

РОЛЬ СЛОВАРНО-ПОИСКОВОГО ФАКТОРА ПРИ ПЕРЕВОДЕ АНГЛИЙСКИХ ТЕРМИНОВ ПО ВЫЧИСЛИТЕЛЬНОЙ ТЕХНИКЕ НА РУССКИЙ ЯЗЫК

В настоящее время перевод остается единственным эффективным способом преодоления трудностей на пути массового распространения информации в современном мире. Современное практически безграничное информационное пространство способствует ежегодному приросту и без того значительного объема переводов, при этом основную массу составляют переводы научно-технических текстов, центральным компонентом которых является термин [5, с. 2].

Количество специальных терминов и терминологических словосочетаний в каждом развитом языке по некоторым данным превосходит в несколько раз количество общеупотребительной лексики и измеряется семизначными числами [4, с. 45]. При этом, термины узких специальностей составляют 60-70 % от всего нового лексического состава языка, из них большая часть представлена терминологическими словосочетаниями [2, с. 13].

Значительную роль в процессе упорядочения и перевода терминологической лексики, несомненно, играют специальные словари, являющиеся результатом многолетней кропотливой работы профессионалов-языковедов и специалистов различных областей знаний.

Однако, поскольку проникновение терминологической лексики в общенациональные языки, обусловленное научно-техническим и социальным прогрессом, компьютеризацией международного общения, как правило, в значительной мере опережает процесс ее изучения и систематизации, даже самые современные словари не успевают отразить все изменения и новинки в составе терминологий различных областей.

В связи с этим в лингвистической литературе широко обсуждается возможность использования корпусов параллельных текстов в терминографии и переводоведении, поскольку они содержат переводы, при выполнении которых переводчиком-лингвистом уже проанализированы различные уровни контекста, установлено, в каких значениях употреблены как уже широко известные, так и новые термины, а также проведены структурный и семантический анализы терминологических словосочетаний.

До недавнего времени в лингвистике анализ данных при сопоставительных исследованиях выполнялся вручную, что обеспечивало невысокую скорость получения и обработки небольшого объема информации. Появление мощной компьютерной техники и обеспечение массового доступа к персональным компьютерам позволяет в настоящее время оперативно выполнять различные операции с электронными текстами.

Настоящая публикация посвящена вопросам научно-технического перевода в рамках подъязыка вычислительной техники, а именно описанию результатов сравнительно-сопоставительного исследования терминов в специальных словарях и в англо-русском корпусе параллельных текстов по вычислительной технике для раскрытия роли словарно-поискового фактора в процессе поиска английской терминологической лексики данной тематики и ее перевода на русский язык.

В качестве источника для выделения терминов и терминологических словосочетаний был разработан англо-русский корпус параллельных текстов, содержащий актуальную на сегодняшний день терминологию по вычислительной технике и состоящий из ста текстов — оригиналов на английском языке и их переводов на русский язык, выполненных профессиональными переводчиками.

Полученный корпус текстов был использован для различных видов анализа терминологической лексики.

На первом этапе с помощью статистических методов из английской и русской частей корпуса данной тематики были выделены 138 наиболее употребительных простых терминов английского языка и 138 наиболее употребительных простых терминов русского языка. Данные словоформы вследствие максимальных численных показателей могут отражать особенности текстов рассматриваемого подъязыка на лексическом уровне.

Дальнейший анализ предполагал проверку наличия выделенных наиболее употребительных терминов подъязыка вычислительной техники по специальным словарям той же тематики. Для выполнения этой задачи было решено использовать самые современные словари, однако, найти большие новые русско-английские словари по вычислительной технике оказалось достаточно затруднительно. Из предлагаемых изданий были выбраны три англо-русских словаря [7; 8; 10] и три русско-английских, среди которых два — толковые [6; 9; 11]. Проведенный статистический анализ показал, что 0,98 английских и 0,91 русских словоформ зафиксированы в выбранных словарях (*application, architecture, keyboard, program, protocol, компьютер, модем, мышь, процессор, устройство*), 0,013 английских и 0,057 русских опорных единиц не включены в словари (*deliver, automatically, вкладка, автоматически*), а 0,033 русских и 0,007 английских словоформ указаны только в составе словосочетаний (*файловый, операционный, системный, добавлять, cordless*).

Использованные для исследования словари послужили также источником для выбора и анализа зафиксированных в них терминологических словосочетаний, основным компонентом которых являются выделенные простые термины. Данные анализа показали, что основное количество терминологических словосочетаний приходится на двухкомпонентные сочетания, однако часто используются и трех-, четырех- и пятикомпонентные составные термины с учетом предлогов. С опорой на

данные словарей в исследовании было решено выделять словосочетания, максимальная длина структурной модели которых составляет пять компонентов.

На втором этапе с помощью компьютерной программы Concordance 3.2¹, которая позволяет автоматически выделить из текстов левую и правую контактную дистрибуцию ключевого термина по заданным параметрам, включая длину контекста, были выделены синтаксически правильные словосочетания с ключевыми терминами в качестве главных компонентов. Выделение словосочетаний осуществлялось в соответствии с двумя критериями: частотным фактором употребления и способами определения терминированности словосочетаний [1, с. 63,64]. В результате данного этапа были получены 582 английских и 767 русских терминологических словосочетаний, которые были распределены по ключевым терминам и по количеству компонентов, входящих в их состав. Проверка полученных словосочетаний осуществлялась по тем же специальным словарям, что и проверка терминов [6–11]. Результаты проверки показали значительное расхождение в количестве зафиксированных в словаре английских и русских терминологических словосочетаний ($f=0,45$ по сравнению с $f=0,21$), что, вероятно, объясняется небольшим объемом русских словарей и стремительным развитием области информационных технологий, которое отражается в языке, как правило, путем появления новых словосочетаний.

Подбор переводных эквивалентов для выделенных терминологических словосочетаний проводился с опорой на четыре источника: русская часть параллельного корпуса и использованные для исследования три англо-русских словаря по вычислительной технике [7; 8; 10]. В результате для 582 словосочетаний русского языка было обнаружено 905 терминов английского языка, 589 из которых ($f=0,65$) взяты из текстов и 316 ($f=0,35$) — из указанных словарей. Следует отметить, что из 589 текстовых русских эквивалентов большинство ($f=0,81$) были определены ранее на основании статистических методов как наиболее употребительные русские составные термины рассматриваемого подязыка.

Сравнительно-сопоставительный анализ эквивалентов, выделенных из русской части текстов, с эквивалентами, зафиксированными в словарях, позволил разделить английские терминологические словосочетания на три группы в зависимости от вариантов их переводов.

В первую группу включено большинство английских составных терминов ($f=0,55$), которые не зафиксированы в словарях. В результате их переводные эквиваленты были выделены только из исследуемых текстов. При этом некоторым английским словосочетаниям соответствуют два

¹ Конкорданс представляет собой тип словаря, в котором каждая словоформа текста характеризуется не только численными показателями (частотой, номером страницы и.т.д.), но и некоторым контекстом, в котором она употреблена [3, с.171]

синонимичных термина, например, *language features* — *средства языка*; *языковые средства*; *boot file* — *файл загрузки*; *загрузочный файл*.

Вторая группа содержит словосочетания ($f = 0,32$), текстовые переводные эквиваленты которых отражены в словарях. В зависимости от наличия или отсутствия других вариантов переводов английских терминологических словосочетаний, в данной группе было выделено две подгруппы.

1. Составные термины, которые имеют постоянные соответствия в русском языке ($f = 0,13$), например, *operating system* — *операционная система*; *network adapter* — *сетевой адаптер*; *subnet mask* — *маска подсети*. При этом большая часть текстовых переводных эквивалентов таких сочетаний ($f = 0,41$) отражена только в одном из исследуемых словарей (по сравнению с $f = 0,32$ в двух и $f = 0,27$ в трех словарях).

2. Составные термины, которым свойственны вариантные соответствия в русском языке ($f = 0,19$). При этом в 0,04 словарных статей наряду с синонимичными вариантами перевода были зафиксированы и случаи полисемии, например, *data file* — 1. *файл данных*, 2. *картотека данных*; *data access* — 1. *выборка данных*, 2. *доступ к данным*.

В третью группу были отнесены словосочетания ($f = 0,12$), 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*; *шинная топология*.

Проведенный анализ переводных эквивалентов свидетельствует о вариативности терминологии подязыка вычислительной техники и отсутствии общепринятых терминов, что значительно усложняет их перевод с английского языка на русский.

Данные сравнительно-сопоставительного анализа терминологической лексики в специальных словарях и в параллельных текстах показали, что, несмотря на высокий процент зафиксированных простых терминов, многие словосочетания остаются неучтенными. Этот факт дает основание рассматривать корпус параллельных текстов подязыка вычислительной техники в качестве источника получения новых терминов и их вариантов перевода, которые еще не зафиксированы в словарях, но уже широко употребляются в научно-технической литературе. Кроме того, параллельные тексты могут оказать помощь не только в установлении значений многозначных слов, но и в определении наиболее предпочтительных для данных текстов терминов из рядов синонимов.

ЛИТЕРАТУРА

1. Головин, Б.Н. О некоторых доказательствах терминированности словосочетаний / Б.Н. Головин // Лексика, терминология, стили: межвуз. науч. сб. / под ред. Б.Н. Головина (отв. ред.). — Горький, 1973. — Вып. 2. — С. 57–65.
2. Даниленко, В.П. Лингвистический аспект стандартизации терминологии / под ред. В.П. Даниленко. — М., 1993.
3. Зубова, И.И. Информационные технологии в лингвистике: учеб. пособие. — Минск, 2001.
4. Марчук, М.В. Развитие и становление значений слов основного терминологического слоя // Проблемы компьютерной лингвистики: сб. науч. статей / под ред. Р.Г. Пиотровского. — Минск, 1997. — С. 44–50.
5. Теория и практика перевода / под ред. Ю.Н. Марчука. — М., 2005. — №1. — С. 2.

СПИСОК СЛОВАРЕЙ

6. Дорот, В. Толковый словарь современной компьютерной лексики. — 2-е изд., перераб. и доп. — СПб., 2001.
7. Зубов, А.В. Словарь компьютерных терминов и вычислительной техники (англо-польско-русско-белорусский). — Белосток, 2003.
8. Масловский, Е.К. *Англо-русский словарь по вычислительным системам и информационным технологиям.* — М., 2003.
9. Першиков, В.И. Русско-английский толковый словарь по информатике. — 3-е изд., перераб. — М., 1999.
10. Пройдаков, Э. Англо-русский толковый словарь по вычислительной технике, Интернету и программированию. — 3-е изд., испр. и доп. — М., 2002.
11. 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.