

УДК 81'255

DOI: 10.18384/2310-712X-2016-2-164-171

АВТОМАТИЗАЦИЯ ПЕРЕВОДА И ТИПОЛОГИЯ ТЕКСТОВ

Марчук Ю.Н.

*Московский государственный областной университет
105005, г. Москва, ул. Радио, д. 10А, Российская Федерация*

Аннотация. Практическая типология переводимых текстов весьма актуальна для использования формальных методов (автоматических словарей, систем машинного перевода), особенно для перевода научно-технических текстов. Актуальность проблемы практической типологии состоит в необходимости повышения качества использования формальных методов, алгоритмических процедур и новых информационных технологий, а также тем обстоятельством, что перевод научно-технических текстов составляет основной массив переводов в мире (от 80% до 90% общего объема переводов).

Ключевые слова: перевод, научно-технические тексты, автоматизированные словари, машинный перевод, повышение качества перевода

MACHINE TRANSLATION AND PRAGMATIC TYPOLOGY OF TEXTS

Yu. Marchuk

*Moscow State Regional University
105005, Moscow, Radio street, 10A, Russian Federation*

Abstract. Pragmatic typology of translated texts is very important for formal methods, such as automatic dictionaries, machine translation systems, especially for translation of scientific and technical texts. Pragmatic typology helps to improve the quality of translation and use of new informational technologies for translation. Its importance lies in the fact that scientific and technical texts constitute from 80% to 90% of the general volume of translations in the world.

Key words: translation, scientific and technical texts, automatic dictionaries, machine translation, better quality of translation

Типология текстов, подлежащих переводу, находится как бы на периферии переводоведческой науки, которая основное внимание уделяет моделям перевода, проблемам адекватности, эквивалентности, соотношению языков и культур и пр. Между тем следует принять во внимание то обстоятельство, что перевод, осуществляемый человеком, есть процесс достаточно медленный, дорогой и с трудом поддается совершенствованию в рамках новых информационных технологий, которые все в большей степени овладевают современным миром коммуникации.

Данные европейских комиссий, изучавших проблемы перевода, а также статистика, накопленная в деятельности Всесоюзного Центра переводов научно-

технической литературы и документации СССР, головной организации по научно-техническим переводам (в настоящее время – Всероссийский Центр переводов) показывают, что основная масса письменных переводов в мире – научно-технические переводы. Это и понятно, поскольку именно такого рода переводы обеспечивают в существенной части научно-технический прогресс человеческого общества.

Чтобы правильно перевести научно-технический текст, требуется хорошее знание предметной области, о которой идет речь в переводе, тем более что новые тексты как правило содержат описание новых достижений науки и техники, хорошее знание двух языков (языка текста и языка перевода), умение четко излагать свои мысли. Такие качества достаточно редко сочетаются у одного человека, поэтому перевод часто требует достаточно больших усилий для корректировки редактором. Представим себе перевод описания нового технического устройства. Ошибки в таком переводе могут не только помешать использованию устройства, но и привести к трагедиям. Неправильный перевод патентов может повлечь судебный процесс.

В настоящее время действуют системы машинного перевода текстов достаточно различного качества, для разных языковых пар. Проблема машинного перевода не нова, в пятидесяти годах прошлого столетия она была одной из актуальных проблем лингвистики и перетерпела долгий путь развития.

Альтернативы машинному переводу нет. Использование автоматических словарей в деятельности человека-переводчика также подключает его к

новым информационным технологиям, но при этом сам процесс остается «человеческим» и не обладает необходимой быстротой и дешевизной. История машинного перевода показывает, что разработка высокоэффективных систем чрезвычайно сложна, требует решения ряда кардинальных проблем моделирования человеческого интеллекта, результат зависит не только и не столько от применения новых решений в рамках информационной технологии, но и от эффективности моделирования мышления. Академик Ю.В. Рождественский отмечает, что машинный перевод является кардинальной проблемой искусственного интеллекта. Это и неудивительно, если мы представим себе, какие задачи приходится решать человеку-переводчику в сфере мышления, межкультурной коммуникации, передачи не только значения слов, но и учета контекста коммуникации, понимаемого в самом широком смысле этого слова [11].

Как сказано выше, типология текстов, подлежащих переводу, не привлекала большого внимания переводоведов. Тем не менее, в теории перевода намечена по крайней мере общая тенденция исследований в этой области. Так, ведущий переводовед В.Н. Комиссаров в своей книге «Современное переводоведение» [5] посвятил специальный раздел текстологическим аспектам переводоведения. В нем отмечается, что большое значение для современного переводоведения имеют результаты исследований в области лингвистики текста. Осуществление перевода и оценка его результатов предполагают умение сопоставлять форму и содержание различных типов текста, учитывать особенности струк-

туры и функционирования текстов в каждом из языков, анализировать соотношение текста как целостного образования с составляющими текст языковыми единицами и структурами. Все это становится возможным на основе положений и подходов лингвистики текста, утверждает автор.

Далее, с точки зрения теории переводоведения, содержательная структура текста рассматривается в трех разных измерениях: вертикальном, горизонтальном и глубинном. Вертикальную структуру текста создает его формально-тематическое содержание, начиная с общего замысла или темы текста. Горизонтальная структура создается формальными и смысловыми связями между высказываниями. Формальная связность текста достигается с помощью различных языковых средств – союзов, повторов и т.п. Смысловое единство текста – его когерентность – обеспечивается логической последовательностью и непротиворечивостью изложения. Непосредственно связана с когерентностью его тема-рематическая структура, или актуальное членение. Особое значение для теории и практики перевода имеет, по мнению В.Н. Комиссарова, детальное описание глубинной структуры содержания текста, отражающей процесс построения речевых высказываний и включения их в текст.

Примером попытки связать классификацию текстов оригинала со стратегией переводчика может служить концепция немецкой исследовательницы К. Райт, сообщает В.Н. Комиссаров. К. Райт предлагает различать три типа текстов: в первом типе доминирует функция описания, во втором – выражения, в третьем – призыва. Эти три

типа текстов, выделяемые на основе функций языка, предлагается дополнить четвертым типом – аудиомедиаальным, к которому относятся тексты, предназначенные для использования в устной форме.

Такая типология текстов непосредственно увязывается с задачами перевода. Предполагается, что первый тип текстов ориентирован в первую очередь на содержание, и при переводе этих текстов (коммерческих, научных, деловых и т.п.) задача переводчика заключается в том, чтобы как можно полнее передать содержание. Второй тип текста считается ориентированным на форму (художественная литература) и задача переводчика заключается прежде всего в сохранении художественно-эстетического воздействия оригинала. Третий тип текста, ориентированный на обращение, стремится к созданию определенного экстралингвистического эффекта и поэтому в переводе должно быть передано обращение к читателю или слушателю. Аудиомедиаальные тексты (тексты радио и телепередач, сценических произведений и т.п.) составляют часть более крупного целого, поэтому при их переводе надо прежде всего учитывать условия среды, в которой они будут использоваться.

Как видим, в данном случае имеет место достаточно глобальная характеристика переводимых текстов.

В.Н. Комиссаров далее приводит переводческую типологию текстов, разработанную немецким переводоведом А. Нойбергом. Он попытался классифицировать переводимые тексты на основе их прагматической ориентации. С такой точки зрения предлагается различать четыре типа текстов,

отличающихся по степени переводимости в прагматическом плане. В первом типе текстов у оригинала и перевода имеются общие цели, основанные на общих потребностях. Содержание оригинала не предназначено исключительно для аудитории исходного языка (научно-техническая литература, рекламные объявления и т.п.). Общность цели означает, что и у аудитории языка перевода могут возникнуть такие же прагматические отношения к сообщаемому, и, следовательно, такие тексты обладают высшей степенью переводимости с прагматической точки зрения.

Второй тип отношений в этой концепции характерен для текстов, исключительно предназначенных для аудитории исходного языка (официальные распоряжения, местная информация, развлекательные материалы и т.п.). Тексты законов, общественно-политическая литература, местная пресса настолько специфичны в прагматическом отношении, что, по мнению А. Нойберта, такие тексты в этом смысле принципиально непереводимы.

Третий тип текстов – художественная литература – может выражать и общечеловеческие ценности. Степень переводимости зависит от жанра: беллетристика и драматургия обладают более высокой степенью переводимости, чем лирическая поэзия.

Четвертый тип текстов создается на исходном языке, но предназначается для перевода на другой язык и изначально ориентирован на аудиторию переводного языка (публикации для зарубежных стран). Этот тип текстов обладает высокой степенью прагматической переводимости.

Подводя итоги вышесказанному, В.Н. Комиссаров констатирует, что

текст выступает как цель, объект и результат перевода и текстологические проблемы должны занимать важное место в современном переводоведении.

Надо заметить, что и в других исследованиях, посвященных теории и моделям перевода, например, в разработках В.Н. Крупнова, Л.Л. Нелюбина, Л.И. Борисовой и многих других, в формулировках и описаниях различных моделей безусловно есть ссылка на те или иные типы переводимых текстов в связи с эффективностью той или иной модели перевода [2; 4; 10]. Однако в самостоятельное направление типология текстов обычно не выделяется. Между тем, когда мы рассматриваем такой вопрос как низкое качество работы современных систем машинного перевода, мы приходим к выводу, что это качество в большей степени зависит от некоторых характеристик переводимых текстов, которые (характеристики) отнюдь не лежат на поверхности. Часто человеку непонятно, почему один текст удовлетворительно переводится системой машинного перевода, а другой, совершенно похожий по тематике, структуре, принадлежности к определенному предметному полю, получает от этой же системы совершенно неприемлемый перевод. Это приводит к мысли, что компьютер тонко реагирует на некоторые особенности текста, которые не улавливаются глобальными характеристиками текстов, подобными тем, которые приведены выше. Вследствие этого особое внимание надо уделить использованию информационных технологий в лингвистике [6].

В первую очередь низкое качество автоматического перевода заметно на переводе лексики. Лексика научно-

технических текстов может делиться, как указала Л.И. Борисова, в двух возможных направлениях. Первое: можно различать общеупотребительную, общенаучную и терминологическую лексику, т.е. три слоя. Однако есть и второй подход: различать только общеупотребительную (общенаучную) и терминологическую лексику, т.е. не отделять общеупотребительную от общенаучной. Л.И. Борисова придерживается второго подхода и убедительно его обосновывает. В научно-технических текстах общеупотребительная лексика часто приобретает значения, не зафиксированные в обычных словарях. Границу между собственно терминами определенного предметного поля и общеупотребительными словами часто практически невозможно провести. С.В. Гринев-Гриневиц убедительно показывает роль терминологии в становлении и развитии человеческого языка [3]. Современные системы машинного перевода, как правило, предлагают пользователю некоторые наборы словарей по разным предметным областям, при этом отделяя их от общеупотребительного словаря [12]. Тем не менее, взаимодействие таких словарей часто происходит таким образом, что появляются самые явные и нелепые ошибки в переводе слов и словосочетаний.

Специалисты по автоматическому переводу уже осознали необходимость более точной типологии текстов, подлежащих переводу. Так, Л.Н. Беляева предлагает довольно детальную типологию переводимых текстов, в рамках которой учитываются такие компоненты как структура текстов, функциональный стиль, регистр, жанрово-стилистические особенности, форма

презентации текстов, отношение субъекта к переводу текста, основная прагматическая функция, иллюкутивные цели и другие характеристики [1]. Особое место в работе Л.Н. Беляевой занимают рекомендации использования систем машинного перевода и систем автоматических словарей в помощь переводчику. Кроме конкретных примеров работы с практическими системами машинного перевода, аспектами пред- и постредактирования машинного продукта, в исследовании Л.Н. Беляевой уже предусмотрен этап предварительного автоматического анализа текста, подлежащего переводу, а именно выделение различного рода формальных признаков деления связного текста на предложения, структурный анализ текста, опора на формальные признаки, позволяющие определить границы анализа и т.п. Можно, однако, отметить, что при этом основная часть работы по типологизации текстов, решения о применении тех или иных машинных словарей и другие аспекты машинной обработки текстов требуют осмысления со стороны человека-редактора. Между тем хотелось бы полностью передать задачу о выборе словарей, способов анализа и синтеза выходного текста и другие задачи перевода на решение системы машинного перевода, ориентированной на данный тип текстов, тип, который сама система могла бы установить.

Определенным шагом в этом направлении можно считать контекстологический словарь, который составлен был в коллективе разработчиков системы АМПАР довольно давно, но идея которого продолжает оставаться актуальной и до настоящего времени [Марчук 1976]. Контекстологический

словарь составляется на основе конкорданса из текстов исходного языка и текстов соответствующих переводов (параллельных текстов). Контекст при этом понимается как основное лингвистическое окружение многозначного слова. Выбор нужного перевода осуществляется работой алгоритмов, исследующих контекст [8; 9]. Такой словарь уже задает определенный тип текстов, что позволяет ориентировать систему машинного перевода также и в анализе других составляющих текста, кроме лексики и выбора терминологического словаря.

Таким образом, типологию текстов, которая удовлетворяла бы требованиям определенной системы машинного перевода и позволяла получать достаточно качественный машинный перевод без существенного вмешательства человека-редактора, можно строить с учетом двух основных факторов. Первый – определение возможностей данной конкретной системы машинного перевода в области перевода лексики. Так, если в ней отсутствует контекстологический словарь, что является основной чертой современных систем (составление такого словаря весьма трудоемкое дело), то типологию можно строить исходя из проверки текстов на наличие/отсутствие определенной категории терминов и/или общеупотребительных слов. Проверка текстов на наличие определенных категорий терминов могла бы помочь определить, какой из тематических терминологических словарей следует использовать в первую очередь. При этом в системе словарей можно исходить из того, что

словарь общеупотребительной лексики будет взаимодействовать с некоторым конечным набором отраслевых словарей, содержащих термины конкретной предметной области. Эти последние словари могут пополняться в зависимости от потребностей новых терминов для конкретных текстов.

Второй фактор – *опора на возможности системы с учетом особенностей конкорданса параллельных текстов, на основе которого построена система*. В данном случае может не потребоваться опытный переводчик текстов. Типологию переводимых текстов в данном случае составляет человек-редактор, исходя из общих характеристик текстов, послуживших основой для составления системы. Примером такой типологии, довольно успешно действующей, является Гренобльская система, предназначенная и используемая в промышленном режиме для перевода текстов текстильной промышленности [7]. При переключении такого рода системы на перевод текстов, обладающих другими характеристиками, основное внимание должно уделяться структуре системы, возможностям адаптации данной структуры к характеру новых текстов.

Таким образом, отметим еще раз необходимость тщательной разработки типологии текстов, подлежащих машинному переводу и другим видам автоматической обработки с помощью новых информационных технологий, для того, чтобы результат был наиболее приемлем с точки зрения качества и требовал бы минимальных усилий пред-, интер- или постредактора.

ЛИТЕРАТУРА:

1. Беляева Л.Н. Теория и практика перевода. СПб.: РГПУ им. А.И. Герцена, 2007. 211 с.
2. Борисова Л.И. Лексические особенности англо-русского научно-технического перевода. М.: Тезаурис, 2005. 215 с.
3. Гринев-Гриневиц С.В. Терминоведение. М.: Академия, 2008. 303 с.
4. Крупнов В.Н. Гуманитарный перевод. М.: Академия, 2009. 159 с.
5. Комиссаров В.Н. Современное переводоведение. М.: ЭТС, 2004. 421 с.
6. Марчук Ю.Н. Информационные технологии в лингвистике. Компьютерная лингвистика. Saarbrücken: Palmarium Academic Publishing, 2015. 131 с.
7. Марчук Ю.Н. Компьютерная лингвистика. М.: Восток-Запад, 2007. 317 с.
8. Марчук Ю.Н. Контекстологический словарь для машинного перевода многозначных слов с английского языка на русский. Ч. I. М.: Всесоюзный Центр переводов, 1976. 264 с.
9. Марчук Ю.Н. Контекстологический словарь для машинного перевода многозначных слов с английского языка на русский. Часть II. М.: Всесоюзный Центр переводов, 1976. 256 с.
10. Нелюбин Л.Л. Введение в технику перевода. М.: МГОУ, 2007. 187 с.
11. Рождественский Ю.В. Философия языка. Культуроведение и дидактика. Современные проблемы науки о языке. М.: ГрантЪ, 2003. 239 с.
12. Хроменков П.Н. Современные системы машинного перевода. М.: МГОУ, 2005. 159 с.

REFERENCES:

1. Belyaeva L.N. Teoriya i praktika perevoda [Theory and practice of translation]. SPb., RGPU im. A.I. Gertsena, 2007. 211 p.
2. Borisova L.I. Leksicheskie osobennosti anglo-russkogo nauchno-tehnicheskogo perevoda [Lexical peculiarities of English-Russian scientific and technical translation]. M., Tezaurus, 2005. 215 p.
3. Grinev-Grinevich S.V. Terminovedenie [Terminology]. M., Akademiya, 2008. 303 p.
4. Krupnov V.N. Gumanitarnyi perevod [Humanitarian translation]. M., Akademiya, 2009. 159 p.
5. Komissarov V.N. Sovremennoe perevodovedenie [Contemporary translation studies]. M., ETS, 2004. 421 p.
6. Marchuk YU.N. Informatsionnye tekhnologii v lingvistike. Komp'yuternaya lingvistika. [Information technologies in linguistics. Computational linguistics]. Saarbrücken, Palmarium Academic Publishing, 2015. 131 p.
7. Marchuk YU.N. Komp'yuternaya lingvistika [Computational linguistics]. M., Vostok-Zapad, 2007. 317 p.
8. Marchuk YU.N. Kontekstologicheskii slovar' dlya mashinnogo perevoda mnogoznachnykh slov s angliiskogo yazyka na russkii. CHast' I. [Contextual dictionary for machine translation of the polysemantic words from English into Russian. Part I]. M., Vsesoyuznyi Tsentr perevodov, 1976. 264 p.
9. Marchuk YU.N. Kontekstologicheskii slovar' dlya mashinnogo perevoda mnogoznachnykh slov s angliiskogo yazyka na russkii. CHast' II. [Contextualise dictionary for machine translation of the polysemantic words from English into Russian. Part II]. M., Vsesoyuznyi Tsentr perevodov, 1976. 256 p.
10. Nelyubin L.L. Vvedenie v tekhniku perevoda [Introduction to translation techniques]. M., MGOU, 2007. 187 p.
11. Rozhdestvenskii YU.V. Filosofiya yazyka. Kul'turovedenie i didaktika. Sovremennye problemy nauki o yazyke [Philosophy of language. Cultural studies and didactics. Modern problems of the language science]. M., Grant', 2003. 239 p.

12. Khromenkov P.N. *Sovremennye sistemy mashinnogo perevoda* [Modern machine translation systems]. M., MGOU, 2005. 159 p.
-

ИНФОРМАЦИЯ ОБ АВТОРЕ

Марчук Юрий Николаевич – доктор филологических наук, профессор, зав. кафедрой теоретической и прикладной лингвистики Московского государственного областного университета; e-mail: marchuk43@migtel.ru

INFORMATION ABOUT THE AUTHOR

Marchuk Yurij N. – PhD of Philology, professor, head of the chair of theoretical and applied linguistics of Moscow State Region University; e-mail: marchuk43@migtel.ru

БИБЛИОГРАФИЧЕСКАЯ ССЫЛКА

Марчук Ю.Н. Автоматизация перевода и типология текстов // Вестник Московского государственного областного университета. Серия: Лингвистика. 2016. № 2. С. 164–171. DOI: 10.18384/2310-712X-2016-2-164-171

BIBLIOGRAPHIC REFERENCE

Yu. Marchuk. Machine translation and pragmatic typology of texts // Bulletin of Moscow State Regional University. Series: Linguistics. 2016. no 2. pp. 164–171. DOI: 10.18384/2310-712X-2016-2-164-171