

Д.В. Степанова (Минск, МГЛУ)

## **К ВОПРОСУ О СОПОСТАВЛЕНИИ АНГЛИЙСКИХ ТЕРМИНОЛОГИЧЕСКИХ СЛОВСОЧЕТАНИЙ И ИХ ПЕРЕВОДОВ НА РУССКИЙ ЯЗЫК**

Изучение результатов переводческого процесса методом сопоставления как текстов переводов, так и переводов с исходными текстами или оригиналами, лежит в основе общей и частной теории переводов. Накопленный в результате сопоставительного исследования целых томов оригиналов и их переводов богатый фактический материал использовался и продолжает использоваться для проведения разноплановых лингвистических исследований, например, в области лексикографии, переводоведения, стилистики, лексикологии. Однако до недавнего времени в языкознании анализ эмпирических данных выполнялся вручную, что ограничивало объем исследуемой информации и обуславливало небольшую скорость ее получения и обработки.

Появление компьютерной техники, постоянное увеличение ее мощностей и обеспечение массового доступа к персональным компьютерам позволяет оперативно выполнять различные операции с электронными текстами.

Особый интерес для теории перевода представляют собой широкодоступные электронные корпуса параллельных текстов, в которых зафиксирован огромный опыт перевода, выполненный человеком. Под корпусом параллельных текстов принято понимать оригинальные тексты и их переводы на другой язык, обладающие лингвистической и смысловой адекватностью по отношению к оригиналу [3, с. 103].

Настоящая статья посвящена описанию опыта использования электронных корпусов параллельных текстов по вычислительной технике для сопоставительного анализа английских терминологических словосочетаний (ТС) и их русских переводов, который проводился в рамках исследования по созданию формальной модели системы с переводческой памятью, т. е. системы машинного перевода, позволяющей накапливать и использовать ранее переведенные текстовые фрагменты.

Как известно, перевод научно-технической литературы представляет собой достаточно трудоемкую работу именно вследствие употребления в специальных текстах значительного числа словосочетаний (по подсчетам разных специалистов 60–70 % от общего числа терминов), перевод которых, как правило, не может быть выполнен пословно и требует применения специальных приемов передачи лексических единиц с одного языка на другой.

Кроме того, научно-технический перевод осложняется и тем фактом, что в силу высоких темпов развития науки и техники даже самые современные специальные словари не успевают отразить все изменения и новинки в

составе терминологий различных предметных областей, что особенно актуально для находящейся в процессе формирования терминологии вычислительной техники.

Материалом для исследования послужили около 600 английских предложных и беспредложных ТС по вычислительной технике, выделенные из английской части составленного автором электронного англо-русского корпуса параллельных текстов данной предметной области.

Подбор русских переводных эквивалентов для выделенных с помощью статистических методов английских ТС проводился с опорой на четыре источника: русская часть электронного корпуса параллельных текстов и три современных англо-русских словаря по вычислительной технике [1; 2; 5]. В результате для 582 словосочетаний языка оригинала было обнаружено 905 терминов языка перевода, 589 из которых (65 %) взяты из текстов и 316 (35 %) – из словарей. Следует отметить, что из 589 русских эквивалентов из текстов большинство (81 %) были определены на основании статистических методов как наиболее употребительные русские составные термины подъязыка вычислительной техники.

Сопоставительный анализ английских ТС рассматриваемой предметной области и их переводов на русский язык осуществлялся в два этапа – анализ переводческих соответствий и анализ структурных моделей словосочетаний.

На первом этапе было проведено сопоставление эквивалентов, выделенных из русской части текстов, с эквивалентами, зафиксированными в словарях, что позволило разделить английские ТС на три группы в зависимости от вариантов их переводов.

В первую группу включено большинство ТС исходного языка (55 %), которые не зафиксированы в словарях. В результате для этих терминов переводные эквиваленты были выделены только из исследуемых текстов. Например, *application layer gateway service* – служба шлюза уровня приложения; *jagged array* – неровный массив; *peer client* – равноправный узел. При этом некоторым английским словосочетаниям соответствуют два синонимичных русских эквивалента, например, *language features* – средства языка; языковые средства; *boot file* – файл загрузки; загрузочный файл.

Вторая группа содержит словосочетания (32 %), текстовые переводные эквиваленты которых отражены и в указанных ранее словарях. В зависимости от наличия или отсутствия других вариантов переводов английских терминов, в данной группе было выделено две подгруппы:

а) ТС, которые имеют постоянные соответствия, или эквиваленты, в русском языке (13 %). При этом текстовые переводные эквиваленты большей части таких словосочетаний (41 %) отражены только в одном из исследуемых словарей, например, *net logon service* – служба регистрации абонентов сети; *network management protocol* – протокол управления сетью; *local network* – локальная сеть. В 32 % случаев текстовые переводные эквиваленты зафиксированы в двух словарях, например, *system resource* – системный ресурс; *system bus* – системная шина; *subnet mask* – маска подсети. В 27 % случаев все четыре рассмотренных источника русских вариантов переводов

предлагают единые эквиваленты, например, *universal serial bus* – универсальная последовательная шина, *machine instruction* – машинная команда, *device driver* – драйвер устройства;

б) ТС, которым свойственны вариантные соответствия в языке перевода (19%), например, *check box* – отмечаемая кнопка, кнопка с независимой фиксацией, позиция для отметки, флаговая кнопка, переключатель, поле для галочки, флажок; *computer name* – имя компьютера, сетевое имя. При этом в 4% словарных статей наряду с синонимичными вариантами перевода были зафиксированы и случаи полисемии, например, *power supply* – 1. питание, 2. источник питания; *data access* – 1. выборка данных, 2. доступ к данным.

В третью группу были отнесены словосочетания (13%), текстовые переводные эквиваленты которых являются синонимичными вариантами словарных переводных эквивалентов, например, *background image* – фоновый рисунок, фоновое изображение; *blank cell* – пустая ячейка, незаполненная ячейка, *bus architecture* – архитектура шины, шинная архитектура.

Проведенный анализ вариантов переводов выделенных английских ТС свидетельствует о вариативности терминологии подязыка вычислительной техники и отсутствии общепринятых терминов, что значительно усложняет передачу единиц исходного языка на язык перевода.

Кроме того, сравнительный анализ терминологической лексики, приведенной в специальных словарях и используемой в параллельных текстах, показал, что, большая часть словосочетаний остается неучтенными. Этот факт дает основание рассматривать корпус параллельных текстов по вычислительной технике в качестве источника получения новых терминов и их вариантов перевода, которые еще не отражены в словарях, но уже широко используются специалистами.

На втором этапе был проведен сопоставительный анализ структурных моделей английских и русских ТС, который позволил установить приемы перевода рассматриваемых лексических единиц на русский язык.

Выполнение данной задачи предполагало, в первую очередь, построение структурных моделей английских и русских терминов с указанием их ядерных компонентов, в качестве которых выступают опорные термины, выделенные формально при помощи статистических методов и выраженные различными частями речи.

Таким образом, в зависимости от грамматических классов опорных элементов были выделены четыре группы ТС: ТС с опорным термином, выраженным именем существительным, например, *scripting language* – язык написания сценариев ( $N_2+N_1 \rightarrow N_1+N_2+N_3^1$ ), ТС с опорным термином, выраженным именем прилагательным, например, *wireless hub* – беспроводной концентратор ( $A_1+N \rightarrow A_1+N$ ), ТС с опорным термином, выраженным глаголом, например, *to create a folder* – создать папку ( $V_1+N \rightarrow V_1+N$ ), и ТС с

---

<sup>1</sup> Структура словосочетания представлена в виде комбинации буквенных символов, каждый из которых означает грамматический класс слова. Используются следующие обозначения классов слов: А-прилагательное, N- существительное, V-глагол, Adv-наречие. Цифрой «1» отмечается опорный термин.

опорным термином, выраженным наречием, например, *automatically generate* – автоматически генерировать ( $AdvI+V \rightarrow AdvI+V$ ). При этом в исследуемых текстах преобладает группа словосочетаний с опорным термином, выраженным именем существительным (88 %).

Результаты проведенного структурно-сопоставительного анализа показали, что лишь в 21 % случаев исследуемые ТС передаются линейно, без изменения структуры, например, *cordless optical mouse* – беспроводная оптическая мышь ( $A_1+A_2+N_1 \rightarrow A_1+A_2+N_1$ ), *hard drive* – жесткий диск ( $A+N_1 \rightarrow A+N_1$ ).

В значительном количестве переводных соответствий (79 %) используются переводческие трансформации, что обусловлено расхождениями в формальных и семантических структурах исходного языка и языка перевода. При этом такие традиционные преобразования, как опущение и добавление семантических компонентов, изменение порядка их следования и замена части речи достаточно часто применяются в различных своих комбинациях.

Наиболее распространенными преобразованиями при переводе английских составных терминов подязыка вычислительной техники являются перераспределение семантических компонентов (25 %), например, *remote storage notification service* – служба уведомлений внешнего хранилища ( $A+N_3+N_2+N_1 \rightarrow N_1+N_2+A+N_3$ ), расширение словосочетания наряду с перегруппировкой семантических компонентов (17 %), например, *application server* – сервер прикладных программ ( $N_2+N_1 \rightarrow N_1+A+N_2$ ) и замена части речи одного или несколько компонентов (15 %), например, *processor system bus* – системная шина процессора ( $N_1+N_2+N_1 \rightarrow A+N_1+N_2$ ).

Анализ отобранного материала свидетельствует о значительных трудностях перевода на русский язык английских ТС, с которыми сталкиваются не только специалисты, но и особенно системы автоматического перевода. Построенные структурные модели с указанием опорных компонентов послужили основой для разработки формальной модели автоматического выделения английских ТС из научно-технических текстов по вычислительной технике и их перевода на русский язык с помощью корпуса параллельных текстов [4].

Полученные результаты могут послужить исходным материалом для лексикографических и грамматических исследований, для составления учебных пособий, специальных словарей, терминологических банков данных, а также могут применяться в системах автоматической обработки естественного языка и для повышения эффективности перевода.

#### Литература

1. Першиков В.И. Русско-английский толковый словарь по информатике. – 3-е изд., перераб. – М., 1999.
2. Пройдаков Э. Англо-русский толковый словарь по вычислительной технике, Интернету и программированию. – 3-е изд., испр. и доп. М., 2002. 640 с.
3. Пустовайтова Л.А. Некоторые результаты исследования переводной неоднозначности английских существительных на материале научно-технических

параллельных текстов // Теория языка и инженерная лингвистика: сб. науч. работ. Л., 1973. – С. 103–112.

4. Степанова Д.В. Построение принципиального алгоритма автоматического выделения английских терминологических словосочетаний и их перевода на русский язык с помощью корпуса параллельных текстов // Актуальные проблемы прикладной лингвистики : в 2 ч. – Минск, 2008. – Ч. 2.

5. Универсальный словарь компьютерной терминологии. Англо-русский. Русско-английский. Таблица расширений имен файлов / Авт.-сост. Л.Н. Качахидзе. – 2-е изд., испр. и доп. М., 2005.