предложения длиной более 20 слов разбить на более короткие.
3. Заменить многозначные лексемы на однозначные.
4. Цепочки существительных разбить на группы с предлогом «of».
5. Добавить союз перед бессоюзно присоединенным придаточным предложением.
6. Сложные предложения, содержащие множественную сочинительную связь, разбивать на простые и/или дублировать предлог.

На третьем, заключительном этапе эксперимента из исходного корпуса текстов объемом в 202 450 словоупотреблений был составлен второй (не пересекающийся с первым) подкорпус из 50 фрагментов предложений с индикаторами переводимости, т. е. с лингвистическими явлениями, которые на предыдущем этапе эксперимента были идентифицированы как наиболее трудные для машинного перевода. Предложения нового подкорпуса и/или их фрагменты были предредактированы с применением сформулированных правил контролируемого языка и затем автоматически переведены с помощью обеих систем машинного перевода. Полученные в результате переводы были прежде всего проверены на корректность терминологии, и в случае необходимости перевод термина был откорректирован либо с помощью составленного нами в процессе эксперимента глоссария, либо помощью интернет-поисковиков по методике [6].

Проведенный эксперимент показал, что правила контролируемого языка для предредактирования при переводе текстов предметной области макро- и мегаэкономики с английского языка на русский при использовании систем МП Google Translate и ПРОМТ значительно повышают каче-

ство продукции обеих систем, но не в равной степени. Для получения высокого качества переводов эти системы в общем случае требуют применения разных наборов правил контролируемого языка. При этом для английских текстов анализируемой предметной области система Google Translate требует значительно меньше усилий по предредактированию (и меньшего количества правил контролируемого языка), чем ПРОМТ, и может быть рекомендована в качестве приоритетной. В табл. 4 иллюстрируется влияние применения одного и того же правила предредактирования на качество машинного перевода обеих систем. Индикаторы переводимости (многозначная лексема «Firm», цепочки существительных «Firm age», «firm establishing year» и координирующий союз «and» выделены жирным шрифтом и подчеркиванием.

Отметим, что союз «and» в данном примере является индикатором переводимости только для системы ПРОМТ, но не для Google Translate. Видно, что применение правила 4 для перефразирования двух словосочетаний исходного текста полностью исправило перевод Google Translate, но не дало желаемого результата для системы ПРОМТ. Размеры статьи не позволяют нам показать правила предредактирования, которые исправляют ошибки системы ПРОМТ при переводе предложения из табл. 4; читателю предлагается проделать это самостоятельно.

Заключение

В статье описаны методика и результаты экспериментального анализа влияния лингвистических характеристик англоязычных текстов по макро- и мегаэкономике на качество их машинного перевода на русский язык онлайн-системами ПРОМТ и Google Translate, разработанными в рамках диаметрально противоположных подходов: лингвистического и статистического соответственно. В процессе эксперимента проверены и подтверждены гипотезы исследования о том, что, с одной стороны, несмотря на различие технологий, автоматизирующих перевод, обе системы не справляются с переводом определенных лингвистических явлений подъязыка (индикаторов переводимости), от которых во многом зависит качество перевода, а с другой стороны, различие в тех-

нологиях обработки текста естественным образом ведет к тому, что эти системы кроме общих характеризуются специфичными для каждой их них индикаторами переводимости. Показано, что индикаторы переводимости можно устранить из исходного текста с помощью предредактирования на основе контролируемого языка, учитывающего индикаторы переводимости, и таким образом значительно повысить качество автоматического перевода без обращения к разработчикам систем. В случаях, когда переводной текст должен удовлетворять высоким критериям качества перевода, может возникнуть необходимость сочетать предредактирование и постредактирование.

Литература

1. Аль-Утайби Султан Маджед. Анализ машинного перевода текстов политического характера (газетного) с русского языка на арабский. Основные семантические трудности и переводческие решения / Аль-Утайби Султан Маджед. – URL: <https://fac.ksu.edu.sa/sites/default/files/2105.pdf> (дата обращения 20.02.2019).
2. Веб-сервис Google Translate. – URL: <https://translate.google.ru> (дата обращения 20.02.2019).
3. Кочеткова, Н.С. Особенности машинного перевода / Н.С. Кочеткова, Е.В. Ревина // Вопросы теории и практики. – 2017. – № 6 (72). – С. 106–109.
4. Микро-, макро- и мегаэкономика. Высшая школа экономики. <http://ecopos.ru/mikro,-makro-i-megaekonomika.php> (дата обращения 20.02.2019).
5. Онлайн-переводчик ПРОМТ. – URL: <https://translate.ru> (дата обращения 20.02.2019).
6. Шереметьева, С.О. К вопросу об электрон-

ных ресурсах профессиональной лексики / С.О. Шереметьева, П.Г. Осминин, Е.С. Щербаков // Вестник ЮУрГУ. Сер. Лингвистика. – 2014. – № 2 (37). – С. 230–232.

7. Muscan, M. On the translatability of cultural specifics in economic texts / M. Muscan. – URL: https://www.researchgate.net/publication/298973806_On_the_translatability_of_cultural_specifics_in_economic_texts (дата обращения 19.02.2019).

8. Povlsen, C. Evaluating Text-type Suitability for Machine Translation a Case Study on an English-Danish MT System / C. Povlsen // Proceedings of the First International Conference on Language Resources and Evaluation (Granada, Spain). – 1998. – Vol. 1. – P. 27–31.

9. Reuther, U. Two in one – Can it work? Readability and Translatability by means of Controlled Language / U. Reuther // Proceedings of EAMT-CLAW03, Controlled Language Translation (Dublin). – 2003. – P. 124–132.

10. Siranush, M. Translation Peculiarities of Economic Texts / M. Siranush. – URL: http://publications.yzu.am/wp-content/uploads/2015/12/Siranush_Meloyan.pdf (дата обращения 19.02.2019)

11. Tagkas, P. Translation of Economic Texts: Challenges and Limitations / P. Tagkas // In: Arslan, F., ed. Contemporary Issues on Linguistics and Language: LILA'14/Linguistics and Language Conference Proceedings Istanbul: DAKAM Publishing, June 2014. – P. 279–287.

12. Underwood, N.L. Translatability Checker: A Tool to Help Decide Whether to Use MT / N.L. Underwood, B. Jongejan // Proceedings of MT Summit VIII (Santiago de Compostela, Spain), 200. – P. 125–133.

Шереметьева Светлана Олеговна, доктор филологических наук, профессор кафедры лингвистики и перевода, Институт лингвистики и международных коммуникаций, Южно-Уральский государственный университет (Челябинск), linklana@yahoo.com

Калининкова Ольга Ивановна, магистрант кафедры лингвистики и перевода, Институт лингвистики и международных коммуникаций, Южно-Уральский государственный университет (Челябинск), olyakalinnikova@mail.ru

Поступила в редакцию 3 марта 2019 г.

DOI: 10.14529/ling190206

EXPERIMENTAL ANALYSIS OF ECONOMIC TEXT TRANSLATABILITY INDICATORS

S.O. Sheremetyeva, linklana@yahoo.com

O.I. Kalinnikova, olyakalinnikova@mail.ru

South Ural State University, Chelyabinsk, Russian Federation

This article analyzes the influence of the specificity of English-language economic texts on the quality of their machine translation into Russian. The study is conducted on the basis of a fairly representative amount of experimental data obtained as a result of applying popular PROMT and Google

Translate online systems to the texts on macro- and megaeconomics, one of the most important economic sub-domains, information sources of which require operative presentation of their content in various national languages. In the course of the experiment, the types and quantitative characteristics of translation errors, typical for each of the above machine translation systems, domain in question and the English-to-Russian translation direction were examined. Translatability indicators characteristic for each of the systems that cause the identified types of errors were detected and compared. Based on the analysis of the obtained data, rules for pre-editing the source text were developed in order to eliminate translatability indicators. It is shown that reformulating of original English texts following the pre-editing rules prior to machine translation allows the user significantly improving translation quality without contacting the developers.

Keywords: machine translation, economic texts, translatability indicators, pre-editing.

References

1. Al-Utajbi Sultan Madzhed. *Analiz mashinnogo perevoda tekstov politicheskogo haraktera (gazetnogo) s russkogo yazyka na arabiskij. Osnovnye semanticheskie trudnosti i perevodcheskie resheniya* [Analysis of machine translation of political texts (newspaper) from Russian into Arabic. The main semantic difficulties and translation solutions]. URL: <https://fac.ksu.edu.sa/sites/default/files/2105.pdf> (accessed 20.02.2019).
2. Web service Google Translate. URL: <https://translate.google.ru> (accessed 20.02.2019)
3. Kochetkova N.S. Osobennosti mashinnogo perevoda [The Features of Machine Translation]. N.S. Kochetkova, E.V. Revina. *Voprosy teorii i praktiki* [Theory and Practice Issues], 2017, no. 6 (72), pp. 106–109.
4. Mikro, makro i megaeconomika. Vysshaya shkola ekonomiki (Micro, macro and mega economics. Higher School of Economics) URL: <http://ecopos.ru/mikro,-makro-i-megaekonomika.php> (accessed 20.02.2019).
5. Online translator PROMT. URL: <https://translate.ru> (accessed 20.02.2019).
6. Sheremetyeva S.O., Osminin P.G., Scherbakov E.S. [On Electronic Resources for Professional Lexics]. *Bulletin of the South Ural State University. Ser. Linguistics*. 2014, no. 2 (37), pp. 230–232.
7. Muscan M. *On the translatability of cultural specifics in economic texts*. URL: https://www.researchgate.net/publication/298973806_On_the_translatability_of_cultural_specifics_in_economic_texts (accessed 19.02.2019).
8. Povlsen C. Evaluating Text-type Suitability for Machine Translation a Case Study on an English-Danish MT System. *Proceedings of the First International Conference on Language Resources and Evaluation (Granada, Spain)*, 1998, vol. 1, pp. 27–31.
9. Reuther U. Two in one – Can it work? Readability and Translatability by means of Controlled Language. *Proceedings of EAMT-CLAW03, Controlled Language Translation (Dublin)*, 2003, pp. 124–132.
10. Siranush M. *Translation Peculiarities of Economic Texts*. URL: http://publications.ysu.am/wp-content/uploads/2015/12/Siranush_Meloyan.pdf (accessed 19.02.2019)
11. Tagkas P. Translation of Economic Texts: Challenges and Limitations. In: Arslan, F., ed. *Contemporary Issues on Linguistics and Language: LILA'14/Linguistics and Language Conference Proceedings Istanbul*: DAKAM Publishing, June 2014, pp. 279–287.
12. Underwood, N.L., Jongejan B. Translatability Checker: A Tool to Help Decide Whether to Use MT. *Proceedings of MT Summit VIII (Santiago de Compostela, Spain)*, 2001, pp. 125–133.

Svetlana O. Sheremetyeva, Doc. of sc., professor of the Linguistics and Translation Department, Institute of Linguistics and Intercultural Communication South Ural State University, linklana@yahoo.com

Olga I. Kalinnikova, graduate student of the Linguistics and Translation Department, Institute of Linguistics and Intercultural Communication South Ural State University, olyakalinnikova@mail.ru

Received 3 March 2019

ОБРАЗЕЦ ЦИТИРОВАНИЯ

Шереметьева, С.О. Экспериментальный анализ индикаторов переводимости экономических текстов / С.О. Шереметьева, О.И. Калинникова // Вестник ЮУрГУ. Серия «Лингвистика». – 2019. – Т. 16, № 2. – С. 44–50. DOI: 10.14529/ling190206

FOR CITATION

Sheremetyeva S.O., Kalinnikova O.I. Experimental Analysis of Economic Text Translatability Indicators. *Bulletin of the South Ural State University. Ser. Linguistics*. 2019, vol. 16, no. 2, pp. 44–50. (in Russ.). DOI: 10.14529/ling190206