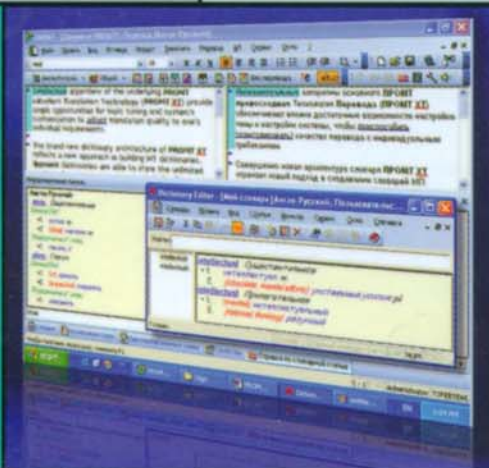


Высшее профессиональное образование

А. Л. Семенов

СОВРЕМЕННЫЕ ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ И ПЕРЕВОД

Учебное пособие



Иностранные
языки

А.А.СЕМЕНОВ

СОВРЕМЕННЫЕ ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ И ПЕРЕВОД

*Рекомендовано
Учебно-методическим объединением
по образованию в области лингвистики
в качестве учебного пособия для студентов
высших учебных заведений, обучающихся
по специальности «Перевод и переводоведение»
направления подготовки «Лингвистика
и межкультурная коммуникация»*



Москва
Издательский центр «Академия»
2008

УДК 519.768(075.8)

ББК 81.1я73

С302

Рецензенты:

военная кафедра Московского государственного института международных отношений (Университета) МИД РФ (начальник военной кафедры — *С. К. Васильев*); доктор филологических наук, профессор, директор Всероссийского центра переводов *И. И. Убин*

Семенов А. Л.

С302 Современные информационные технологии и перевод : учеб. пособие для студ. перевод. фак. высш. учеб. заведений / А. Л. Семенов. — М. : Издательский центр «Академия», 2008. — 224 с.

ISBN 978-5-7695-4459-0

В учебном пособии представлены основные теоретические положения переводческой деятельности как аналитико-синтетического процесса. Излагаются принципы компьютеризации текстовой деятельности. Рассматриваются теоретические основы практики использования современных информационных технологий письменного перевода. Наряду с лингвистическими проблемами затрагиваются наиболее актуальные организационные вопросы переводческой деятельности.

Для студентов старших курсов переводческих факультетов, слушателей системы повышения квалификации переводчиков, а также специалистов, заинтересованных в самостоятельном развитии навыков практического перевода.

УДК 519.768(075.8)

ББК 81.1я73

Учебное издание

Семенов Аркадий Львович

Современные информационные технологии и перевод

Учебное пособие

Редактор *Е. А. Кречетова*. Ответственный редактор *Г. Е. Конопля*
Технический редактор *Е. Ф. Коржуева*. Компьютерная верстка: *Н. А. Рогоза*
Корректоры *Л. В. Будюкина, О. Н. Яковлева*

Изд. № 101109585. Подписано в печать 20.12.07. Формат 60×90/16. Бумага офсетная № 1. Печать офсетная. Гарнитура «Таймс». Усл. печ. л. 14,0. Тираж 3000 экз. Заказ № 25769.

Издательский центр «Академия». www.academia-moscow.ru

Санитарно-эпидемиологическое заключение № 77.99.02.953.Д.004796.07.04 от 20.07.2004. 117342, Москва, ул. Бултерова, 17-Б, к. 360. Тел./факс: (495) 330-1092, 334-8337.

Диапозитивы предоставлены издательством.

Отпечатано в ОАО «Саратовский полиграфкомбинат». 410004, г. Саратов, ул. Чернышевского, 59. www.sarprk.ru

Оригинал-макет данного издания является собственностью Издательского центра «Академия», и его воспроизведение любым способом без согласия правообладателя запрещается

© Семенов А. Л., 2008

© Образовательно-издательский центр «Академия», 2008

ISBN 978-5-7695-4459-0 © Оформление. Издательский центр «Академия», 2008

ПРЕДИСЛОВИЕ

В последнее десятилетие в формах и стиле систем информационного обеспечения коммуникативной деятельности произошли кардинальные изменения, основной причиной которых стала компьютеризация текстовой деятельности. Именно так — «Компьютеризация текстовой деятельности» — называется новая учебная дисциплина, которая в настоящее время предлагается во многих европейских учебных заведениях различного уровня. Эта дисциплина имеет прикладной характер и часто совмещается с другими профессиональными дисциплинами. В связи с этим выпускаются оригинальные учебники, некоторые из которых приведены в списке литературы.

Необходимость подобного издания в России назрела уже давно. Естественно, его содержание должно соответствовать существующим условиям обмена информацией в делопроизводстве российского типа — с одной стороны, и программам подготовки переводчиков — с другой.

К сожалению, не только учебников, но и научной литературы по этой теме в нашей стране чрезвычайно мало. Изданные ранее книги в настоящее время недоступны из-за малых тиражей. В предлагаемом учебном пособии мы даем достаточно подробный обзор известных книг ведущих российских ученых в области машинного перевода (Ю. Н. Марчук) и автоматических переводных словарей (И. И. Убин). Мы с удовлетворением можем отметить, что российская школа новейших технологий в переводе продолжает развиваться, и в данном учебном пособии используем материалы, которые были подготовлены в ходе кандидатского диссертационного исследования П. Н. Хроменковым.

Исходя из наиболее распространенных учебных программ по курсу перевода, новое учебное пособие кроме обзорной главы включает четыре главы, соответствующие ведущим направлениям использования информационных технологий в переводческой деятельности. Все эти главы объединяет общая ориентированность на межъязыковую коммуникацию и ее реализацию в процессе письменного перевода.

В первой части рассмотрены основы современной информационной теории коммуникации и основные положения общей теории перевода в терминах аналитико-синтетического процесса.

Выбор стратегии перевода должен определяться целью коммуникации и техническими возможностями осуществления перевода. Следует отметить, что современные технические средства перевода пока еще остаются достаточно дорогими и не всегда доступны начинающим переводчикам. Поэтому мы старались максимально наглядно показать не только принцип их действия, но и методы использования в переводческой деятельности.

Во второй части мы последовательно рассмотрели работу с системой машинного перевода PROMT, автоматическим переводным словарем Lingvo и ТМ-инструментом TRADOS 5 Freelance.

Вторая глава этой части написана совместно с доктором филологических наук, профессором И. И. Убиным.

Небольшая заключительная глава посвящена эргономике труда переводчика в условиях использования новых информационных технологий.

В учебно-методическом плане первая часть пособия может быть рекомендована преподавателям для использования при подготовке краткого лекционного курса и практических или лабораторных занятий по различным спецкурсам, связанным с компьютеризацией переводческой деятельности.

Вторая часть больше рекомендуется студентам старших курсов для подготовки к практическим занятиям в компьютерных классах. Им было бы полезно найти собственные примеры оригинальных текстов и применить для их перевода различные функции программных средств, описанных в пособии.

Опыт подготовки таких учебных пособий невелик. Но за этим должны последовать еще много других. Чтобы они были интереснее и содержательнее, необходима критика первого опыта. Мы будем благодарны за любые конструктивные замечания и пожелания в адрес автора и издательства.

Автор выражает благодарность за советы и благожелательную критику доктору филологических наук, профессору И. И. Убину, а также доктору Хану Ван де Стааю (Академия перевода, Маастрихт, Нидерланды) за консультативную помощь в подготовке четвертой главы. Особая благодарность — кандидату технических наук П. А. Тимофееву за техническую поддержку.

ЧАСТЬ I

ТЕОРЕТИЧЕСКИЕ ОСНОВЫ ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ В ПЕРЕВОДЕ

Глава 1

ИСТОКИ НОВЫХ ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ В ПЕРЕВОДЕ

1.1. Первые опыты машинного перевода

Кропотливый труд переводчика, перенасыщенный множеством рутинных монотонных операций, в сочетании с творческой идеей переводческой деятельности, требующей интеллектуального подхода, всегда был в центре внимания революционеров науки и технологии.

Сразу же с появлением одной только идеи электронной вычислительной машины зародилась мечта о ее использовании в качестве электронного переводчика. Уже в 1946 г. эта мечта получила свое реальное воплощение в трудах американских ученых У. Уивера и А. Д. Бута. Они предлагали различные способы использования вычислительных машин для перевода текста с одного языка на другой. Очередным стимулом для развития научных исследований в области машинного перевода стала декларативная работа У. Уивера «Translation», опубликованная в 1949 г. В прогностическом плане ее автор обосновал возможность и целесообразность создания систем машинного перевода на базе ЭВМ¹.

Знаменательный 1946 год в истории науки и техники принято считать началом эпохи завоевания компьютером интеллектуальной сферы — текстовой деятельности человечества. К сожалению, при этом остался в тени тот факт, что еще в 1933 г. в СССР П. П. Троянскому было выдано авторское свидетельство № 40995 на «Машину для подбора и печатания слов при переводе с одного языка на другой или на несколько других одновременно». Однако этот факт вряд ли можно считать началом истории машинного перевода.

В 1954 г. был впервые продемонстрирован машинный перевод на практике — широко известный Джорджтаунский эксперимент.

¹ См.: Уивер У. Новая башня // Машинный перевод. — М., 1957.

7 января в конторе фирмы «ИВМ» в Нью-Йорке публично был осуществлен перевод с русского языка на английский при помощи выпущенной этой фирмой электронной счетной машины IBM-701. Для испытания был специально подготовлен словарь из 250 русских слов, записанных латинскими буквами. Слова были подобраны таким образом, чтобы каждое из них имело одно или, в крайнем случае, два английских значения. Помимо слов были включены также некоторые падежные окончания. В отдельных случаях давались только корни, а некоторые слова были представлены полностью даже с личными окончаниями (глаголы) или во множественном числе. Помимо английских значений в словаре для каждого слова указывались три специальных кодовых числа, используемых для управления работой машины. Для перевода были выбраны простейшие русские фразы, состоящие из слов, которые входили в словарь, и разработаны шесть синтаксических правил, обеспечивающих с помощью кодовых чисел правильный перевод.

Этот опыт вызвал огромный интерес во всем мире. Но большинство ученых считало, что речь может идти лишь о переводе научного и технического текста, причем и в этих, более простых условиях, перевод объемных отрывков, не говоря уж о целых статьях или книгах, расценивался как практически невозможный.

В СССР работы по автоматическому переводу начались в 1954 г. В конце 1955 г. на электронной вычислительной машине БЭСМ Академии наук СССР были проведены первые опыты автоматического перевода с английского языка на русский. Систему правил перевода (так называемый алгоритм перевода) разработала филолог И. К. Бельская. Программу составили математики С. Н. Разумовский, Н. П. Трифонов, Л. Н. Королев и другие. Это был первый в мире машинный перевод с английского языка на русский. Очень важно то, что он представлял собой перевод не отдельных фраз, а целых фрагментов текста научной книги.

Начальный этап развития машинного перевода приходится на период с 1946 по 1957 г. Это время характеризуется повышенным вниманием к машинному переводу, что было обусловлено экспериментами, проводимыми и в США, и в СССР. Особый интерес у разработчиков вызывало построение морфологического компонента и машинных словарей; уже сложилось определенное представление о необходимости подробных алгоритмов синтаксического и лексического анализа и синтеза для систем МП¹. Однако этот этап не завершился созданием промышленных систем МП.

¹ Для удобства изложения, но вместе с тем, чтобы не нарушать восприятие материала, мы решили использовать сокращение МП только в сочетании со словом «система». В случаях, когда речь идет о сущности машинного перевода и о процессе, сокращение не используется.

По мере углубления и расширения работ по машинному переводу становилась ясной огромная сложность поставленной задачи, появились некоторые сомнения в практической возможности получения высококачественного, полностью автоматического перевода.

Следует заметить, что работы по использованию ЭВМ в переводе начались с проектирования и создания систем МП, а не автоматических словарей, что казалось бы более естественным, поскольку словари проще в реализации, чем целостные системы перевода. Это объясняется несколькими причинами, основные из которых перечислены ниже.

1. Отсутствовало четкое представление о выделении в процессе традиционного перевода отдельных этапов. Не было даже общих оценок по определению «веса» и трудоемкости каждого этапа и соответственно трудозатрат человека-переводчика на решение его задач.

2. Существовала общая недооценка уровня сложности перевода как интеллектуальной деятельности.

3. Не были разработаны приемы формального описания морфологии, синтаксиса и семантики для целей машинного перевода.

4. Из-за отсутствия опыта создания крупных лексических массивов для систем МП не было понимания огромной трудоемкости этих задач. Так, У. Уивер в своих выкладках предполагал возможность создания в то время системы МП со словарем объемом 4 млн слов. Заметим, что самые крупные словари современных промышленных систем МП в настоящее время в среднем не превышают 250 тыс. словарных статей.

5. Не была разработана теория в узком (теория создания систем МП) и в широком смысле (теория применения ЭВМ на всех этапах перевода).

6. Технические возможности вычислительной техники в то время не соответствовали грандиозности поставленной задачи.

Все это привело к тому, что первые теоретики и практики машинного перевода, обосновав и доказав теоретическую и практическую возможность применения ЭВМ в переводе, существенно недооценили ни теоретические, ни практические трудности, преодоление которых необходимо для создания крупных промышленных систем МП, не обратили внимания на возможность иного (помимо подобных систем) пути автоматизации перевода, а именно — использования автоматических словарей разного типа.

После 1957 г. начинается новый период развития машинного перевода, связанный с первой реализацией основных идей. Но уже через десять лет наступило некоторое разочарование в практических результатах таких работ. Сам по себе этот период характеризовался бурным развитием исследований, сопровождавшихся и апробацией результатов. Проводились теоретические исследования по автоматическому анализу и синтезу на морфологиче-

ском и синтаксическом уровнях. Центр тяжести поисковых работ переместился с лексики на синтаксис. Возникла необходимость использования информации семантического уровня для получения перевода высокого качества, и, как следствие, были сформулированы основные идеи семантического анализа и синтеза текста перевода. Предпринимались попытки перехода к промышленной эксплуатации систем МП и создания для этой цели обширных словарей по разным тематикам. В течение нескольких лет, начиная еще с 1956 г., коллективом ученых под руководством Ю. А. Моторина разрабатывалась система МП, послужившая прототипом системы АМПАРа. Одновременно велись исследования по машинному переводу в Институте точной механики и вычислительной техники Академии наук СССР, которыми руководили Д. Ю. Панов (со стороны вычислительной техники и программирования) и И. К. Бельская (лингвист). Был создан алгоритм перевода с английского языка на русский. В теоретическом плане были достигнуты весьма обнадеживающие результаты в области структурно-прикладной лингвистики. Но практические достижения — собственно сам перевод — не могли быть признаны удовлетворительными из-за их низкого качества.

Этот (второй) период развития машинного перевода завершился опубликованием отчета Наблюдательного комитета по автоматической обработке текстов Национальной академии наук США по поводу экономической целесообразности машинного перевода и перспектив научных исследований в данной области. В докладе была показана почти полная его практическая непригодность без постредактирования. Исходя из языковой ситуации в СЛИА с учетом стоимости постредактирования, его авторы пришли к выводу об экономической невыгодности машинного перевода и вообще о недостижимости удовлетворительного результата в какие-либо реально планируемые сроки.

Пессимизм этого заключения не был принят в нашей стране, не поддержали его и некоторые ученые США и других стран. Однако отрицательное отношение авторов доклада к работе систем МП невольно отразилось на всех направлениях исследований в этой области и даже на самой идее использования ЭВМ при переводе.

Падением интереса к машинному переводу в конце 60-х гг. прошедшего столетия открывается третий — современный — этап развития компьютеризации переводческой деятельности. Всеобъемлющее усиление роли информации на разных языках и необходимость оперативного перевода большого количества текстов в условиях научно-технической революции заставили специалистов разных профилей вновь обратить свое внимание к возможностям машинного перевода. Но это обращение уже состоялось на принципиально новом уровне. Если первый период характеризовался

определением задач и путей развития машинного перевода, второй — разработкой теоретических проблем, созданием компонентов систем МП и переходом к разработке и внедрению их промышленных образцов, которые должны были заменить человека в осуществлении перевода, то современному этапу свойственно отсутствие противопоставления ЭВМ человеку. Системы МП по своим функциям и структуре становятся более гибкими и уже рассматриваются в одном ряду с другими способами и средствами автоматизации перевода. Это было обусловлено, с одной стороны, неизбежным осознанием информационных потребностей общества, не последней из которых является необходимость преодоления языковых барьеров, а с другой — более четким пониманием реальных возможностей промышленных систем МП.

Началу этого нового периода уже предшествовали и разработка и достаточно широко проведенная проверка теоретических основ машинного перевода. Это позволило сознательно избрать путь совершенствования технологичности самих систем и интерактивного включения человека-переводчика в некоторые технологические переходы между автоматическими процессами. Произошло естественное и удобное подчинение систем автоматического перевода качественным возможностям человека-переводчика.

1.2. Автоматические переводные словари

В середине 60-х гг. прошедшего столетия со всей очевидностью было признано, что полностью автоматический машинный перевод удовлетворительного качества в современном мире не может быть получен. Но уже имеющийся опыт разработки систем МП подсказывал, что многие рутинные операции в технологии перевода могут быть возложены на электронную вычислительную машину. Причем машина в состоянии выполнять эти операции не только быстрее чем человек, но и качественнее. Изменилось само отношение к электронной вычислительной машине: ее стали рассматривать как инструмент для автоматизации труда переводчика.

Однако полное признание такой подход к технологии перевода получил только к середине 70-х гг. К этому времени была четко осознана необходимость создания специальных автоматических устройств для повышения производительности труда переводчика и включения этих устройств в информационное обслуживание науки и техники.

Начало работ по созданию автоматических словарей в помощь человеку-переводчику знаменовало появление нового направления в стратегии автоматизации переводческой деятельности. Если системы МП своей основной функцией нацелены на замену человека-переводчика машиной, работающей либо в автоматиче-

ском режиме, либо совместно с человеком-редактором, то появление автоматических словарей означало переход к диалогу человека и ЭВМ, в котором последняя играла подчиненную роль. Этот переход был обусловлен целым комплексом социальных, лингвистических и технических факторов, которые сложились в середине 60-х гг. прошлого столетия. Первый среди них — быстрый рост объема научно-технической и деловой документации на разных языках и необходимость оперативного перевода значительной ее части; второй — невозможность выполнить эту работу традиционными способами, т. е. силами человека-переводчика в сжатые сроки; третий — неспособность созданных на то время систем МП обработать весь поток информации; четвертый — возросшие возможности вычислительной техники.

Влияние последнего фактора на развитие машинного перевода и вычислительной лексикографии, к сожалению, недооценивается некоторыми лингвистами, что совершенно неправомерно. Если при теоретических исследованиях лингвисты действительно редко учитывают реальные возможности вычислительной техники, то выбор оптимальных решений конкретных лингвистических задач в рамках промышленных систем МП и автоматических словарей в значительной степени зависит именно от уровня развития вычислительной техники. Это с особой наглядностью проявилось в середине 60-х гг., когда возросшие возможности ЭВМ позволили создать автоматические переводные словари объемом в десятки, а затем и в сотни тысяч лексических единиц, что было абсолютно невозможно в 50-х гг., когда только начинались работы по машинному переводу. Развитие вычислительной техники оказывает влияние на развитие вычислительной лексикографии и сейчас. Создание разнообразных, простых в обращении и дешевых компьютеров с достаточно большими возможностями привело к появлению нового направления в разработке автоматических переводных словарей в помощь человеку — словарей индивидуального пользования.

В настоящее время существует большой выбор компьютерных переводных словарей, выпускаемых на CD, которые по полноте содержащейся информации ни в чем не уступают своим традиционным аналогам (т. е. классическим бумажным словарям). Среди наиболее популярных можно выделить электронные словари семейства МультиЛекс компании «МедиаЛингва», систему электронных словарей Lingvo компании «АВВУУ», а также целый ряд электронных словарей зарубежных производителей. Их лексической базой зачастую служат авторитетные и признанные среди переводчиков традиционные бумажные словари (это характерно, например, для словарей семейства МультиЛекс), что позволяет им все более и более прочно внедряться в процесс современного перевода.

Все переводчики, за редким исключением, являются активными пользователями персонального компьютера, что само по себе располагает скорее к работе с компьютерным вариантом словаря, нежели с книгой.

Разработчики охватывают все более широкий спектр специальных областей перевода, что позволяет компьютерным словарям отвечать требованиям все большего числа пользователей. Свойство «обратимости» делает возможным поиск переводных эквивалентов в обратном направлении, что не только увеличивает ассортимент словарей в нашем пользовании, но и повышает качество перевода, так как в результате поиска мы получаем весьма большое число вариантов перевода с самыми различными оттенками значения в зависимости от контекста. Другими важными атрибутами компьютерных словарей, о которых нельзя не упомянуть, являются:

- совместимость с текстовым редактором Word и другими приложениями (в соответствии с конкретным продуктом) и запуск поиска непосредственно из этих приложений;

- возможность поиска слова в его текстовой форме;

- поиск словосочетаний и устойчивых выражений;

- удобство просмотра искомого слова или словосочетания во всем множестве контекстов (сбор информации из корпуса различных словарных статей);

- формирование истории запросов;

- пометка выборочных словарных статей для быстрого доступа к ним в последующей работе;

- пополнение содержания словарей в форме комментариев к словарным статьям;

- создание пользовательских словарей любого направления перевода и любой тематики по заданным шаблонам;

- возможность прослушивания произношения иностранных слов и устойчивых словосочетаний.

Таковым является базовый набор свойств компьютерных переводных словарей, которым отдают предпочтение сегодняшние пользователи. Хотя не все переводчики перешли на компьютеризованную работу со словарями, и традиционные бумажные словари еще сохраняют свою актуальность, тем не менее автоматизация поиска переводных эквивалентов существенно сокращает временные затраты на выполнение перевода. И. И. Убин считает, что использование автоматических переводных словарей может ускорить процесс перевода на 40 %¹.

¹ См.: Убин И. И. Современные средства автоматизации перевода: надежды, разочарования и реальность // Перевод в современном мире. — М., 2001. — С. 60—69.

1.3. ТМ-инструменты (Translation Memory)

В начале 80-х гг. прошедшего столетия руководитель японской национальной программы по машинному переводу Макото Нагао разработал новую концепцию перевода, основанного на примерах (Example based translation). Идея эта довольно проста. Известно, что в области научной коммуникации языки характеризуются однообразностью синтаксических конструкций и значительно терминологизированным лексическим составом. При замене лексики и терминологии часто повторяются одни и те же конструкции. Это наталкивает на мысль о том, что при накоплении достаточно большой коллекции ранее переведенных фраз велика вероятность того, что большая часть последующих текстов будет аналогична уже переведенным вручную. Для «обучения» системы должен быть сформирован большой массив исходных текстов и их переводов (такие пары называют билингвами), который должен быть введен в сверхмощную многопроцессорную ЭВМ. При переводе новых текстов из массива билингвов должны выбираться аналоги фрагментов этих текстов, которые могут быть использованы для формирования выходного текста.

Историю развития машинного перевода идея Макото Нагао в силу уже сложившегося отношения к этой проблеме, вообще, изменить никак не могла. Но на развитие современной технологии перевода с использованием компьютера она оказала, пожалуй, наибольшее влияние за последние 20—30 лет. Сейчас существует множество систем, которые так или иначе реализуют предложенную концепцию — это системы класса «память переводов» (Translation Memory) или ТМ-инструменты.

1.4. Информационные технологии и бизнес

Информация всегда играла значительную роль в обществе и, вообще, в человеческой деятельности на всех ее уровнях. Но как составляющая технологической деятельности информация стала необходимым элементом не так уж давно — может быть сто или немногим более лет назад. В те времена роль информации в технологии была столь невелика, что оценкой ее значимости никто не занимался.

Понятно, что определение роли того или иного параметра в технологической цепочке какого-либо производства становится важным только тогда, когда это хотя бы в малейшей степени оказывается связано с материальными и финансовыми затратами на производство и их окупаемостью. Как правило, об участии в технологическом процессе судят еще и по наличию и сложности оборудования, необходимого для выполнения той или иной операции.

Примерно 65—70 лет назад информационные технологии начали оказывать материально значимое влияние на деловую активность в бизнесе. С одной стороны, информационные технологии зачастую приносили существенную прибыль. С другой — они в свою очередь требовали заметных капитальных вложений. Именно в это время появился и конкретный инструмент, необходимый для реализации сложных информационных технологий, — компьютер.

По всем законам формальной логики можно считать, что история технологического подхода к информации насчитывает уже 70 лет. Но действительно значимую роль в бизнесе информационные технологии начали играть лишь в конце прошлого столетия — в середине 90-х гг. В это время многие исследователи пришли к выводу о том, что именно информационные технологии являются фактором резкого повышения эффективности экономик наиболее развитых стран.

Для того чтобы заставить информацию работать наравне с основным оборудованием и трудовыми ресурсами, потребовалось коренным образом изменить подход к пониманию роли технологии и информации.

По мере увеличения мощности и расширения сферы применения информационных технологий компании все чаще рассматривали их в качестве ключевого фактора успеха. Рост значимости информационных технологий наиболее наглядно отражается в структуре корпоративных расходов. Например, по данным Бюро экономического анализа Министерства торговли США, в 1965 г. на информационные технологии приходилось менее 5 % капитальных затрат американских компаний. В начале 80-х гг. после массового внедрения персональных компьютеров эта доля возросла до 15 %. К 90-м гг. она превысила 30 %, а в конце XX в. уже составляла более 50 %. Ежегодные затраты компаний во всем мире на аппаратные средства, на программное обеспечение и обслуживание информационных систем составляют около одного триллиона долларов США, а с учетом телекоммуникационных систем — более двух триллионов¹.

Переводческий бизнес не является исключением. Конкуренция и борьба за конкурентное преимущество в переводческом бизнесе так же важны, как и в любом другом. Отличительной особенностью здесь является только то, что резервы достижения конкурентного преимущества в этом специфическом бизнесе весьма ограничены. Если в других сферах в конкурентной борьбе определенную роль могут играть параметры сырьевой базы, специфика трудовых ресурсов, условия поставки готовой продукции, то в

¹ См.: Carr N. IT Does Matter? Information technology and the corrosion of competitive advantage. — Boston, 2004.

переводческом бизнесе практически отсутствуют такие факторы, как сортность готовой продукции, сроки ее реализации, значимость географии размещения трудовых ресурсов и т. п. Остается только, пожалуй, единственный резерв в борьбе за достижение конкурентного преимущества — использование новых специфических технологий переводческого процесса.

Как и в любом другом, в переводческом бизнесе основу информационных технологий составляет применение вычислительной техники и программного обеспечения. Как и любой другой, переводческий бизнес может много выиграть или сильно пострадать от компьютеризации его технологии. Задача каждого участника переводческого бизнеса — правильно оценить положение дел на своем участке и рационально использовать мощные информационные технологии (или определить необходимость их применения).

Нет смысла тратить пространство и время для создания впечатляющей картины современного процесса перевода информации с одного языка на другой. Пожалуй, уже не найдется людей, которым пришлось бы доказывать преимущества использования компьютера. Будем исходить из того, что без компьютера уже просто не возможна никакая текстовая деятельность. В сознании современного человека произошел необратимый реверс: в наши дни так же трудно представить себе рабочее место переводчика с пишущей машинкой, как 20 — 25 лет назад было трудно представить себе его рабочее место с персональным компьютером на столе.

Мы говорим о компьютеризации переводческой деятельности, имея в виду в основном, собственно, использование компьютера. В действительности, оставаясь главным звеном современного рабочего места, компьютер создает довольно сложную рабочую среду, компоненты которой имеет смысл рассмотреть подробно. Сама по себе замена пишущей машинки на компьютер даже в самой малой части огромной переводческой проблемы не означает какой-либо существенной новизны в процессе перевода и не имеет ничего общего с понятием информационной технологии в переводе.

Новые информационные технологии в переводе предполагают использование компьютера лишь как средства реализации сложнейшего специализированного программного обеспечения. Было бы трудно подвергнуть полной инвентаризации программное обеспечение лингвистических процессов. В настоящее время постоянно появляются все новые программы, обновляются или даже кардинально перерабатываются уже известные. Реально возможному обзору можно подвергнуть лишь основные принципиально необходимые компоненты программного обеспечения, составляющие отличительные признаки новейших информационных технологий в переводе.

К основным компонентам современной информационной технологии в переводческом бизнесе следует прежде всего отнести информационно-справочную базу (электронные словари, справочники, энциклопедии, нормативно-терминологические источники), средства накопления, архивирования, поиска и восстановления фрагментированных переводов (системы «памяти переводов»), средства подготовки, форматирования и конвертирования текстовых данных, средства локализации текстов переводов и др.

Контрольные вопросы и задания

1. Какие факторы и тенденции в развитии науки заложили основы для начала первого этапа в истории машинного перевода?
2. Какое событие называют Джорджтаунским экспериментом? Какой была основная цель этого эксперимента?
3. Объясните, что стало итогом второго этапа развития машинного перевода.
4. Дайте характеристику третьего этапа развития машинного перевода.
5. Что нового в идеологию автоматизации переводческой деятельности внесли автоматические переводные словари?
6. Проанализируйте функции вашего автоматического словаря и выпишите в блокнот его основные атрибуты.
7. Перечислите компоненты информационной технологии, значимые для переводческого бизнеса.
8. Составьте паспорт программно-лингвистического обеспечения своего автоматизированного рабочего места в виде двух столбцов. В левый запишите элементы рабочего места, которые у вас уже имеются, а в правый — элементы, на ваш взгляд недостающие.

ЛИНГВИСТИЧЕСКИЕ ОСНОВЫ ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ В ПЕРЕВОДЕ

2.1. Перевод как аналитико-синтетический процесс

Аналитико-синтетическую обработку информации необходимо понимать в широком смысле, т. е. распространяя ее практически на все сферы коммуникативной деятельности. Естественно, при этом мы должны исходить из определения языка как материального воплощения информации и рассматривать его как деятельность, связанную с материальными результатами. Достижение этих результатов может идти разнообразными способами, в частности способом перевода.

Исходя из представления о языке в лингвистике как о деятеле, большинство процессов получения информации и ее обмена можно описать в терминах аналитико-синтетической обработки. В этом смысле не требует дополнительных пояснений отношение к аналитико-синтетической обработке материализованной информации, под которой принято подразумевать классификацию и систематизацию, направленные на кодирование, библиографическое описание и аннотирование информации, обобщение, реферирование и перевод с одного языка на другой. Во всех этих случаях результатом аналитико-синтетической обработки информации является вторичная информация.

Но и порождение первичной информации в любом виде непосредственно связано с аналитико-синтетическим процессом. Если источником вторичной информации служит информация первичная, то источником последней является целенаправленная трудовая деятельность. С позиций понимания коммуникативной деятельности, в результате которой возникает первичная информация, текст оказывается первичным коммуникативным инструментом.

Первичная коммуникативная деятельность состоит из совокупности этапов, главные из которых — ориентирование в условиях достижения цели коммуникации, сличение полученного результата с намеченной целью и отображение коммуникативной деятельности, например, в виде текста. Продуктом первичной коммуникативной деятельности и ее предметно-знаковым продолжением является текст.

Перевод полностью относится к сфере речевой деятельности, сохраняя все ее специфические особенности, главная из которых обусловлена тем, что он осуществляется в среде двух языков.

С позиций аналитико-синтетической обработки информации может быть принято определение перевода, разработанное А. Ф. Ширяевым: «Перевод есть вид специализированной речевой деятельности, опирающийся на использование двух языков в целях воссоздания на одном языке речевых высказываний, эквивалентных по форме и содержанию речевым высказываниям на другом языке»¹.

Аналитико-синтетический процесс складывается из трех основных этапов: анализа информации, ее переработки и синтеза вторичной информации. Вторичная информация — понятие относительное и специфичное. В некоторых случаях, когда происходит материальное отторжение вторичной информации, она может рассматриваться и как первичная. С позиций системного подхода в какой-либо конкретной системе это связано с переходом информации с одного уровня на другой.

Этот процесс хорошо понятен в отношении перевода. Перевод как документ в любом виде в системе обмена информацией в конкретной предметной области может быть признан первичной информацией. С позиций аналитико-синтетического процесса специфика перевода заключается еще и в особом соотношении предметов анализа и синтеза. Основные характеристики такого соотношения — эквивалентность и адекватность, поскольку перевод выполняется в целях воссоздания на одном языке речевых высказываний, эквивалентных по форме и содержанию речевым высказываниям на другом языке. И по форме, и по содержанию перевод (вторичная информация) призван адекватно замещать источник (первичную информацию).

Традиционно до последнего десятилетия мировой лингвистикой рассматривались три жанра письменного перевода: переводы Библейских текстов, литературный перевод и специальный перевод. В настоящее время последний, третий, жанр принято подразделять на множество предметных направлений — политический, консультативный, юридический, технический, медицинский, военный перевод и др.

Все эти основные направления (возможно, история выделит еще и другие, формируемые запросами и интересами общества) укладываются в три сферы теории перевода:

- буквальный перевод (духовной литературы);
- свободный перевод (художественной литературы);
- прагматический трансфер (тематических материалов для информационных процессов).

Многие десятилетия научный подход к теории перевода стремится идти по пути его компьютеризации. И уже внедрены существенные достижения. Но от чисто математического подхода

¹ Ширяев А. Ф. Картина речевых процессов и перевод // Перевод как лингвистическая проблема. — М., 1982. — С. 12.

60-х годов теория постепенно возвращается в лингвосемиотический фрейм: по сей день основными остаются две классические области — семантика и дискурс.

Эти области лингвистики имеют большое значение не только для теории, но и для практики перевода. Понимание текста как коммуникативного дискурса стало исходным условием многих плодотворных теорий перевода. Коммуникация реализуется посредством текста. Невозможно адекватно перевести изолированные слова или даже предложения, пока они не станут частью полного дискурса, который обычно входит в конкретный контекст ситуации. Кроме того, важно отметить, что перевод дискурса возможен только в том случае, если известны эквивалентные структуры в том языке, на который осуществляется перевод. А это знание складывается из сравнения, из сопоставительной лингвистики или, точнее, из контрастивной текстологии. Чтобы перевести отдельные предложения какого-либо сообщения, необходимо знать не только то, имеются ли в передающем языке и в языке перевода соотносящиеся лексические и грамматические единицы, но и то, какие стилистические условия приняты в тексте сообщения. Это положение приобретает серьезное сдерживающее значение для автоматизации и компьютеризации процессов перевода. Если для правильного понимания предложения перед его переводом надо учитывать весь окружающий текст, то естественно, система МП (или автоматический переводчик) должна каким-то образом сверять перевод каждого предложения со всем переведенным текстом. Но это действие требует достаточно развитого интеллекта, которым еще не обладает ни одна компьютерная система.

2.2. Модели процесса перевода

Одним из самых устойчивых затруднений является то, что в процессе перевода нет прямого доступа к действиям переводчика. Все исследования основываются на косвенных фактах в виде продукта или результата процесса. Перевод — это процесс и продукт, предмет и цель исследования, творчество и его результат.

Самым привлекательным направлением в теории перевода является выход за пределы сравнений различных версий перевода и лингвистических систем в область понимания того, как перевод действует в сумме всех коммуникативных интеракций, как может происходить коммуникация, когда включаются различные коды и что делает переводчик для ее поддержания.

Такая коммуникативная или социолингвистическая точка зрения на перевод достаточно нова. Она позволяет выполнять исследования завершеного текста, а не его отдельных избранных исследователем элементов. Предшествовавшие ей взгляды на про-

цесс перевода (в различные периоды долгой истории перевода его трактовали и как «метаморфозы», и как «замены» и «интерпретации») не могли учесть бóльшую часть переменных в этом сложном процессе. Только модели перевода, рассматривающие этот процесс в фрейме межъязыковой коммуникации, в состоянии учитывать максимальное количество переменных. Но следует иметь в виду, что перевод — это особый вид коммуникации, отличающийся прежде всего тем, что в него включены по меньшей мере два кода: два языка или два диалекта. Модели этого процесса должны отражать взаимные соотношения между различными коммуникационными системами.

2.3. Основные теоретические модели

Рассматривая аналитико-синтетический процесс безотносительно к его объектам и результатам и учитывая наличие в нем трех основных этапов, представляет интерес сопоставление этих этапов с этапами выполнения перевода с одного языка на другой. Касаясь, вообще, перевода как одного из видов двуязычной коммуникации, О. Каде выделяет следующие этапы в модели перевода¹:

- декодирование текста на языке оригинала с целью перекодирования;
- перекодирование, т.е. подстановка знаков языка перевода вместо знаков языка оригинала;
- реализация текста на языке перевода.

В устном переводе (синхронном) А. Ф. Ширяев также видит три этапа²:

- ориентирование в исходном тексте;
- поиск и принятие переводческих решений;
- осуществление переводческих действий.

Также выделяя в самом общем виде три этапа (анализ текста оригинала; замена кода языка оригинала на код языка перевода; синтез текста перевода), И. И. Убин проводит параллель между тремя фазами традиционного перевода и тремя этапами работы систем МП: анализом, трансфером, синтезом³.

В. Н. Комиссаров считал, что «реальный процесс перевода развивается, по-видимому, линейно, и самостоятельными этапами в этом процессе оказываются лишь комплексы действий, необходимых для перевода отдельного отрезка текста... Однако в преде-

¹ Каде О. Проблемы перевода в свете теории коммуникации // Проблемы теории перевода в зарубежной лингвистике. — М., 1978. — С. 73.

² Ширяев А.Ф. Картина речевых процессов и перевод // Перевод как лингвистическая проблема. — М., 1982. — С. 9.

³ Убин И.И. Лингвистические основы создания автоматического переводного словаря: автореф. дис. ... д-ра филол. наук. — М., 1989. — С. 7.

лах каждого такого комплекса можно различать два вида действий, составляющие хотя и несамостоятельные, но достаточно ясно выделяемые этапы, которые могут, с известными оговорками, описываться реально. К первому такому этапу будут относиться действия переводчика, связанные с “извлечением” информации из оригинала. Ко второму — вся процедура выбора необходимых средств в языке перевода»¹.

Ситуационная модель перевода. Эта модель описывает процесс перевода в терминах эквивалентности. Основной ее особенностью является то, что она учитывает условные стили в ассоциации с различными коммуникативными ситуациями, т.е. при переводе текста необходимо найти ситуационно эквивалентный параметр в другом языке. Это, может быть, несколько сближает данную модель с методами компаративной лингвистики. В дальнейшем процесс развивается через ситуационную эквивалентность и по своей методике близок к стилистической адаптации.

Большая часть политических переводов выполняется по этой модели, так как политический дискурс маркирован четче других, и его обычно ожидает больше совпадений в других языках.

Некоторые лингвисты (Ю. Найда и другие) представляют аналогичную модель в терминах грамматики предложения, и она эксплицирует лингвистические процедуры, которые могут быть востребованы для достижения «динамической эквивалентности» между передающим языком и языком перевода сообщения. Его модель имеет вид, представленный на рис. 1.

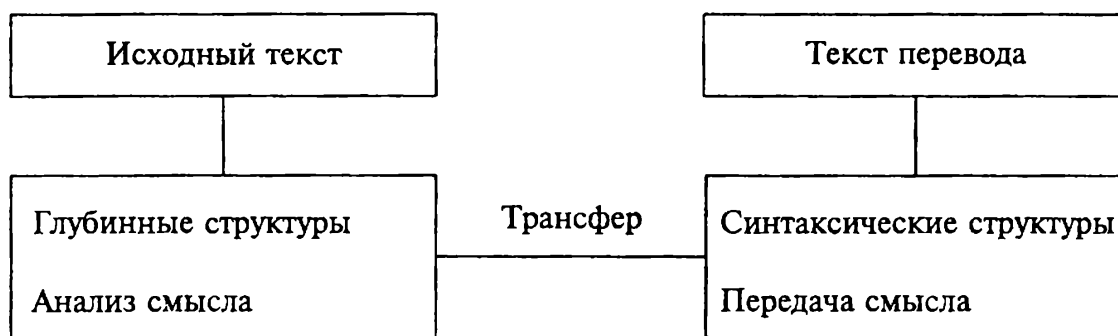


Рис. 1

Эквивалентность рассматривается здесь не как случайные совпадения между ситуационно аналогичными текстами, а как результат управляемого когнитивного процесса, который может быть доказан лингвистической теорией. Эта мысль на более современном уровне раскрыта В. Н. Комиссаровым². Если было бы возмож-

¹ Комиссаров В. Н. Слово о переводе. — М., 1973. — С. 162.

² См.: Комиссаров В. Н. Когнитивные аспекты перевода // Перевод и лингвистика текста. — М., 1994. — С. 7—22.

но свести синтаксические и семантические комплексы текста или отдельной синтаксемы к серии ядер и компонентов, появилась бы возможность свободно оказывать влияние на трансфер между кодами. Такой подход к проблеме ближе других стоит к психолингвистике. Реализовать его в каких-либо формализованных компьютерных системах в настоящее время пока невозможно. Мы еще раз убеждаемся в том, что процесс перевода развивается через глубинные структуры и включает в себя трансформационный анализ и синтез, что доступно только развитому интеллекту.

Многоуровневая модель перевода. Среди известных моделей перевода привлекает к себе внимание идея многоуровневой иерархической структуры, которую многие лингвисты рассматривают как релевантную процессу понимания перевода. Модель Кэтфорда-Холлидея выделяет мысль о том, что эквивалентность перевода зависит от возможности формального совпадения между лингвистическими единицами на различных структурных уровнях. Подобная модель имеет следующий обобщенный вид (рис. 2):

Контекст (передающий текст)		Ситуация (язык перевода)
Грамматические	Транскоди- рование	единицы
Лексические		единицы
Графические		единицы
Фонетические		единицы

Рис. 2

Процесс перевода в данном случае развивается через формальные соответствия, но способы достижения эквивалентности эта модель не раскрывает.

Коммуникативная модель перевода. Эта модель, предложенная Гертом Ягером, включает в себя реалии дискурса и его анализ в терминах лингвистики и семиотики. Эта модель схематически показана на рис. 3.

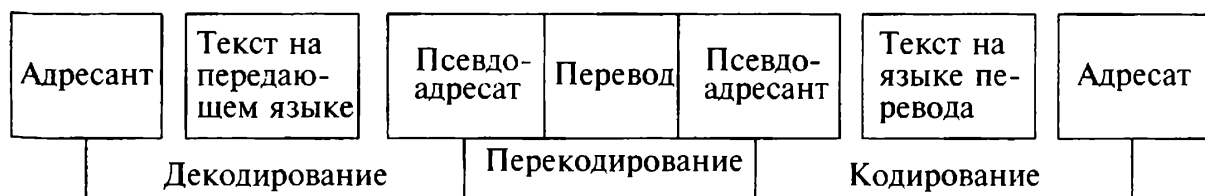


Рис. 3

Первостепенно важным компонентом в межъязыковом посредничестве здесь признается переводчик и его билингвальная компетенция. Он — единственный интерфейс в этой коммуникации, работающий в режиме переключения. Процесс перевода развивается через способность переводчика к декодированию и кодированию, которой он достигает различными методами обучения письменному и устному переводу.

Но и эта модель представляется слишком жесткой, чтобы воспринимать комплексность перевода как набор интегрированных процессов. Объектом анализа уже оказываются самые очевидные воздействующие факторы: адресант, адресат, код. Но такие факторы, как объект/событие/сигнификат, канал связи/среда и контекст/ситуация вовсе не учитываются или учитываются поверхностно. Это довольно серьезный недостаток, так как синтаксических аспектов текста совершенно недостаточно для описания дискурса без участия аспектов референционной семантики и коммуникативной прагматики текста. С позиций современной лингвистики такая модель должна быть дополнена аспектами контрастивной текстологии, без потери каких-либо уже достигнутых преимуществ.

Комплексная модель перевода. В этой модели обусловлены все факторы, которые должны быть учтены для того, чтобы концептуальный фрейм мог квалифицированно представлять лингвистическую теорию перевода. Однако из нее не следует явно, насколько



Рис. 4

учтены психолингвистический трансфер (Найда) или семиотическое перекодирование (Ягер). Но если вернуться к определению билингвального контакта, то становится понятным место этих лингвистических параметров в рассматриваемой модели.

Соответствия, которые могут наблюдаться по завершении этих процессов между лингвистическими единицами, встречающимися в параллельных текстах, являются результатом сознательного или подсознательного выравнивания (переключение кодов), если переводчик счел его необходимым, средствами ссылок на внутренние факторы дискурса и внешние объекты/события, а не формальными корреляциями, обусловленными взаимным соответствием двух языковых систем. Схематически эту модель отражает рисунок 4.

Такие соответствия в обобщенном виде всегда должны быть подчинены определенному типу дискурса, который может быть избран в сложившейся ситуации.

* * *

Итак, перевод является открытой гибкой вариативной процедурой, и критерии оценки его эффективности должны включать степень практической реализации билингвальной компетенции, знание внутриязыковых условий переключения и осведомленность в предметной области, к которой относится предмет коммуникации.

Большинство моделей до сих пор учитывали роль переводчика в процессе перевода как некий «черный ящик» или как «декадирующий/кодирующий аппарат перевода». В действительности переводчик на основе своего интеллектуального потенциала, используя всю билингвальную компетенцию, принимает участие как в ситуационном контексте адресанта, так и в ситуационном контексте адресата, как можно больше аппроксимируя и минимально (но согласуясь с предметом текста) используя в любой конкретной ситуации формальные и стилистические условности каждой языковой системы.

По этой модели процесс перевода развивается от одного дискурса к другому через коммуникативную компетенцию билингва, который не всегда должен быть (и не всегда бывает) профессиональным устным или письменным переводчиком. При этом происходит поиск эквивалентности на всех уровнях от синтагматико-грамматического до парадигмо-семантического, но, что чрезвычайно важно отметить, весь процесс осуществляется с учетом законов прагматико-стилистической целесообразности. Способность к аппроксимации посредника содержит в себе аналитический компонент по отношению к тексту на передающем языке и его адресанту и синтетический компонент по отношению к тексту на языке перевода и его адресату.

2.4. Типы переводов

Обзор моделей процесса перевода наглядно обнаруживает условия, которые необходимы для того, чтобы перевод состоялся. Любая инвентаризация выявленных условий могла бы быть полезной только для описания разнообразных межъязыковых ситуаций, которые требуют участия билингвального посредника.

Традиционное разделение переводов по типу жанров: библейский, литературный, технический (политический, военный) — можно справедливо критиковать за чрезмерную ориентированность на стилистические критерии. По своему участию в коммуникации перевод должен рассматриваться с позиций факторов, определяющих успех или неуспех в коммуникации. Проблемы и практические аспекты перевода изменяются, например, в диапазоне видов взаимодействия между адресантом и адресатом. Они могут зависеть от занимаемых социальных положений (автор — читатель, бизнесмен — клиент, профессор — студент) — во всех этих случаях цель перевода будет сильно различаться и самые большие отличия будут связаны с прагматическим компонентом.

Другим обуславливающим фактором является тип ситуации. Прагматический контекст вносит большие различия в формы речи, например в зависимости от национального и политического климата, образовательного уровня, традиций билингвальных контактов внутри языка и в отношениях с другими языками.

И конечно, одним из особо значимых факторов оказывается сам переводчик, который может быть профессионалом или непрофессионалом, компетентным или неопытным, оплачиваемым специалистом или случайным помощником, штатным или штатным сотрудником и др. В любом случае процедуры, которые он выполняет для совершения своей работы, не могут быть унифицированы. Они могут варьироваться в широких пределах от буквального перевода до резюмирования, от калькирования до полного переосмысления сообщения. В такой ситуации возникает проблема: если процедуры перевода не могут быть унифицированы, не может быть и каких-либо устойчивых критериев оценки его адекватности, основанных на анализе процедур. Качество перевода оценивается по ситуациям адресанта и адресата, по типам текстов и билингвальной компетенции устного или письменного переводчика.

Единственной реально выполнимой методикой обеспечения стабильного качества перевода является подбор переводчика по его билингвальной компетенции для выполнения конкретной работы и последующее контрольное редактирование. Качество перевода находится в прямой зависимости от оценки текста оригинала, билингвальной компетенции переводчика и требований (ситуации) адресата перевода.

Те же принципы лежат и в основе переводческого сервиса, предлагаемого Бюро переводов Комиссии Европейского Сообщества. Его технологию отличает то большее значение, которое придается оценке требований заказчика (адресата), что заметно усиливает роль прагматического компонента. В Бюро переводов Комиссии ЕС различаются пять категорий перевода:

- «сырой» устный перевод (rough oral) в целях извлечения исходной информации (часто выполняется штатным переводчиком в присутствии заказчика);
- «содержательный» перевод (abstracting the content) для получения полной информации о содержании сообщения на передающем языке;
- «сырой» письменный перевод (rough written) в служебных целях для последующей аналитической обработки;
- «чистой» письменный перевод (careful written) для документального использования и длительного хранения;
- перевод для публикации (for publication) — аутентичный перевод с учетом многоязычного делопроизводства этой международной организации.

Все это доказывает жизненную стойкость типологии перевода и переводческого труда в отрыве от грамматического и лексического анализа текста (хорошо известный «грамматический перевод»), но со строгим контрастивным текстологическим анализом.

Если исходить из того, что перевод является межъязыковой текстовой аппроксимацией, требующей определенной билингвальной компетенции, то следует указать, каким образом могут быть достигнуты средства переключения. Без знания грамматического перевода, подкрепленного практикой, не может быть достигнута билингвальная компетенция необходимой профессиональной степени.

Опыт Всероссийского центра переводов указывает на то, что большая часть переводов выполняется специалистами в предметной области, к которой относится текст оригинала, имеющими прикладные знания иностранного языка. Профессиональные переводчики, получившие грамматические знания, не становятся автоматически переводчиками-практиками. Но в этом противопоставлении нет противоречия по той причине, что дискурсный анализ предполагает грамматическую корректность. Выше уже шла речь о совпадениях дискурсов в профессиональных областях. В наших примерах это в полной мере подразумевается в параметре «ситуация адресанта/адресата». Знание ситуации обеспечивает необходимый грамматический фон для билингва в одном из его языков, а некоторая степень единства дискурса в состоянии обеспечить достижение требуемого качества в переводческих процедурах.

Вместе с тем перевод, как межъязыковая аппроксимация на уровне текстов, требует координированного использования раз-

личных знаний и навыков за пределами владений по меньшей мере двумя языками. Многие современные лингвисты в инвентаризацию составляющих билингвальной компетенции вводят еще один компонент — оценку (самооценку) широты кругозора, которая в данном случае может и должна быть описана в терминах социолингвистики и лингвострановедения.

2.5. Коммуникативные особенности аналитико-синтетического процесса

Аналитико-синтетическую обработку информации мы рассматриваем в широком смысле, т.е. распространяя ее практически на все сферы коммуникативной деятельности. Естественно, мы придерживаемся понимания языка как материального воплощения информации и определяем его как деятельность, связанную с материальными результатами, достижение которых может происходить самыми разнообразными путями. Говоря словами В.Гумбольдта, «язык есть не продукт деятельности людей (Ergon), а деятельность (Energia)»¹.

Технологический подход к информационной деятельности может быть построен на деятельностном представлении языка в лингвистике. Большинство процессов получения информации и обмена ею можно рассматривать как технологические операции аналитико-синтетической обработки текстов независимо от языка. Перевод с одного языка на другой — это тоже в общем смысле аналитико-синтетическая обработка информации.

Не требует в этом смысле дополнительных пояснений отношение к аналитико-синтетической обработке материализованной информации, под которой принято понимать классификацию и систематизацию (связанные с кодированием, библиографическим описанием и аннотированием информации), обобщение, реферирование и перевод с одного языка на другой. Во всех этих случаях результатом аналитико-синтетической обработки информации становится вторичная информация.

Тексты сообщений существуют и функционируют во взаимодействии между собой и с окружающим миром. Взаимодействие с окружающей средой — это имманентное условие реализации системы, а потому оно является одним из основных системных признаков. Постигание механизмов коммуникативного функционирования текстов достигается только при соотношении с прагматически определенной средой. И ни о каком функционировании вообще речи идти не может, ибо всегда необходимо ясно представлять среду, в которой оно осуществляется.

¹ Гумбольдт В. Избр. труды по языкознанию. — М., 1984. — С. 72—73.

2.6. Лингвистическое понимание среды и системы в коммуникативном процессе

Исследования среды, в которой функционирует текст, представляет интерес не ради оценки окружающей ситуации, а для понимания прагматических условий порождения и восприятия самого текста. Здесь особенно важна та закономерность, которая состоит в том, что связи определенного вида между элементами множества, образующего систему (внутренние связи), преобладают над связями между этими же элементами и окружающей средой (внешние связи).

Влияние фактора «окрестных условий» подробно исследовано Г. П. Мельниковым на примере предметной области кибернетики¹. Важное теоретическое значение имеет его тезис о малоизученном аспекте взаимоотношений системы и среды на переходе «действительность — возможность». В аспекте такого перехода для текста сообщения необходимо наличие другого объекта — среды. Только в этом случае можно будет экспериментально или теоретически указать валентности и системы, и среды, которыми они будут обладать, если при заданных внешних условиях окажутся во взаимодействии.

Что же составляет среду функционирования текстовой информации (сообщения)? Ее можно охарактеризовать как сложноструктурную системную среду, состоящую из ряда отдельных подсистем, каждая из которых имеет независимые или малозависимые между собой связи с текстом. Поскольку отношения между системой и средой формируются под воздействием системности реального мира, они приобретают функциональный характер, т. е. направлены на выполнение какой-либо функции. В конкретном случае функционирования текста в определенной предметной области требуется реализация функции обеспечения коммуникации в общепринятых терминах и с аспектами, заданными прагматическими параметрами. Это создает необходимость взаимодействия текста с объективно существующей системой понятий в данной предметной области, но, кроме того, еще с системами (национальной, международной, государственной, военной, технической и др.) политики и социальных отношений и, наконец, с общей лексической и грамматической системой языка. Все эти составляющие окружающего мира и образуют среду функционирования текстов сообщений.

Дихотомия «система — среда» приобретает смысл только при наличии валентных связей, которыми располагают и текст, и среда. По их направлениям и благодаря их участию осуществляются

¹ См.: Мельников Г. П. Системология и языковые аспекты кибернетики. — М., 1978. — С. 49.

обменные процессы, суть которых на самом низком уровне заключается: со стороны текста — в лингвистическом обеспечении коммуникативного процесса, а со стороны среды — в формировании потенциального энергетического воздействия текста как инициатора коммуникативных реакций.

По характеру обменных процессов можно судить о динамическом состоянии коммуникативного взаимодействия и об уровне развития коммуникативных отношений в среде предметной области. Если по отношению к тексту сообщения обменные процессы строятся на входящих потоках, то можно говорить о развитых коммуникативных отношениях и активной текстовой деятельности в данной предметной области. Если потоки уравновешены, то скорее всего такая текстовая деятельность соотнесена с неразвивающейся предметной областью.

Наблюдения за текстовой деятельностью в среде конкретной предметной области дают возможность утверждать, что возможно и третье состояние, когда обменные процессы между текстовой деятельностью и средой функционирования сообщений строятся на исходящих потоках по отношению к текстам. В этом случае текстовая деятельность приобретает продуцирующую функцию, инициированную какими-то косвенными экстралингвистическими факторами.

Это глубоко специфический аспект функционирования текста сообщения, поскольку отражает характер его связей с системой лингвистических и экстралингвистических отношений в предметной области и смежных предметных областях.

Во-первых, в условиях сильной интерференции высокотехнологичных предметных областей (представляющих, например, усложненные технологические процессы) происходит не простое взаимопроникновение терминологии на понятийном и лексическом уровнях, а ее взаимообусловленность. Примеры можно легко найти и в социальной сфере, но наиболее наглядным и простым является построение терминов робототехники по моделям тех предметных областей, где робототехника находит применение.

В английском языке специфический термин предметной области микроэлектроники *clean-room electronics* (интегральные схемы, изготавливаемые в особо чистых условиях) вызвал появление термина *clean-room robotics* (роботы для работы в особо чистых условиях). У этих двух терминов совпадает только синтаксическая конструкция. Объем понятий сильно различается: в первом случае понятие касается объекта, во втором — субъекта, но окружающая среда — коммуникативные процессы в предметных областях, связанных с применением робототехники для производства интегральных схем, — у этих терминов одинаковая и сходные подчинительные связи.

В данном случае коммуникативный процесс затрагивает две лексические подсистемы, и если в одной из них имеется устойчивая лексическая единица, а в другой ее нет, но в системе понятий все же существует некое аналогичное материальное содержание, то среда функционирования текста может становиться регламентирующим фактором, диктующим формы коммуникативного процесса. Под воздействием опыта коммуникативной текстовой деятельности в предметной области микроэлектроники и правила, выведенного средой функционирования текстов сообщений, влияние лексической подсистемы робототехники оказалось активнее влияния системы понятий предметной области. То есть коммуникативные связи между лексическими подсистемами, осуществляемые в единой коммуникативной среде, обобщили модель построения лексической единицы текста как коммуникативного средства.

Во-вторых, глубокое проникновение, например, робототехники, в сферу производственных отношений создает условия для построения лексических единиц по моделям, соотнесенным с понятиями служебной иерархии: *blue collar robot* (рабочий робот) и *white collar robot* (робот-координатор).

Описанные наблюдения и примеры характеризуют окружающую среду как совокупность обобщающих каналов связи, по которым осуществляются обменные процессы. Обобщающее воздействие каналов связи, поддерживаемых средой, наполняет обменные процессы опытом реализации качества, накопленного в различных сферах среды и элементах системы.

И. В. Арнольд предложила семиотический подход к исследованию взаимодействия языковых систем и среды¹. Нам представляется возможным распространить сделанные ею выводы на соотношение сил влияния текста и среды, рассматривая текст как четко оформленную подсистему.

Если ограничиться узким прикладным пониманием семиотического членения на синтактику, семантику и прагматику, мы можем наблюдать три всеобъемлющих аспекта взаимодействия текста сообщения со средой. Принято считать, что синтаксический аспект раскрывает только внутренние связи. С учетом вероятности многоуровневой структуры плана содержания текста (например, правила поведения в метро кроме прагматического инструктивного уровня могут еще быть использованы как информация о развитии городского транспорта или как информация о действующих административных отношениях в сообществе) и возможности самостоятельного функционирования одного текста на разных информационных уровнях смысл синтаксических связей

¹ Арнольд И. В. Основы научных исследований в лингвистике. — М., 1991. — С. 125.

значительно расширяется из-за комплексности и структурной сложности составляющих их информативных элементов.

Без учета окружающих условий процессы формализации текста сообщения в коммуникативном процессе, конечно, можно было бы объяснить внутрисистемными отношениями, т. е. на уровне синтактики. Понимание внутренних связей безусловно дает возможность описать принципы функционирования текста сообщения. Но в этом случае описание было бы статичным и представляло бы более или менее идеализированную формализованную модель, которая в состоянии отобразить принципы, но не результаты функционирования текста как коммуникативной единицы или даже его формальных подсистем, что лишает исследование практической ценности.

В новейшей истории текстовой деятельности можно найти примеры осознанного и объективного стремления к формализации текстовой деятельности. Самым ярким подобным примером может быть компьютеризация текстовой деятельности, связанная с высочайшей степенью формализации: автоматическая подготовка писем, ведение корреспонденции в среде электронной почты, автоматическая подготовка спецификаций, руководств и др. И в литературном детективном жанре бурно развивается компьютерная индустрия компиляции детективных историй.

Однако с учетом среды появляется возможность продуктивного анализа семантических связей текста сообщения с системой понятий предметной области и прагматических связей, а также с участниками коммуникации и с той же системой понятий.

Кроме общего представления о семантических связях как о соотношении лексического состава языка и предметного мира, при системном подходе к пониманию принципов и результатов функционирования текста в соотнесенной с ним коммуникативной среде важен учет влияния диалектики количественно-качественных превращений. Накопление знаний в предметной области активизирует коммуникативный процесс. С расширением и углублением коммуникативной компетенции, выражающейся в активизации текстовой деятельности, происходит перераспределение внутренних связей в виде развития способов текстовой реализации плана выражения. Качественные изменения, усиливающиеся в текстовых сообщениях по мере развития коммуникативной компетенции, отражаются на изменении их связей со средой.

Перераспределение внутренних связей под воздействием количественных изменений носит диахронический характер, соотносимый с прогрессом предметной области. Но диахрония количественных изменений имеет более стереоскопическую суть нежели прогресс предметной области, поскольку классификация на номинативной основе всегда вещественнее, чем на понятийном уровне. Вместе с тем такой естественно существующий дисбаланс

становится инициатором динамического взаимодействия между текстовой деятельностью и средой.

В текстовой деятельности количественные изменения происходят в двух плоскостях.

1. Расширяются синтагматические связи, соотносимые с углублением исследовательских процессов в предметной области, т. е. с детализацией имеющихся знаний.

2. Углубляются парадигматические связи, обусловленные расширением исследований, т. е. появлением новых направлений в исследованиях.

Изменения в первой плоскости связаны с усилением синонимии, во второй — с полисемией и омонимией.

Исследования прагматического компонента системы внешних связей дают возможность оценить участие адресата и адресанта коммуникации в обменных процессах между текстом и средой. Этот аспект представляет интерес своей непосредственной связью с кругом социолингвистических проблем, проблем соотношения лексики речи и лексики письменного языка, а также с большинством экстралингвистических проблем.

Предопределяя направления и проблематику научных исследований, социальные условия влияют на изменение (скорее всего, расширение) объема понятий и формирование их систем. Так, роботизация бытового обслуживания развила различное отношение к роботам в США и Японии. Роботы приобретают элементы искусственного интеллекта и способны выполнять некоторые сложные логические последовательности операций. Более того, стало возможным речевое управление роботом. С этими процессами связано появление таких синтаксических конструкций, как *ear robot* (робот с акустическими сенсорами) или *oral robot* (робот с голосовым управлением).

Семантика подобных синтаксических конструкций ясна, и они легко адаптируются в других языках. Но форма их выражения, например, в японском языке, подтверждает важность социолингвистического компонента функционирования одного и того же текста в разных условиях. Возможности японского языка позволяют его носителям подчеркнуть отношение культуры к результатам деятельности человека. Большинство новых понятий, чуждых традиционным антропоцентрическим представлениям японцев об окружающем мире, выражаются как иноязычные, хотя многие из них принадлежат словотворчеству японского подъязыка робототехники. Примерами таких терминов могут быть *wobotto* (женщина-робот) или *rumi robotto* (робот-прислуга). Построенные по англоязычным моделям, они происходят из системы понятий, сформировавшихся в результате развития робототехники в Японии.

Как известно, экономия выразительных средств, вообще свойственная языку и имеющая тенденцию к усилению с развитием

индустриализации, наиболее заметно проступает в устной речи. Характерной чертой последних 15—20 лет является активное проникновение устного речевого стиля в письменный язык. Для русского языка этот период менее продолжителен: факты нормативного использования речевых синтаксисов в письменных текстах с достаточной для объективной статистики частотой обнаруживаются только в последние 6—7 лет. В предметной области робототехники существует много примеров нормализации речевых форм. Но вопрос о существовании устных и письменных форм — это чаще всего вопрос синонимии.

В прагматическом аспекте взаимосвязей текста сообщения и среды важен учет экстралингвистических проблем. Их значение и влияние находятся в прямой зависимости от особенностей каждой конкретной предметной области. В робототехнике значение экстралингвистических проблем определяется спецификой взаимоотношений в системе «человек—машина» и особенностями функционирования интеллектуальных систем в антропоцентрической среде.

Таким образом, функционирование текстов сообщений в коммуникативных процессах в конкретных предметных областях не может быть понято и объяснено в практических целях без досрочного глубокого анализа их взаимодействия с окружающей средой. Собственно и само понятие «функционирование текста» предполагает его неразрывную связь со средой. Текст сообщения и среда его реализации соподчинены функционально. Все их взаимосвязи и влияния предусматривают обеспечение процесса общения и организацию целенаправленного взаимодействия в определенных регламентированных социальных условиях.

2.7. Контрастивная суть аналитико-синтетического процесса

Контрастивная текстология в состоянии объяснить многие наиболее сложные проблемы аналитико-синтетического процесса. С особой остротой такие проблемы выступают в литературном переводе. В целях достижения адекватности синтезируемого текста кроме чисто лингвистических задач самого жанра должны быть учтены традиции и методы теории литературы и предусмотрены трудности перехода содержания произведения через лингвистические и культурные границы. Эта область аналитико-синтетического процесса нуждается не только в литературном и лингвистическом фундаменте, но и в основах семиотической теории и в опыте межкультурных трансформаций.

Прагматический компонент литературной деятельности по синтезу текста лежит на самой ее поверхности: переводу подлежат те

литературные произведения, которые могут обладать познавательной функцией.

В то же время семиотика имеет сильные связи с логикой и общей теорией знаков и релевантна спецификации параметров значений лексических единиц и структуры текста, поскольку сравнительные возможности социологии, психологии и других составляющих необходимы для понимания различий и сходств на разных уровнях коммуникации обществ, групп и индивидуумов.

Представленная ниже модель (рис. 5) специфицирует эти параметры и отображает их взаимосвязи. В ней аналитико-синтетический процесс помещен в среду контрастивной текстологии, которая, в свою очередь, складывается из четырех групп факторов: независимого анализа дискурса, контрастивного анализа, литературной стилистики и контрастивной филологии.

К сожалению, известно очень немного работ, посвященных исследованиям перевода как аналитико-синтетического процесса с лингвистических позиций. Эта область еще не определила свою принадлежность: филологи не осмеливаются вести лингвистические исследования, а лингвисты выводят филологические исследования за пределы своих интересов. Поэтому правая зона предложенной модели представляет собой наиболее мало изученную область аналитико-синтетического процесса. Вместе с тем эта часть модели показывает, что необходимо сделать для достижения понимания комплексного процесса перевода.

В зарубежной лингвистике (А. Лафевер, Г. Виддоусон и другие) предпринимаются попытки аналитических анализов с позиций контрастивной текстологии. Все известные в настоящее время исследования в области контрастивной текстологии ведутся с чисто лингвистической точки зрения. Хотя основным их методом является сравнение стилистических параметров, они безусловно от-

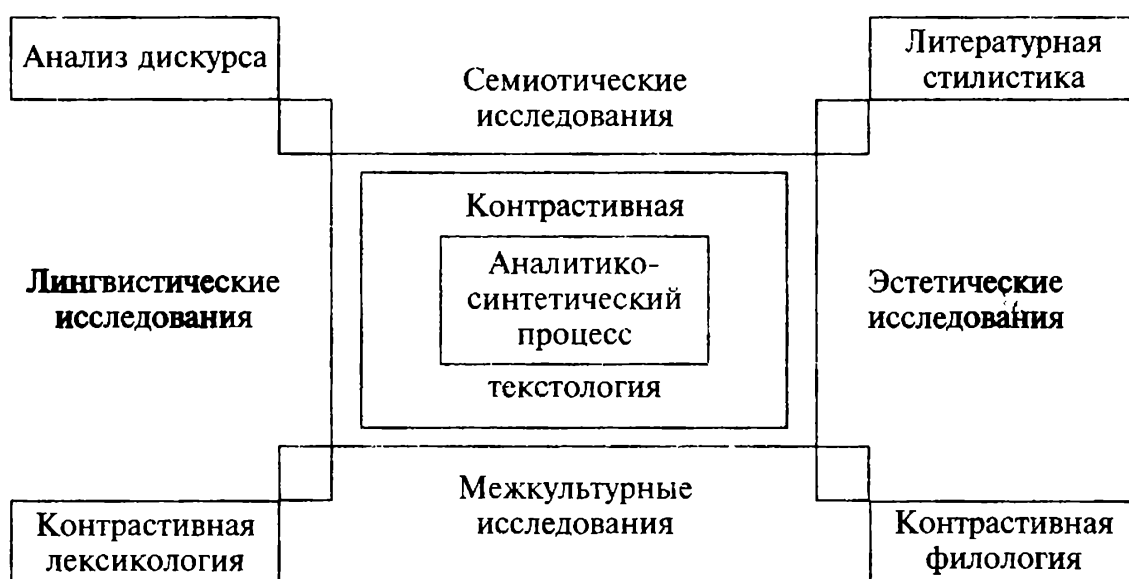


Рис. 5

носят принадлежность изысканий к области контрастивной текстологии и причисляют методы контрастивной текстологии к методике анализа литературного перевода.

Задачей обработчика литературного произведения является синтез эквивалентного текста для читателя, который не понимает его на языке оригинала. Этот процесс комплексной аппроксимации в принципе и включает в себя множество психолингвистических и социолингвистических переменных. То, что можно было бы назвать аналитико-синтетической обработкой текстовой информации, представляет собой набор решений, которые аналитик сознательно или подсознательно принимает относительно уровня соответствия стилистического жанра или степени эквивалентности.

Аналитико-синтетическую обработку можно рассматривать как сервисную услугу для читателя произведения, написанного на неизвестном ему языке, воспринимаемую отдельно от оригинала. Результатом такой обработки первичного текста становится лингвистическое произведение, имеющее отношение к одному, двум или нескольким языкам и включающее понимание и интерпретацию оригинала. И понимание, и интерпретация достигаются через сравнение текстов, т. е. через контрастивную текстологию.

Начиная с парадигматического аспекта контрастивной текстологии, в аналитико-синтетическом процессе необходимо проверить семантические соответствия во всем тексте. Изучение семантической эквивалентности в пределах одного языка или между несколькими языками — одно из основополагающих условий понимания проблем перевода.

В синтагматическом аспекте контрастивной текстологии необходим учет грамматико-комбинаторных связей, которые формируют когерентный дискурс. На этом этапе требуется анализ, уже выходящий за рамки предложения и даже за рамки целого текста, если речь идет о переводе как о динамическом коммуникативном процессе, а не как о пословном перекодировании текста.

Третьим аспектом контрастивной текстологии, от которого зависит понимание структуры дискурса в аналитико-синтетическом процессе, является прагматика. Под прагматикой в данном случае понимаются контекстологические связи между текстом и участниками коммуникативного акта. Этот аспект участвует в создании интерфейса взаимодействия писателя и читателя. Его особенность состоит в том, что именно на этом этапе могут быть найдены или потеряны экстремальные риторико-стилистические элементы, которые обеспечивают адекватность или в значительной мере разрушают синтезируемый текст как литературное произведение. Это — экстремальная точка, где аналитико-синтетический процесс переходит в процесс интерпретации. При выполнении перевода текст, предложенный переводчиком, будет адекватным оригиналу в меру

точности поставленной переводческой задачи. Не менее важна определенность стратегии прочтения оригинала. Существенную роль играют предварительный анализ и учет читательской стратегии, т. е. читательской прагматики. В том случае, когда при восприятии (чтении) более значимой оказывается цель интерпретации, нежели цель извлечения информации, объективный целесообразный перевод может быть достигнут. Тогда это будет перевод, который адекватно представит перцептивный потенциал оригинала.

Прагматическая стратегия аналитико-синтетического прочтения оригинала должна строиться с ориентиром не на анализ использованной в оригинале литературной стилистики или эстетики, а на ясное понимание того, почему использована именно эта стилистика или эстетика.

Контрольные вопросы и задания

1. По каким параметрам перевод относят к аналитико-синтетическому процессу в лингвистической технологии?
2. Какие основные теоретические модели перевода вы знаете?
3. Попробуйте составить модель своих действий при переводе.
4. Назовите коммуникативные особенности перевода как аналитико-синтетического процесса.
5. Что составляет среду функционирования текстовой информации?
6. Как вы понимаете взаимосвязь текста сообщения и среды его функционирования в прагматическом аспекте?
7. Укажите три аспекта контрастного характера аналитико-синтетического процесса.

ОБЛАСТЬ ПРАКТИЧЕСКОЙ РЕАЛИЗАЦИИ ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ

3.1. Лингвистический ингредиент

Перевод — понятие очень широкое, вмещающее в себя самые различные результативные уровни: от художественного перевода, связанного со сложнейшими творческими процессами, до буквального перевода отдельных фрагментов информации. Соответственно использование средств компьютеризации текстовой деятельности в художественном переводе обретает характер вспомогательного инструментария. Скорее всего, в этом виде перевода технологический подход будет невозможен, хотя создание информационно-справочной базы и умение пользоваться ею продуктивно поможет и переводчику художественной литературы. Тексты, описывающие технологии и продукты этих технологий, сами становятся частью таких технологий и частью таких продуктов. Обязательный и неотъемлемый элемент любого товара — инструкция для пользователя на языке страны его реализации. Такие тексты по своей сути являются лингвистическим ингредиентом материального продукта или технологии. Обычно эта часть товара требует перевода на другие языки. Следовательно, для создания и обработки таких текстов (редактирования, реферирования, аннотирования, перевода) вполне можно использовать технологический подход.

Лингвистический ингредиент может быть представлен в виде сопутствующей текстовой информации, web-страницы, пользовательского интерфейса и в ряде других форм.

При экспорте/импорте продукта возникает непростая проблема перевода лингвистического ингредиента и адаптации его к другой языковой культуре. А это означает, что кроме перевода неизбежно решение еще многих задач, связанных с оформлением текста — строжайшее сохранение его форматов или, наоборот, изменение форматов в соответствии с другими правилами; сохранение или изменение иллюстративного материала; транслитерация некоторых надписей и т. п. Лингвисты считают, что перевод (как репродукция на другом языке) составляет лишь 25 % всех затрат по переработке лингвистического ингредиента при экспорте/импорте товаров, услуг и технологий.

Для организации современной технологии перевода важно разделить все возможные технологические операции на две группы: «ручные» (доступные только человеку) и формальные (рутинные),

подлежащие автоматической интерпретации с помощью компьютера. К сожалению, такое деление не может быть абсолютным, а производится лишь с большой долей вероятности. Четкой линии раздела мы добиться не сможем, но более или менее технологичное распределение «обязанностей» между переводчиком и компьютером в современном мире вполне достижимо.

Существует множество «рабочих» определений перевода как деятельности и как продукта этой деятельности. Каждый автор такого определения подчеркивает в нем те параметры, которые важны для него при решении поставленных задач. В самом общем утилитарном смысле под переводом до недавнего времени мы понимали процесс, в результате которого смысл, стиль и целенаправленность текста на одном языке с наибольшей полнотой репродуцируются на другом языке. Этому определению вполне соответствовала ситуация, когда в обязанности переводчика входило представить конкретный текст на другом языке. В своей работе он мог использовать пишущую машинку, несколько словарей, а в лучшем случае — консультировался с авторами исходного текста для уточнения смысла и терминологии. Производительность всего процесса определялась почти исключительно персональной квалификацией переводчика. Для ее повышения не было никаких объективных возможностей. Но и объемы переводов были не такими огромными, с какими приходится иметь дело сейчас. Новые товары появлялись редко, нечасто происходило и обновление уже существовавших моделей. Это легко продемонстрировать очень доступным и понятным примером. Если всего 25—30 лет назад для чистки зубов у нас был выбор между зубным порошком или пастой, то сейчас почти ежедневно мы знакомимся с рекламой все новых и новых сортов, видов и разновидностей зубных паст. На их упаковках содержится исчерпывающая информация о составе, даются инструкции по употреблению, рекомендации по хранению и т. п. Без подобного лингвистического ингредиента товар просто перестает существовать. Никому не придет в голову купить просто белый тюбик, в котором будто бы, согласно самому искреннему уверению продавца, находится зубная паста! Если кто-то и позволит себе с большим чувством сомнения купить тюбик без соответствующей информации на доступном языке, все же успеха такой товар иметь не будет.

Запросы на переводческую продукцию резко возросли вместе с объемами промышленного производства. В переводческом бизнесе также произошли значительные изменения. Появились компьютеры. Даже самое простое неспециализированное программное обеспечение позволяет редактировать тексты и печатать их без тех проблем, что были связаны с пишущими машинками. Но все же «человеческий фактор» в переводческом бизнесе заменить или хотя бы изменить его роль практически невозможно. Считается,

что человек может перевести в день около 2 тыс. слов независимо от того, будет он создавать текст с помощью простой шариковой ручки, пишущей машинки или компьютера.

Но стал совершенно иным характер промышленного производства. Появилось много однотипных товаров, сопровождаемых сходными лингвистическими ингредиентами. Тексты, разъясняющие особенности новых разновидностей товаров, в значительной мере повторяются. Зачастую они представляют собой прежние тексты в новой редакции. Ученые, занимающиеся исследованиями компьютеризации текстовой деятельности, считают, что очередной обновленный промышленный продукт, а следовательно, и его описание отличаются от предыдущей версии в среднем на 30 %. Большинство текстов, сопровождающих обновленную продукцию, представляют собой лишь новую редакцию прежнего текста, который уже был переведен.

Это позволяет переосмыслить организационные процессы в переводческом бизнесе и подсказывает пути автоматизации повторного перевода с помощью специализированных программных средств — переводческого инструментария. Так появился целый класс сервисных программ под названием «переводческие инструменты». В отличие от систем МП эти программные средства самостоятельно никаких переводческих операций не выполняют. Но они оказывают активную и эффективную организационную поддержку переводческого процесса, создают комфортную среду и, беря на себя рутинные операции, повышают производительность труда переводчика.

К таким программным средствам прежде всего относятся системы, накапливающие в памяти фрагменты текстов и варианты их переводов, — это своеобразная память выполненных переводов. Во всем мире такие компьютерные программы называют Translation Memory или сокращенно ТМ-tools. В российском переводе их чаще всего называют ТМ-инструментами. Кроме ТМ-инструментов в настоящее время широко распространены компьютерные «системы терминологического обеспечения» (Terminology Management Systems или TMS). Эти системы автоматически просматривают исходный текст и предлагают варианты переводов терминов, которые уже были приняты переводчиком в выполненных ранее и внесенных в память системы переводах. Разработаны программы для адаптации переводного текста в условиях иноязычной культуры — программы локализации.

Существует еще целый класс сервисных программ для упрощения условий перевода и редактирования текстов в специфических форматах. Например, тексты web-страниц насыщены дополнительной информацией (знаки форматирования, ссылки и др.), которая затрудняет применение средств автоматизации перевода. Сервисные программы берут на себя решение проблемы

конвертирования электронного текста в удобную для перевода форму, а затем уже в переводе восстанавливают все прежние специфические знаки форматирования. Таким образом происходит подстановка текста на другом языке при полном сохранении внешнего вида страницы.

Самостоятельный класс программ представляют собой автоматические переводные и толковые словари, энциклопедии и справочники.

3.2. Терминологичность информации

Не меньшее значение для повышения качества и производительности переводческого бизнеса имеет оперативное терминологическое обеспечение.

Существует много электронных переводных словарей, которые интегрируются с текстовым редактором MSWord. Они хорошо зарекомендовали себя и получили всеобщее признание и широкое распространение не только среди переводчиков. Кроме того, они постоянно пополняются. Во многих случаях можно совершенно официально регулярно пополнять через Интернет лицензионную копию словаря на своем персональном компьютере. Периодически и довольно часто в продажу поступают дополненные новые версии имеющихся словарей. Но в современной ситуации этого не всегда бывает достаточно. Многие бюро переводов и индивидуальные переводчики создают свои узкоспециализированные словари терминов и профессионализмов.

Все предлагаемые современным рынком электронные переводные словари состоят из множества тематических разделов. Каждый раздел представляет собой словарь — зачастую электронную версию какого-либо известного «бумажного» словаря. При поиске нужного термина можно воспользоваться всеми имеющимися разделами-словарями или подключить к поиску лишь некоторые словари из коллекции. Расположив словари в определенном порядке, можно организовать иерархию их приоритетов, соответствующую избранной многозначности. Как правило, разработчики предлагают еще «пользовательский словарь». Этот раздел технически выполнен как и все остальные разделы-словари, но наполнять его терминами и эквивалентами может сам пользователь. Такой словарь будет работать наравне с остальными и с тем приоритетом, которым его наделит сам пользователь.

Большинство ТМ-инструментов обладают функцией терминологического обеспечения. При обнаружении в переводимом тексте слова или словосочетания, которое в прежних переводах было помечено как термин, ТМ-инструмент обращает внимание переводчика на этот фрагмент переводимого текста. Таким образом,

реализуется еще один очень важный параметр повышения качества перевода — принцип единства терминологии, который должен неукоснительно соблюдаться. Естественно, в условиях локальной компьютерной сети принцип единства терминологии будет четко выдержан всем составом переводческого бюро.

В настоящее время уже существуют готовые терминологические базы данных, которые могут быть подключены к ТМ-инструментам в виде дополнительных, сменных или временных блоков. Как правило, такие терминологические базы создаются производителями экспортируемой продукции. При выдаче заказа на перевод сопроводительной документации к своей продукции фирмы обычно передают переводчику терминологическую базу в виде временного блока для ТМ-инструмента переводчика. Это не только ускоряет процесс перевода, повышает его качество, но и унифицирует документацию фирмы.

3.3. Системы машинного перевода

Обособленную и в высшей степени специфическую группу лингвистического программного обеспечения составляют системы МП. В отличие от ТМ-инструментов, которые помогают переводчику лишь организационно рационально использовать свой переводческий потенциал, системы МП выполняют перевод, а человек при этом необходим для организации работы компьютера. В связке «человек — компьютер» происходит противоположное распределение функций.

В отношении к машинному переводу в настоящее время нет определенности. Если вы не верите в его реальность, вы не только не одиноки в этом, но даже не окажетесь в меньшинстве. Если вы признаете возможность машинного перевода и верите в его прогресс, вы также не будете в меньшинстве и у вас найдется много сторонников. Секрет этого двойственного отношения скрывается в понимании назначения машинного перевода и его целесообразности.

Для достижения коммерчески оправданных результатов применения какой-либо системы МП прежде всего необходима мудрость ее пользователя. И вся мудрость состоит лишь в том, чтобы использовать его только тогда, когда это целесообразно. Вряд ли целесообразно обращаться к этой системе для перевода небольшого юридического документа. Профессиональный переводчик сделает это быстро и качественно. А если документ еще и типовой, то он использует свой ТМ-инструмент. Но очень часто возникает необходимость принятия оперативного производственного решения по материалам какого-либо объемного текстового массива на иностранном языке. Что же делать в таком случае? Выпол-

нять перевод, который отнимет много времени и будет стоить очень дорого? А ведь вполне возможно, что по прочтении перевода оперативное решение будет однозначно отрицательным... В такой ситуации машинный перевод оказывается единственно правильным способом. Возможно, его качество будет невысоким, но тем не менее он даст информацию не только для принятия производственного решения, но и для выделения той части текста, которая впоследствии может понадобиться в производственных целях и которую будет целесообразно передать в бюро переводов для выполнения качественного перевода.

Современные системы МП работают примерно по такому же алгоритму, по которому в самых общих чертах работает обычный переводчик. Обычно система выполняет перевод по предложениям. Уже на этом этапе возможны потенциальные ошибки. Переводчик видит весь текст или значительный его фрагмент и понимает контекст. Отсутствие контекста порождает ошибки. Далее система МП производит синтаксический анализ для нахождения наиболее вероятного подлежащего, сказуемого и других членов предложения. К сожалению, далеко не все реальные предложения могут корреспондироваться с моделями, имеющимися в блоке синтаксического анализа системы. Подлежащее, выраженное инфинитивом, очень часто интерпретируется системой МП как сказуемое или его часть, и перевод оказывается некачественным. После синтаксического анализа входного предложения на языке оригинала следует конструирование грамматической конструкции для языка перевода и замена в этой конструкции слов одного языка на эквиваленты в другом языке. Но в условиях многозначности ошибки просто неизбежны.

В этой кратко представленной технологии машинного перевода можно подметить критические точки, в которых несложно увидеть резервы повышения его качества. Если текст оригинала написан четким логичным стилем, если предложения полные и короткие, то синтаксис текста перевода будет вполне удовлетворительным. Как мы уже отмечали, в наше время уже есть солидные фирмы, документацию которых готовят «по совместительству» не разработчики и производители продукции, а квалифицированные лингвисты.

Если задуматься над лексическими проблемами систем МП, то и здесь имеются резервы для повышения качества перевода. Просмотрев текст, направляемый на перевод, лингвист этой системы может указать приоритетный список словарей, которыми должна пользоваться система при переводе. Конечно, такие словари должны быть тщательно распределены по тематикам. Они требуют постоянной корректировки и пополнения, но при современном развитии лексикографии и оснащении ее мощными компьютерными средствами это перестает быть проблемой.

Уже сейчас существует немало фирм, которые готовят свою документацию с учетом того, что она будет переводиться с помощью систем МП. В таких фирмах выработан собственный стиль с ограниченным количеством грамматических конструкций и словарем, исключающим возникновение проблем многозначности. При подготовке документа важным критерием оценки языкового стиля является оценка его переводимости именно той системой МП, которой располагает данная фирма. Учитываются возможные проблемы и предугадываются ошибки, которые могут быть устранены в процессе постредактирования.

Качество машинного перевода и рентабельность его систем значительно повышают две дополнительные технологические операции — предредактирование и постредактирование. В процессе предредактирования происходит доработка текста, подлежащего переводу. Этот этап напоминает операцию перефразирования, которой часто пользуются переводчики. Мысленно переводчик перестраивает данную фразу в более удобную для перевода форму: упрощает грамматические конструкции, уточняет многозначные лексемы. Как правило, крупные фирмы, издающие довольно большие объемы сопроводительной документации, заинтересованы в переводе своей документации сразу на несколько языков. Тщательная лингвистическая проработка и предредактирование исходного текста экономически целесообразны, поскольку повышают качество сразу нескольких переводов.

Нельзя не учитывать еще и тенденцию к формализации языка оперативной коммуникации. В электронной почте, в коротких сообщениях в мобильной телефонной связи, в автоматических подтверждениях на запросы чаще всего употребляются языковые «штампы», что создает выгодную ситуацию для использования систем МП.

Дополнительный шанс для подобных систем предоставляется в связи с новыми тенденциями в делопроизводстве. Компьютеризация текстовой деятельности не только приносит комфорт в современное делопроизводство, но и повышает технологичность текстов. Компьютер упрощает редактирование текста при его подготовке, способствует его дальнейшей формализации, делает его более технологичным. Все это расширяет сферу применения систем МП.

Нельзя не заметить, что постоянно, можно сказать ежедневно, в Интернете появляется все больше сайтов, располагающих системой МП, которой можно воспользоваться в онлайн-режиме. Сейчас уже не стоит вопрос о возможности или невозможности обращения к такому переводу. Существуют только две взаимосвязанные проблемы — определение целесообразности использования той или иной системы МП и поиск средств повышения качества получаемого перевода.

3.4. Основные принципы локализации при переводе

Получение текста на другом языке чаще всего еще не означает завершения технологического цикла перевода. Одним из главных условий выполнения качественного перевода является его функциональная равнозначность по сравнению с текстом на языке оригинала. Следует убедиться, что текст перевода занимает такую же позицию в структуре информации, какая отведена оригиналу. Кроме того, необходимо, чтобы информационное наполнение нового текста было выражено языком, привычным для его потребителя. В отдельных случаях еще требуется соблюдение некоторых формальных условий: таких, как объем текста, допустимый уровень экспрессивности, этическая адекватность и т. п. Например, текст должен физически уместиться в какую-либо рамку на этикетке товара или его звуковая форма по продолжительности должна соответствовать видеоряду рекламного ролика. Довольно часто образность некоторых метафор в разных языковых культурах не совпадает и метафора, переведенная буквально, может сильно повлиять на функциональность всего перевода.

Для решения перечисленных (и многих других) проблем существует специальная технологическая операция — локализация текста перевода. О ее важности свидетельствует тот факт, что сейчас на тему локализации в мире выходит множество книг довольно внушительных объемов. Самой популярной в европейских странах является дважды переизданная за последние пять лет работа Б. Эсселинка «Практический курс локализации»¹.

Такие мелочи на первый взгляд, как последовательность написания календарной даты, точка или запятая в качестве разделителя в десятичных числах и тому подобное, могут существенно повлиять на информационную значимость текста. Особенно наглядно такие несовпадения проявляются, например, при переводе, связанном с продажей компьютерного программного обеспечения в другую языковую среду и в другую массовую культуру.

Тексты системных сообщений, сообщений об ошибках, тексты меню и служебной информации распределены внутри программ. Для человека-переводчика, пожалуй, будет практически нереальным обнаружить все эти текстовые включения, перевести и технически грамотно поместить их в текстах оригинала и перевода.

Развитие информационной технологии в этом направлении проходило в два этапа. Сначала специалисты в области локализации пошли по пути создания вспомогательного переводческого инструмента. Это была сервисная программа, которая извлекала из локализуемой программы всю текстовую информацию, подлежащую переводу. После перевода она возвращала соответствующую

¹ Esselink, B. A practical Guide to Localization. — John Benjamins, 2000.

щие тексты на языке перевода на прежние места в локализуемой программе. На современном этапе развития локализации принята другая технология. При разработке программного обеспечения предусматривается сервисный блок, который сразу же собирает все текстовые сообщения в одном месте с соответствующими адресами и служебной информацией, описывающей правила вывода этих сообщений на экран. Отпала необходимость поиска текстовых включений по всей программе, повысилась надежность самого программного обеспечения и упростилась его адаптация. Собственно, произошло приближение к традиционной технологии перевода — выполняется требование предпереводческой подготовки материала, направленного на перевод. Это совпадает с операциями предредактирования в машинном переводе или рефразирования в традиционном переводе.

Мы в общих чертах рассмотрели ситуацию перевода текстовых сообщений программного обеспечения. Но она типична для нашего времени. Сообщения на небольшие дисплеи сейчас выводят даже стиральные машины. И их тоже необходимо переводить, и происходит это практически по той же технологии.

Еще чаще, даже не подозревая этого, мы сталкиваемся с аналогичной ситуацией, когда нам необходима информация из Интернета. Любая страничка в наиболее распространенном формате HTML принципиально ничем не отличается от компьютерной программы. Только в ней будет больше текста и меньше команд. И если при подготовке компьютерной программы программист-разработчик уже позаботился об этапе «предредактирования» и упростил работу переводчика, то страничку из Интернета никто заранее не готовил для перевода. Но для переводчика, владеющего ТМ-инструментом, проблем здесь не прибавится. Все современные ТМ-инструменты располагают сервисной программой, которая может нейтрализовать специальные команды языка HTML (тэги) на время перевода и восстановить их действие в тексте перевода после завершения всех переводческих операций. Такая сервисная программа в составе ТМ-инструмента называется фильтром. Наиболее универсальные ТМ-инструменты имеют в своем составе достаточно большое количество фильтров, так как в различных информационных технологиях могут использоваться самые разнообразные форматы представления текста (например, XML и другие).

Одна из важнейших проблем, относящихся к локализации, — несовпадение объемов текстов на различных языках. Приведем простой пример. Словосочетание *японский язык* представлено на русском языке 13-ю символами, а на английском (*Japanese*) — 8-ю символами. На японском языке эти слова будут обозначены тремя иероглифами. Конечно, иероглифы сами по себе значительно сложнее кириллицы или латиницы, но, безусловно, места они занимают значительно меньше.

Гораздо чаще, чем это может показаться на первый взгляд, огромной проблемой становится подбор более удобного синонима или способа сокращения текста. Даже европейские языки, зачастую имеющие общее происхождение, могут различаться объемами содержательно одинаковых текстов на 30—35%. Если необходимо перевести довольно большой текст, для управления объемом могут быть использованы обычные редакторские способы, в чем переводчику помогут навыки экономии выразительных средств.

Существуют и более сложные проблемы адаптации текста перевода в другой языковой среде и культуре. Среди них одна из наиболее понятных — оформление текста. Можно назвать немало языков, участвующих в переводе, в которых направления строк не совпадают: слева направо и наоборот; горизонтальное направление и вертикальное. Для решения таких проблем уже не достаточно владения навыками подбора более удачных синонимов или какими-то особыми редакторскими приемами.

Переводчику приходится обращаться к специальным программным средствам локализации перевода. Многие системы локализации основаны на тех же принципах действия, что и ТМ-инструменты (действие по подобию). В них расширены функции форматирования, добавлены функции графического редактирования. Например, программы могут предоставить возможность изменения размеров информационных окон для вывода требуемого объема текста, автоматически ввести функцию прокрутки текста при невозможности изменения размеров окна и др. Современные системы локализации позволяют выполнять все эти операции независимо от разработчика продукции, руководство по использованию которой переводится на другой язык, т. е. для этого не требуются исходные файлы. Системы могут также изменять информацию, сопровождающую готовую продукцию.

Для того чтобы обеспечить своей продукции наиболее широкий рынок, многие дальновидные разработчики предусматривают вероятность последующего перевода сопроводительных текстов на другие языки и необходимость адаптации в другой языковой культуре. Так, японцы иероглифические тексты традиционно пишут строками сверху вниз, а строки располагают справа налево. Но в технической документации иероглифическое написание соответствует европейской традиции представления текста, что упрощает процессы локализации.

3.5. Рентабельность технологии перевода

В деятельности бюро перевода производственного предприятия могут сложиться две принципиально разные ситуации. В одной из них переводы в основном выполняются с иностранного языка на

родной для того, чтобы воспользоваться товарными или технологическими преимуществами поставщика из другой языковой среды. Во второй — переводы выполняются на иностранные языки для продвижения отечественной продукции или услуг на международные рынки. Вообще-то, обе ситуации должны стремиться к сближению.

Продвижение своего товара, как правило, сопровождается предоставлением документации на языке рынка продаж, а потому и исключает необходимость его перевода. Однако все зависит от диверсионной политики предприятия, в которой могут преобладать те или иные подходы. Совершенно очевидно, что в настоящее время должна развиваться тенденция к преобладанию второй ситуации.

По интенсивности преобладания второй ситуации современные бюро перевода делятся на четыре категории.

К первой категории относятся бюро, которые выполняют перевод документации для «событийной» презентации продукции предприятия, подготавливая так называемый «пилотный» комплект документации. В подобном случае не приходится ожидать высокой рентабельности процесса перевода. Это будет «штучный» перевод и его технологичность практически невозможно повысить никакими автоматизированными средствами, если не брать в расчет ставшие уже обязательными переводческими инструментами электронные словари, справочники, энциклопедии и информационные ресурсы Интернета. Но такой перевод обнаружит «узкие места» и в подготовке оригиналов документов, и в самой организации процесса перевода. Если будут зафиксированы все «вот на будущее нужно учесть...», то такое бюро перевода сможет претендовать на место в следующей категории.

Ко второй категории относятся бюро, которые способны влиять на качество подготовки оригиналов и корректировать свои технологические процессы в зависимости от особенностей представляемых текстов. Это способствует выработке характерного стиля работы бюро перевода. Документация всей последующей модификации продукции выдерживается в одном стиле, а расширяющийся ее ассортимент узнается по единому стилю всего товарного ряда. Технология такого бюро перевода уже не может обойтись без внедрения ТМ-инструментов.

Для завоевания позиции в следующей третьей категории необходимо адекватное реагирование на потребности и особенности рынка. Особое значение при этом приобретают презентативные и культурологические проблемы перевода, при решении которых невозможно обойтись без внедрения систем локализации.

К высшей четвертой категории относятся те бюро перевода, которые способны оперативно реагировать на любые из-

менения в политике продвижения продукции своего предприятия: выполнять редактирование документации в процессе модификации продукции, осуществлять локализацию в зависимости от изменений рынка, оперативно подготавливать дубликации на другие языки и проводить унификацию документации одновременно на всех предполагаемых к использованию языках. Такое бюро обретает функции унифицированной мультитекстной службы, для которой характерно использование сетевых версий всех видов программного обеспечения переводческой деятельности.

Приведенная категоризация никак не связана с качеством выполнения переводов. Предполагается, что в любом из рассмотренных случаев конечным продуктом деятельности бюро будет качественный перевод документа на другой язык. Выделение обозначенных категорий характеризует технологичность переводческого процесса и степень его оснащённости новыми информационными технологиями. Конечно, главными основаниями для внедрения дорогостоящих технологий являются объёмы переводческой деятельности бюро и интенсивность и разветвленность экспортной политики предприятия. По параметрам рентабельности то или иное бюро переводов может достаточно уверенно позиционировать себя в любой из описанных категорий, объективно соотнося свою деятельность с объёмами производства своего предприятия и долей лингвистического компонента в его продукции.

Контрольные вопросы и задания

1. Что такое лингвистический ингредиент товарной продукции?
2. На какие две группы делятся все технологические операции современного перевода? Каков основной принцип этого деления?
3. В чем причины стереотипности современной текстовой информации?
4. Подберите несколько инструкций по использованию однотипных товаров и попытайтесь количественно (в процентах) оценить совпадение их текстов.
5. В подобранных инструкциях подчеркните термины и определите:
а) количественное соотношение терминов и слов общей лексики в каждой инструкции; б) соотношение повторяющихся в разных текстах терминов и общего количества терминов во всех инструкциях.
6. В чем вы видите целесообразность использования систем МП?
7. Назовите совпадения и различия технологий машинного и традиционного перевода.
8. Чем продиктована необходимость локализации и каково ее назначение?
9. Перечислите факторы, которые определяют рентабельность использования новых информационных технологий при переводе.

ЧАСТЬ II

ПРАКТИЧЕСКИЕ ОСНОВЫ ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ В ПЕРЕВОДЕ

Глава 4

СИСТЕМЫ МАШИННОГО ПЕРЕВОДА

4.1. Системы машинного перевода и Интернет

Спад интереса к машинному переводу в конце прошлого века привел к тому, что некоторые системы даже прекратили свое существование. Но с появлением и стремительным расширением всемирной сети Интернет многие информационные технологии потребовали внедрения в промышленных масштабах какого-либо оперативного (online) переводческого инструмента. Под требования Интернета подходил только машинный перевод, и это обстоятельство вызвало интерес к новым разработкам. Многие системы, которые ранее являлись промышленными, получили коммерческое развитие и вышли на рынок информационных технологий в качестве полноценного программного продукта, обладающего свойствами товара.

В стадии своего зарождения и первичного развития Интернет был практически полностью англоязычным. Но по мере превращения узкопрофессиональной сети во всемирную паутину количество его языков стремительно приблизилось к количеству языков мира. Известно, что только для 6 % населения Земли английский язык является родным. Но в момент выхода Интернета в мир 90 % всей его информации было на английском языке. В настоящее время количество сайтов на других языках стабильно увеличивается: около четверти информации в Интернете представлено на испанском языке, более трети — на немецком, французском и японском языках. Число сайтов на русском языке также постоянно растет.

В результате увеличивающегося интереса и возрастающих потребностей в осуществлении перевода больших массивов информации многие владельцы наиболее мощных и популярных сайтов приступили к массовому внедрению технологий машинного пе-

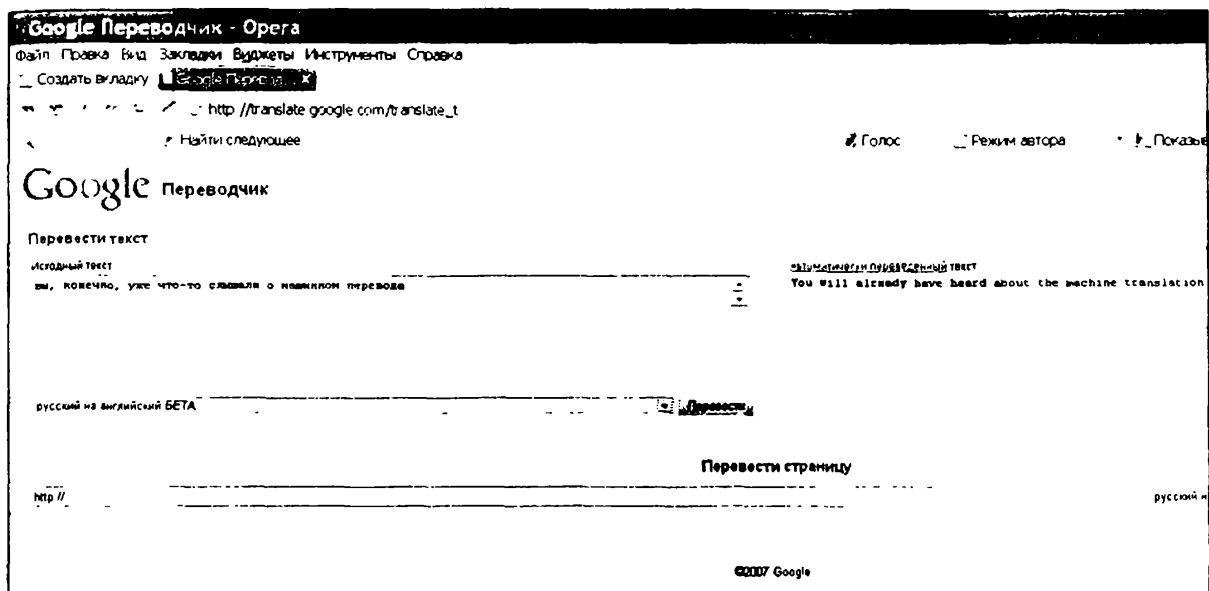


Рис. 6

ревода. Для этого выбираются технологии, позволяющие осуществлять перевод в режиме реального времени. Такие системы применяются в основном на сайтах, связанных с интенсивным обменом информацией в актуальных областях международной деятельности.

Но главным современным достижением является то, что системами МП, работающими в режиме реального времени, оснащены самые распространенные поисковые системы Интернета. На рис. 6 показано, как предложение «Вы, конечно, уже что-то слышали о машинном переводе» переведено на английский язык системой МП поисковой системы Google. Перевод следует признать вполне удовлетворительным, хотя грамматическая форма сказуемого достаточно сложна.

Из интерфейса можно сделать вывод о том, что система в состоянии выполнять переводы целых страниц; возможны и несколько направлений перевода. Первоначальный вариант может быть отредактирован непосредственно в окне системы, использован в среде Интернета или скопирован в текстовый файл.

Рисунок 7 иллюстрирует ответ поисковой системы на запрос «CompuServe».

Предположим, нас заинтересовал второй результат поиска. Как видите, мы могли бы сразу получить ответ на русском языке (имеется предложение системы [Перевести эту страницу]), но мы сначала открыли эту страницу (рис. 8).

Убедившись, что данная страница для нас интересна, запрашиваем у системы ее перевод на русский язык. Результат представлен на рис. 9.

Заметим, что в этом компьютере поисковая система Google настроена на русский язык, поэтому перевод автоматически выполняется на русский язык, но система способна делать переводы на

множество языков. Если у пользователя возникают сомнения в правильности перевода или по каким-то другим причинам ему необходимо посмотреть оригинал, он может вызвать его, отметив заинтересовавший фрагмент перевода. Это и показано на рис. 10.

Поисковая система Google выполняет все эти действия бесплатно. Но, вообще, стоимость машинного перевода такого качества приблизительно в 25 раз ниже стоимости традиционного перевода. Конечно же, качество переводов, выполненных человеком, значительно выше, но получить такие переводы в режиме реального времени в процессе работы в Интернете практически невозможно, и услуги системы МП являются единственным и достаточно хорошим решением проблемы извлечения справочной информации.

Этот факт служит основой для расширения востребованности машинного перевода в режиме реального времени в Интернете.

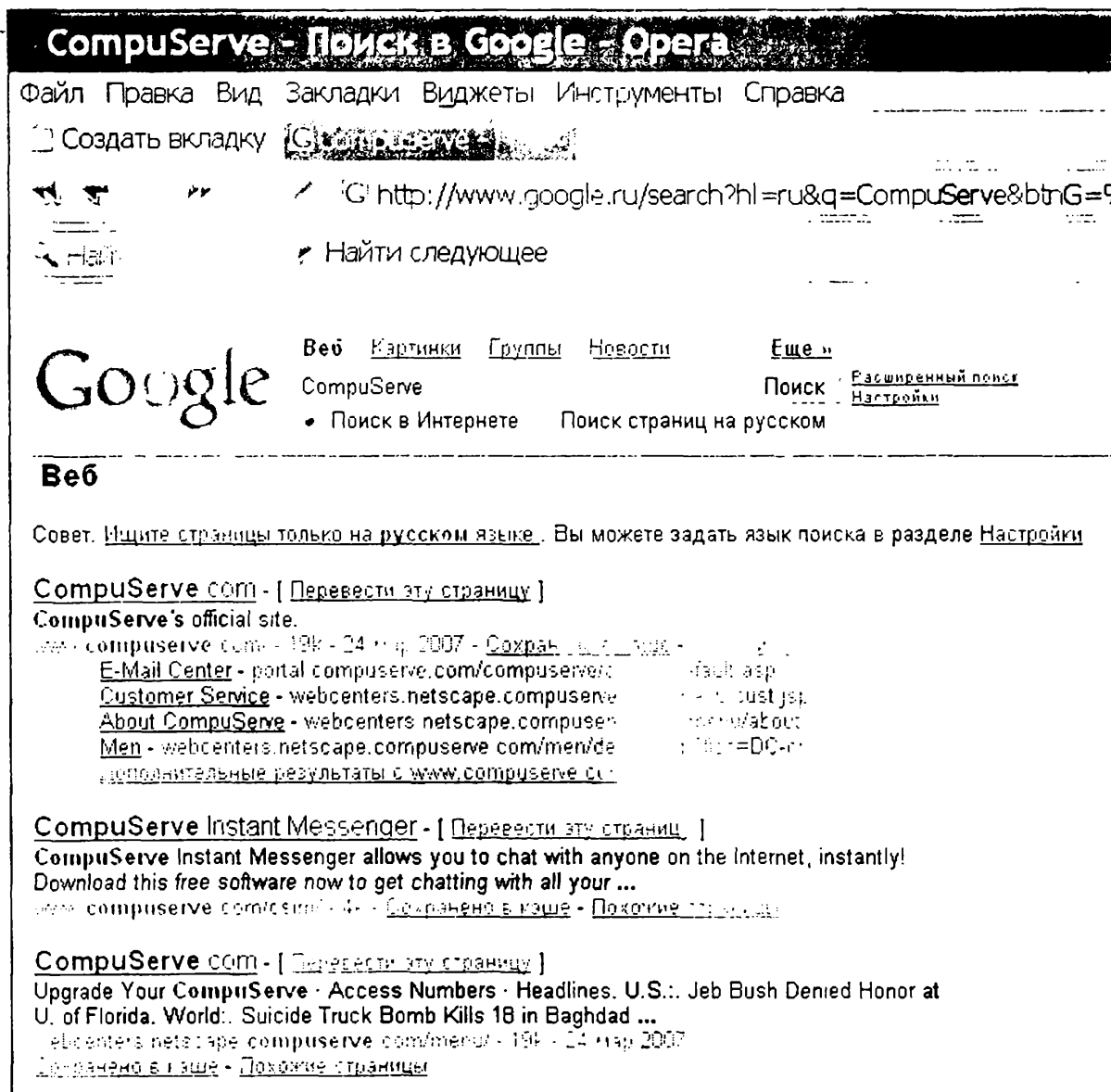


Рис. 7

Instant Messenger

CompuServe

CompuServe Instant Messenger

The coolest place to be online! And it's **FREE!**

Use your Buddy List and join millions of people online with the fastest communication system to hit the net.

- Faster than email
- Talk to your friends (on AOL and CompuServe) one at a time or in a group
- Create your Buddy List so you always have friends online
- Talk with your friends (live) while you work or surf the net

The next best thing to being there!

Get it Now!

[Upgrade](#) [Forget your Password?](#)

Warning: Studies suggest Instant Messenger could be habit forming!

Copyright ©2007 CompuServe Interactive Services, Inc. All rights reserved.
[Legal Notices](#) | [Privacy Policy](#)

Рис. 8

Эта страница была переведена автоматически, исходный язык - английский. ВЕТА
 просмотрите исходную веб-страницу или наведите курсор мышки на текст, чтобы просмотреть исходный текст

Instant Messenger

CompuServe

CompuServe Instant Messenger

Классных место быть в онлайн! И к тому же **FREE!**

Воспользуйтесь Buddy Перечень и объединить миллионы людей в сети с коммуникационной системы динамично попасть чистым.

- Быстрее, чем по электронной почте
- Поговорите со своими друзьями (от AOL и Компьюсерв) по одному или в группе
- Создать свой список Buddy так что вы всегда есть друзья в сети
- Поговорите со своими друзьями (в прямом эфире), а Вы работаете или чистый серф

Следующий Лучше всего быть там!

Get it Now!

[Обновление](#) [Забудьте пароль?](#)

Исследования показывают Instant Messenger может

Рис. 9

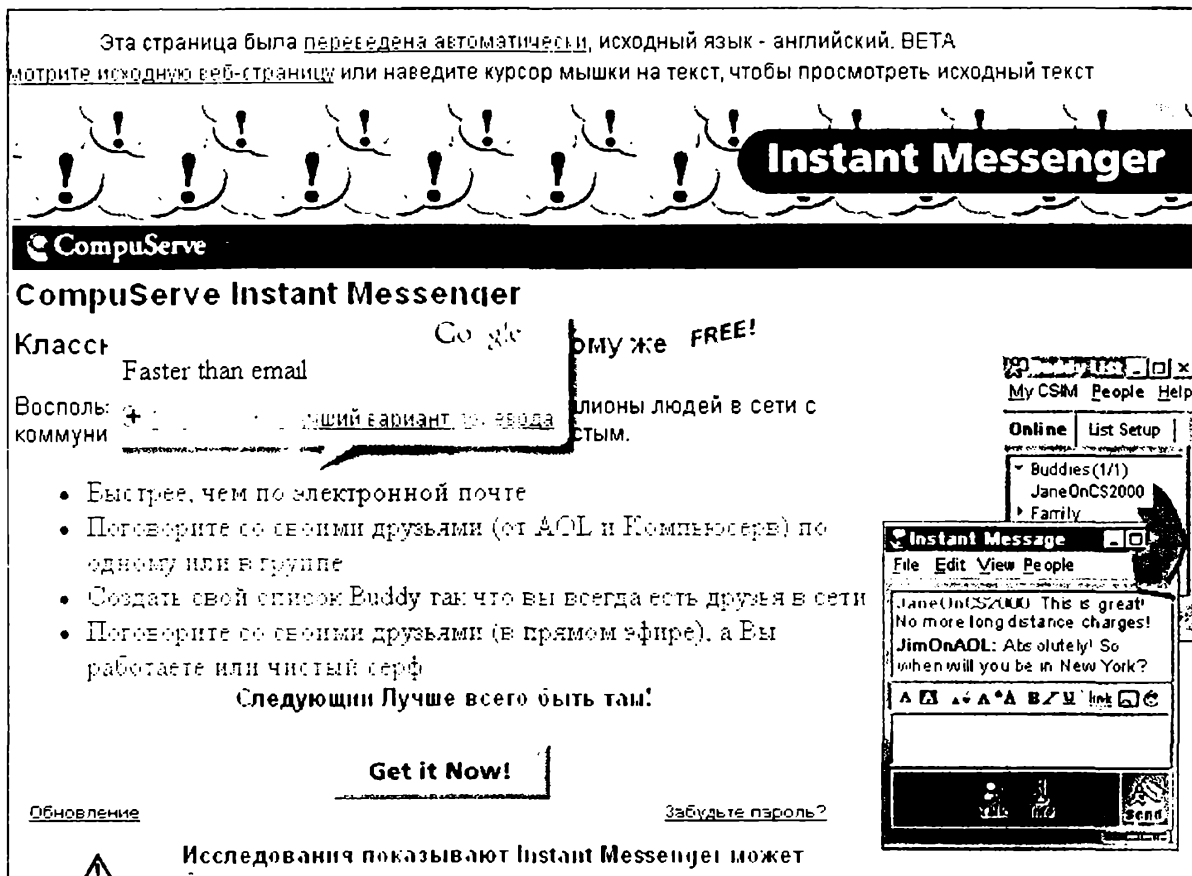


Рис. 10

Для пользователей предоставляются возможности получения бесплатного перевода определенной части текста или даже всей страницы.

Несмотря на бурное развитие информационных технологий в области машинного перевода во всем мире существуют всего несколько систем, профессионально работающих с русским языком.

В нашей стране, к сожалению, количество фирм, занимающихся машинным переводом, чрезвычайно мало. Бесспорный лидер на российском рынке фирма «ПРОМТ». Кроме ряда коммерческих продуктов эта фирма предлагает и систему МП, работающую на сайте фирмы в режиме реального времени. Она позволяет переводить небольшие тексты электронных писем и сразу же отправлять их по указанным адресам. Онлайновая версия располагает весьма ограниченными возможностями, однако полная коммерческая версия вполне конкурентоспособна по сравнению с самыми совершенными системами МП преуспевающих западных разработчиков.

В России, так же как и в других развитых странах, ведутся работы по созданию систем искусственного интеллекта, в частности в области программного обеспечения для сети Интернет. Фирма «ПРОМТ» разработала систему Интернет-переводчик, получившую название PROMT Internet, которая состоит из браузера, про-

граммы для перевода фрагментов текста при работе с браузерами Internet Explorer и Netscape Navigator, а также любыми Windows-приложениями. В настоящее время идет расширение ее возможностей: увеличение количества рабочих языков и направлений перевода, добавление сервисных средств, сохраняющих полное форматирование исходного документа и предоставляющих возможность передвигаться по ссылкам, отображенным на странице перевода.

Подводя итог описанию современного состояния машинного перевода в сети Интернет, хотелось бы отметить, что в последнее время системы МП все больше используются в различных службах информационной сети. Интернет становится той сферой жизнедеятельности, в которой возможно реальное создание мирового сообщества без границ и каких-либо национальных ограничений. Единственное препятствие для плодотворной работы в этом направлении — языковой барьер. Решить эту проблему помогут современные системы МП.

Если после создания Европейского сообщества для перевода все возрастающей документации на языки всех входящих в него народов обратились к помощи системы МП SYSTRAN, то в настоящее время в распоряжении ученых и бизнесменов уже имеется целый ряд программ машинного перевода, разрабатываемых в разных странах и поддерживающих самые различные языки. Среди разработчиков систем МП заметно усиливается тенденция к расширению номенклатуры рабочих языков. В своих исследованиях ученые стремятся охватить как можно большее число не только европейских, но и таких редких языков, как африканские и восточные. Российские разработчики программного обеспечения для систем МП в дополнение к стандартным парам, в которых присутствует русский язык, реализуют новые направления перевода, использующие исключительно иностранные языки, как, например, программа REVERSO фирмы «ПРОМТ», которая выполняет переводы с французского языка на немецкий и обратно.

4.2. Система машинного перевода как лингвистический процессор

Системы МП занимают важное место среди лингвистических автоматов, входя в класс текстовых процессоров. Кроме них к этому классу относятся:

- автоматизированные информационные системы и системы управления базами данных;
- автоматические словари;
- редакционно-издательские системы;
- системы автоматической компрессии текстов;

- системы автоматической адаптации текстов;
- автоматизированные обучающие системы;
- экспертные системы;
- системы шифровки и дешифровки текстов;
- системы автоматической атрибуции текстов.

Являясь лишь частью лингвистических процессоров, системы МП тем не менее играют важную роль в развитии лингвистических автоматов. В настоящее время наибольшее развитие получили системы, относящиеся именно к классу текстовых процессоров. Можно говорить об их непосредственной связи с другими текстовыми (в частности, с автоматизированными информационными системами и автоматизированными переводными словарями) и речевыми процессорами.

Таким образом, они гармонично сочетаются с различными лингвистическими процессорами и являются неотъемлемой частью систем искусственного интеллекта, полномасштабное применение которых планируется в компьютерах пятого поколения. Однако уже сейчас, как отмечалось выше, эти системы входят в единые комплексы, способные выполнять самые разнообразные задачи.

Современные системы МП подразделяются на следующие группы:

- системы прямого типа;
- трансферные системы;
- системы семантического типа.

Подавляющее большинство действующих коммерческих систем МП относится к трансферному типу. Кроме этого, мы полагаем, что, учитывая степень современного теоретического обоснования, совершенно уместно говорить о наличии доминирующего типа систем, а именно трансферного, и об их модификациях — системах, основанных на использовании примеров и параллельных текстов. Все коммерческие программы машинного перевода в той или иной степени представляют собой системы, использующие трансфер.

Системы прямого типа в настоящее время практически не применяются, став достоянием истории. Системы, основанные на использовании баз знаний и языка-посредника, на сегодняшний день все же являются экспериментальными проектами будущего, находясь, по сути, на этапе лабораторных исследований. В связи с недостаточно отлаженными и апробированными механизмами действия и сложной структурой настройки об их промышленном и коммерческом внедрении пока говорить рано.

Сопоставим системы МП в классификации по их функциональному назначению:

- исследовательские прототипы, для которых известен принцип действия системы безотносительно к условиям ее функционирования;

- экспериментальные системы, прошедшие проверку на экспериментальных тестовых массивах;
- промышленные системы, выдержавшие тестовые испытания, наделенные определенными помехоустойчивыми, качественными и скоростными характеристиками и уже работающие с реальными текстами;
- коммерческие системы, обладающие всеми свойствами промышленных систем, предназначенные для коммерческой продажи и представляющие собой законченный самостоятельный продукт, имеющий свойства товара.

Наибольший практический интерес представляют коммерческие системы МП.

4.2.1. Системы прямого перевода

С точки зрения лингвистического обеспечения данные системы МП относятся к процессорам полулексического (полуграмматического) типа. В них выделяются уровни морфологического анализа и морфологического синтеза, но отсутствуют модули полного синтаксического и семантического анализа и синтеза, без которых невозможно воссоздать структуру предложения в терминах той или иной формальной грамматики или семантики.

В лингвистическом обеспечении этих систем МП учитываются элементы синтаксиса и семантики естественного языка при разрешении омонимии, определении грамматических характеристик ряда слов, выборе переводного эквивалента, трансформационных преобразованиях и т. д.

Поскольку полный грамматический анализ и синтез в таких системах не производится, для них характерно невысокое качество перевода, обеспечивающее его понятность в пределах 60—70 %. Результатом работы этих систем является перевод-подстрочник или черновой перевод, как правило, пригодный лишь для сигнального информационного обеспечения специалистов. Он неизбежно требует интенсивного постредактирования, обычно с привлечением оригинала. К тому же ряд систем этого типа предполагает значительное редактирование входных текстов и/или использование весьма ограниченного лексикона и упрощенного синтаксиса.

Качество работы подобных систем в основном определяется степенью разработанности словарей, которые насчитывают десятки и сотни тысяч слов и словосочетаний. В связи с этим весьма актуальна проблема автоматизации разработки и пополнения словарей для систем этого типа.

В настоящее время системы прямого типа или пословно-оборотные системы, как мы уже отмечали, практически не используются, уступив место системам трансферного типа.

4.2.2. Трансферные системы

С середины 60-х гг. прошедшего столетия началось бурное развитие синтаксической теории и активное ее использование при проектировании новых и модернизации существующих систем МП.

Для трансферных систем МП характерен развернутый синтаксический анализ и синтез, причем анализ осуществляется в категориях входного языка, а синтез — в категориях выходного языка. Чтобы преобразовать результаты анализа в категории, используемые для синтаксического синтеза, требуется трансфер, поэтому системы, в которых применяются такие межъязыковые операции, относятся к типу трансферных.

Следует отметить, что современные трансферные системы МП, согласно концепции переводных соответствий, предложенной Ю. Н. Марчуком, целесообразно отнести к системам, основанным на переводных соответствиях.

Как уже отмечалось ранее, исследования в области машинного перевода начались благодаря успехам по дешифровке сообщений. Язык — семиотическая система, элементы которой, хотя она и не является кодом в полном смысле этого слова, могут быть описаны в терминах теории связи. Таким образом, согласно этой теории, для каждой языковой пары существует определенный набор переводящих соответствий. Вследствие отсутствия единой трактовки понятия «единица перевода» и ее корреляции с термином «переводные соответствия» Ю. Н. Марчук предлагает считать данные понятия тождественными. В дальнейшем переводные соответствия будут нами трактоваться как единицы перевода, взятые в динамике переводческого процесса и реализуемые в условиях конкретного перевода определенного текста в рамках данной пары языков.

Переводные соответствия подразделяются на:

- эквивалентные — эквиваленты, появившиеся в силу языкового тождества;
- вариантные — контекстуально-зависимые эквиваленты;
- трансформационные — требующие различных преобразований для правильной передачи смысла.

Перевод эквивалентных соответствий осуществляется довольно просто и заключается в поиске тождественных эквивалентов в словарях. Такие переводные единицы обрабатываются системой МП наиболее полно и правильно.

Вариантные соответствия находятся системой МП не так легко, но в принципе любой из выданных машиной вариантов передает основной смысл предложения. В зависимости от правильного выбора алгоритма можно говорить о большей или меньшей степени потери информации.

Трансформационные соответствия наиболее трудно поддаются машинному переводу вследствие необходимости коренных преоб-

разований выходного предложения. В случае ошибки алгоритма следует говорить уже не о потере информации, а о невозможности правильной передачи смысла. Традиционно такие соответствия относятся к наиболее сложным элементам перевода, при машинной обработке которых отмечается наибольшее число ошибок.

В 1984 г. М. Нагао предложил принцип перевода с помощью примеров или параллельных текстов (Example-Based Machine Translation), который во многом перекликается с концепцией переводных соответствий, выдвинутой Ю. Н. Марчуком.

Принцип действия систем МП, основанных на использовании параллельных текстов, заключается в создании архива переводов, где хранятся уже переведенные предложения, как на исходном, так и на выходном языках. Кроме этого структуры предложений, которые имеют наиболее общий характер, представлены в виде эталонов (как во входном, так и в выходном вариантах), по аналогии с которыми строятся реальные предложения. На практике этот метод выглядит следующим образом. В случае совпадения структуры реального предложения со структурой эталонного вместо основных его компонентов, таких как подлежащее и сказуемое, в выходном предложении по образцу также ставятся подлежащее, сказуемое и т. д., а перевод остальных элементов осуществляется в обычном режиме.

В настоящее время создаются системы МП, использующие комбинированные принципы перевода (multi-engine machine translation). В них наряду с системами трансферного типа предусмотрен архив переводов. Таким образом, при обработке текстов осуществляется комбинированный подход к определению структуры предложения, опирающийся как на результаты анализа входного текста, так и на данные поиска в архиве, где хранятся уже готовые переводные структуры.

Предпосылкой для дальнейшего развития систем, основанных на переводных соответствиях, служат исследования в области решения проблемы автоматического распознавания параллельных текстов. Система STRAND, разрабатываемая в Университете Мэриленд, способна автоматически распознавать и скачивать из сети Интернет параллельные тексты, т. е. оригинал и его перевод, выполненные человеком-переводчиком. В большинстве случаев программа может в автоматическом режиме отличать настоящий перевод от интерпретации. Учитывая глобальный характер сети Интернет и скорость работы программы STRAND, которая за три недели в состоянии переработать до миллиона параллельных текстов, можно прогнозировать значительное расширение архивов и распространение систем МП этого типа. Данная методика позволяет не только создавать обширные архивы параллельных текстов в довольно короткие сроки, но и осуществлять постоянное их пополнение в автономном режиме.

Оценивая состояние современных систем МП трансферного типа, следует отметить доминирующую роль модели «текст — текст», предложенной Ю. Н. Марчуком¹ и являющейся продолжением его концепции переводных соответствий.

4.2.3. Системы семантического типа

В 80-х гг. прошедшего столетия появились системы МП, принцип действия которых основан на использовании семантических баз знаний. Определяющим отличием этих систем от других программ машинного перевода является наличие компонента, включающего экстралингвистические знания. К этому классу относятся системы семантического типа, использующие модель «смысл ↔ текст», разработанную А. К. Жолковским и И. А. Мельчуком.

Результатом исследования в области венгро-русской системы МП стало понятие интерлингвы или языка-посредника, сформулированное И. А. Мельчуком. Проблемы, возникшие в связи с порядком слов в венгерском языке, привели к отказу от пословного метода перевода и предпочтению изучения обычных синтагматических структур (таких, как посессивность и адъективная перегласовка).

За подобными исследованиями большого количества языковых пар естественно должны были последовать поиски синтаксических форм, в некоторой степени присущих всем языкам, но наиболее подходящих для какого-либо одного. Из ряда межъязыковых структур выбирались бы те, которые могли соответствовать языкам входных и выходных текстов. Подобный анализ лексических сходств и различий выявил целый ряд межъязыковых синтаксических единиц (например, передающих сравнение, отрицание, преувеличение и т. д.). С этой точки зрения язык-посредник мог быть представлен в качестве суммы всех соответствий рассматриваемых языков, адаптированной для использования в многоязычных системах.

Последующее развитие И. А. Мельчуком своих идей вылилось в создание его модели «смысл ↔ текст». К сожалению, в настоящий момент нет ни одной действующей системы МП семантического типа, которую можно было бы отнести к промышленным или коммерческим. В некоторых лингвистических центрах ведутся исследовательские работы по созданию системы МП, основанной на использовании баз знаний. В них введен язык-посредник. Однако для коммерческих целей эта продукция пока не предназначена.

¹ См.: Марчук Ю. Н. Модель «текст — текст» и переводные соответствия в теории машинного перевода // Проблемы компьютерной лингвистики. — Минск, 1997. — С. 21 — 29.

Таким образом, говорить о серьезном применении систем МП семантического типа несколько преждевременно. Это направление, возможно, получит широкое распространение лишь после прорыва в области создания систем искусственного интеллекта, когда станут возможными формализация и представление экстралингвистических знаний.

* * *

К основным тенденциям развития современных систем МП следует отнести:

- распространение их номенклатуры — увеличение числа рабочих языков и тематики;
- переход к созданию речевых систем, их связь с текстовыми системами МП и другими речевыми и текстовыми лингвистическими процессорами в интегральных системах;
- сокращение сроков разработки, распространение промышленных и коммерческих систем МП трансферного типа;
- поэтапное совершенствование систем МП с оценкой их эффективности;
- массовое внедрение в сети Интернет и дальнейшее их распространение во Всемирной паутине.

Хотелось бы также отметить, что создатели современных систем МП в последнее время преследуют в основном практические, утилитарные цели. Победа направления «текст — текст» над моделью «смысл ↔ текст», создание специальных «облегченных» грамматик для перевода обыденной речи, массовое внедрение систем МП трансферного типа, основанных на переводных соответствиях — все это свидетельствует, с одной стороны, о невозможности реализации новых интеллектуальных систем, способных уже сейчас решать задачи промышленного и коммерческого применения, а с другой — о востребованности переводов, выполненных системами трансферного типа, которые, несмотря на их несовершенство, все же признаны их потенциальными потребителями.

В настоящее время, как мы полагаем, основной областью применения систем МП станет сеть Интернет, через которую будут осуществляться не только сам перевод, но и дальнейшее распространение новых информационных технологий, неотъемлемой частью которых являются эти системы.

Мы предвидим в сети Интернет настоящий расцвет электронной коммерции. В обозримом будущем следует ожидать роста торговли через Сеть и, возможно, даже бума онлайн-торговли, особенно в Европе, которая по этому показателю отстает от США на полтора-два года. Такой рост может быть связан в первую очередь с глобальной реализацией проекта WebTV. Как известно, телевидение является одним из основных средств массовой инфор-

мации во всем мире. В США многие покупки совершаются с помощью интерактивного телевидения. Если же удастся воплотить в жизнь план соединения телевидения с глобальной сетью Интернет в планетарном масштабе, то это откроет необъятные возможности не только для получения информации, но и для развертывания поистине интернациональной торговли.

Выход на электронный рынок может стать одним из решений значимых проблем нашей экономики, а именно в тех наукоемких отраслях промышленности, которые, несмотря на все потрясения, сохранили свой потенциал и могут представить на мировой рынок конкурентоспособную продукцию, пользующуюся спросом из-за высокого качества и относительно низкой цены. Сейчас мы переживаем эру перехода в область информационных технологий, и Россия может занять в этом «виртуальном мире» достойное место, подобающее ее огромным природным и человеческим ресурсам. Самое важное не упустить время и внедрить данные технологии в производство, тем более, что в последние годы наблюдается снижение стоимости компьютеров и соответствующего программного обеспечения.

Бурный рост торговли в Европе и в какой-то степени в России сдерживается двумя основными факторами: высокой стоимостью соединения и соображениями безопасности при перечислении денежных средств за сделанную покупку. Даже при пользовании банкоматами накладываются определенные ограничения на максимальную сумму, которую можно снять с кредитной карты, так как банк не может полностью быть уверен, что данную операцию совершает действительно истинный владелец карточки. Такого же рода трудности возникают и при торговле в сети Интернет. В ближайшее время, по прогнозам ученых, эти опасения останутся в прошлом. Сейчас ведутся активные работы по созданию биометрических технологий аутентификации пользователей. Фирма «Compaq» совместно с «Identificator Technology» создала идентификатор по отпечаткам пальцев Identificator DFR-200, который способен распознавать пользователя по отпечаткам его пальцев и может быть подключен практически к любому порту компьютера. Компания «Sensar» совместно с лондонским банком «Citibank» проводит эксперимент по апробации технологии сканирования радужной оболочки глаза.

Проблема, связанная с высокой тарифной платой в Европе за пользование телефонной линией при подключении к Интернету, близка к своему разрешению благодаря проведению в ряде стран Интернет-забастовок и бойкотов в Сети. В результате таких акций тарифная плата за доступ в Интернет снижается на десятки процентов. Стоимость подключения к Сети в Европе резко сокращается, становясь в некоторых случаях просто символической. Несколько иная ситуация в России. Если в Москве плата за предос-

тавление вполне удовлетворительного трафика снизилась до 200—300 рублей в месяц, то на расстоянии каких-нибудь 100—120 километров от столицы она выше в 5—6 раз.

Как только ценовые проблемы будут решены или сбалансированы и новые технологии станут доступны широкому кругу пользователей, наступит настоящий расцвет Интернет-торговли. Но он станет возможным только в том случае, если торговля будет вестись по всему миру, а достижение этой цели невозможно без использования систем МП, так как потенциальные покупатели могут и не владеть английским языком, а расходы на перевод, выполненный человеком, будут несоизмеримо высокими.

Все эти факторы способствуют росту капиталовложений в исследования и разработку современных систем МП. Причем главное направление такого рода исследований будет предусматривать разработку программ для использования в сети Интернет. Основной областью применения систем МП должна стать сеть Интернет; программное обеспечение для машинного перевода будет интегрировано не только в веб-браузеры, но и в другие программные продукты, предназначенные для работы в Сети. Принимая во внимание современное состояние разработок в области машинного перевода, следует отметить, что в ближайшем будущем, если конечно не произойдет революционного прорыва в области создания искусственного интеллекта, свое основное применение в сети Интернет найдут системы трансферного типа, так как в подобных случаях не требуется 100%-ной точности перевода, а временной фактор и фактор снижения себестоимости играют решающую роль.

4.3. Современные требования к системам машинного перевода

Оценкой эффективности систем МП занимаются на протяжении длительного времени. Предлагались и предлагаются самые различные методы: от статистических до учета мнения конечного пользователя. Несмотря на все многообразие создаваемых методик, на настоящий момент нет единой общепринятой методики оценки эффективности таких систем. Высказываются мнения о принципиальной невозможности сопоставления различных систем МП в силу неравноценности их программ еще на начальном этапе сравнения. Как можно сравнивать две системы, из которых разработкой одной занималась целая группа лингвистов и программистов при значительной финансовой помощи спонсоров, а другая была создана энтузиастами? Для проведения действительно адекватного и четкого сопоставления по всем значимым параметрам нужны немалые капиталовложения, направленные на про-

ведение идентичной настройки сравниваемых систем и выработки единых критериев.

Ю. Н. Марчук отмечает, что до сих пор нет единого критерия оценки эффективности систем МП, и предлагает учитывать стоимость системы и оценивать ее работу за определенный промежуток времени.

Кроме практического подхода к оценке переводов, выполненных с помощью систем МП, что до настоящего времени проводится в большей степени эмпирическим путем, скорее интуитивно, основываясь на профессиональном опыте и знаниях оценивающего, существуют и теоретические аспекты данной деятельности. В теории перевода адекватность переводов традиционно сводится к категориям семантической (смысловой) полноты и точности, дополняемым стилистической эквивалентностью, включающей, в частности, принцип соответствия текста перевода стилистическим нормам языка перевода. Именно с учетом этих параметров чаще всего и выводится оценка качества перевода. Ю. В. Ванников полагает, что существуют различные типы адекватности перевода: семантико-стилистическая, функциональная и дезидеративная. Кроме этого адекватность, в зависимости от коммуникативной установки оригинала, может относиться к валоративному, инициативному, информационному и селективному типам¹.

Соотнесенность перевода с оригиналом также может быть выявлена на основе их преобразований в базисные структуры и последующего сравнения с учетом некоторого числа допустимых расхождений.

А. И. Новиков считает целесообразным сравнивать тексты оригинала и перевода на основе денотатной структуры предложений².

При таком подходе создается денотатная структура для каждого из текстов. Узлами такой структуры являются понятия, а отношения между ними — предикатами, отражающими связь между узлами. Сравнение происходит на основе выявления близости или отдаленности между соответствующими денотатными структурами.

Э. И. Королев, придерживаясь традиционного подхода к оценке качества перевода, отмечает, что основное требование к системе МП заключается в обеспечении адекватности и понятности переведенного текста без обращения к оригиналу³. Названные характеристики должны проверяться на представительных массивах переводимых текстов.

¹ См. подробнее: Ванников Ю. В. Виды научно-технического перевода: общая характеристика, функции, основные требования. — М., 1988.

² Новиков А. И. Семантика текста и ее формализация. — М., 1983. — С. 131—132.

³ Королев Э. И. Промышленные системы машинного перевода. — М., 1991. — С. 15.

Качество системы МП зависит от многих факторов, но, естественно, главным образом от ее семантической силы, т.е. от полноты словарей и степени развития процедур анализа и синтеза. Кроме того, качество перевода повышается при использовании операций предредактирования, интерредактирования и особенно постредактирования, без которых сегодня трудно представить промышленную систему МП.

Такие промышленные системы должны обеспечивать высокую скорость перевода — порядка десятков и сотен тысяч слов в час. Естественно, что скорость перевода в системе зависит также от сложности лингвистического и программного ее обеспечения (в первую очередь от языка программирования), от наличия в комплексе системы автоматических корректоров входного текста, операций пред-, интер- и постредактирования.

Э. И. Королев считает, что изначально промышленные системы МП должны обладать современной техникой сканирования, хранения и передачи текстовой информации. Они должны быть связаны с компьютерными редакционно-издательскими системами, прежде всего с программами автоматической коррекции орфографии.

Проблемами понятности и адекватности перевода, выполненного системой МП, занимались не только отечественные ученые, но и их зарубежные коллеги. Существуют различные зарубежные методики оценки понятности переведенного машиной текста. Представим для примера 5-балльную шкалу оценки перевода, предложенную Нагао:

1 — смысл предложения понятен и не возникает никаких вопросов; грамматика, словоупотребление и стиль соответствуют общей структуре текста и не требуют постредактирования;

2 — смысл предложения понятен, но возникают большие проблемы с грамматикой, словоупотреблением и стилем;

3 — общий смысл предложения понятен, но трактовка некоторых его частей вызывает сомнение из-за неправильных грамматического строя и словоупотребления, а также стилистических ошибок; требуется обращение к оригиналу;

4 — в предложении имеется большое количество грамматических, словоупотребительных и стилистических ошибок; его смысл можно понять с трудом после внимательного изучения;

5 — смысл предложения непонятен.

Однако, несмотря на многообразие предлагаемых оценочных суждений, ни одно из них не может считаться объективным. Уровень понимания текста реципиентом всегда зависит от индивидуальных, а значит, субъективных факторов (уровня образованности, степени знакомства с предметной областью, представленной в тексте, и т.п.).

Понятие адекватности подразумевает обязательное обращение к тексту оригинала с тем, чтобы выяснить, насколько точно пе-

ревод соответствует первоисточнику. В такого рода экспериментах обычно участвуют специалисты, хорошо владеющие как языком оригинала, так и выходным языком, на котором собственно и выполняется перевод. Критерий адекватности служит для подтверждения правильной передачи смысла оригинала, так как нередки случаи, когда реципиент прекрасно понимает содержание текста, но это содержание не в полной мере соответствует содержательной стороне исходного текста. В качестве примера приведем 7-балльную оценочную шкалу адекватности, также предложенную Нагао:

1 — содержание предложения на исходном языке (ИЯ) в полной мере соответствует содержанию выходного предложения; необходимости в преобразованиях нет, носитель языка полностью понимает смысл выходного предложения;

2 — содержание предложения на ИЯ соответствует содержанию выходного предложения; носитель языка понимает его смысл, но необходимо некоторое его преобразование;

3 — содержание предложения на ИЯ передано в выходном предложении правильно, но требуются преобразования в порядке слов;

4 — несмотря на правильную передачу общего смысла исходного предложения, в выходном варианте возникают проблемы с согласованием времен, координацией между членами предложения, правильным употреблением наречий;

5 — содержание предложения на ИЯ не совсем правильно передано в выходном предложении; некоторые выражения отсутствуют; имеются нарушения согласования и употребления главных и придаточных предложений, а также дополнительных членов предложения;

6 — содержание предложения на ИЯ передано в выходном предложении неправильно;

7 — содержание выходного предложения полностью не соответствует содержанию предложения исходного; его структура не соответствует грамматически правильной структуре предложения, подлежащее и сказуемое отсутствуют.

Критерий адекватности наряду с критерием понятности — важнейшие элементы оценки качества перевода. К сожалению, в настоящее время не создано методик, позволяющих проводить адекватный и объективный анализ переведенных текстов. Формализация и автоматизация данного процесса требует больших временных и финансовых затрат и на современном уровне развития техники представляется невозможной ввиду трудности отражения и использования экстралингвистических знаний в компьютерных системах. Оценка, произведенная человеком, в той или иной мере субъективна и может широко варьироваться в зависимости от личности исследователя.

Кроме учета названных критериев и подсчета объема редакторской правки существуют различные подходы к оценке действующих систем МП, предусматривающие значимость разнообразных параметров.

Для исследования большинства коммерческих систем МП независимые эксперты применяют принцип «черного ящика», когда предположение о внутренней структуре системы МП и о ее типе делается на основании оценки выполненных ею переводов. Так как большинство поступающих в продажу систем МП обладают свойствами продукта, то их разработчики прилагают максимальные усилия для того, чтобы принципы функционирования системы и алгоритмы перевода составляли коммерческую тайну. В этом случае выявление основных принципов функционирования может основываться исключительно на результатах переводов, полученных опытным путем и на их последующем анализе.

В противовес принципу «черного ящика» используется принцип прозрачности системы, иначе именуемый «glass box». Он применяется при оценке эффективности системы ее создателями и разработчиками и состоит в поэтапном отслеживании прохождения анализа и синтеза, в процессе которого можно определить, какой модуль или какой алгоритм функционирует неправильно. Такая процедура необходима при доводке системы и введении в нее поправок.

Широкое применение получило использование тестовых массивов для контроля качества работы программы. При этом могут быть введены как реальные, так и искусственно созданные тексты для проверки правильности перевода того или иного языкового явления.

В качественной и многоплановой оценке эффективности системы МП важную роль играют и сама по себе целевая направленность исследования, и состав его непосредственных участников. Реальной оценке подлежит не качество перевода в целом, а его приемлемые результаты, предназначенные для конкретных, узких областей и задач, в которых предполагается их использовать. В связи с этим целесообразно обозначить потенциальных участников и заказчиков исследований в области машинного перевода. Многие считают приемлемой следующую градацию: исследователи, спонсоры исследований, разработчики, покупатели, переводчики, конечные пользователи переводов.

Исследователи занимаются изучением основополагающих проблем машинного перевода и принципиальными возможностями их создания и совершенствования. Суть их интересов сводится к выявлению новых феноменов в функционировании систем, разработке исследовательских прототипов и их последующей проверке, в модификации существующих моделей.

Цель оценочных исследований, инициируемых спонсорами, заключается в уточнении факторов, свидетельствующих либо в пользу продолжения финансирования проекта, либо, наоборот, о необходимости прекращения инвестиций в связи с его неперспективностью или высокими затратами. Иногда от результатов таких анализов зависит не только судьба какого-либо конкретного проекта, но и уровень финансирования целого научного направления.

Разработчики систем МП активно сотрудничают с создателями систем. Эта категория специалистов занимается практической доработкой, превращая исследовательский прототип в реально действующую систему, обладающую свойствами реального продукта и пригодную для коммерческого распространения. Основопологающим фактором проводимых ими исследований является прагматический аспект функционирования программы, простота ее интерфейса, помехоустойчивая работа алгоритмов и приемлемое качество переводов. Создатели систем МП занимаются качественной модернизацией уже действующих систем в рамках функционирования теоретической модели. Довольно часто в целях повышения качества перевода они проводят исследования, направленные на повышение эффективности системы.

К категории покупателей относятся все группы пользователей, начиная от крупных компаний и переводческих служб и заканчивая конечными пользователями и частными лицами. Люди этой категории наряду с лингвистическими факторами оценки эффективности систем МП широко применяют и экстралингвистические критерии, к которым относятся стоимость самой системы, затраты на ее эксплуатацию и обучение персонала, возможность и стоимость обновлений, эффективность работы службы поддержки клиентов, конкурентоспособность разработчиков системы. Эта довольно большая группа пользователей оценивает систему и с прагматической точки зрения, основываясь на экстралингвистических факторах ее функционирования.

В отличие от пользователей-покупателей, пользователей-переводчиков в большей мере интересуется лингвистическая сторона функционирования систем. Рассматривая качество перевода, такие специалисты в первую очередь оценивают лингвистическое обеспечение, пользовательский интерфейс и пути своевременного пополнения системы. В силу своей специализации переводчики способны оказать весьма значимую помощь при оценке и доработке лингвистических компонентов программы.

Конечными пользователями можно считать и покупателей, и переводчиков, и любого человека, который в той или иной степени нуждается в системах МП в своей деятельности. Принимая во внимание широту понятия «конечный пользователь» и невозможность его полного раскрытия в рамках данной работы, мы

хотели бы рассматривать конечного пользователя как специалиста, не владеющего языком, на котором составлен текст оригинала. Его оценка качества перевода зависит от нескольких факторов, но прежде всего от цели ознакомления с переводимым текстом, и от того, является ли он специалистом в области, с которой связан перевод. Целью ознакомления с переводом может быть лишь его просмотр для извлечения основного смысла. При подобном «сигнальном» чтении к качеству перевода предъявляются минимальные требования — он должен быть в целом понятен и должен отображать наиболее общий смысл. При детальном ознакомлении с текстом перевода он должен соответствовать определенным требованиям: наиболее полно и правильно передавать терминологию, перевод которой должен учитывать все новейшие изменения в словарном составе обоих языков; грамматический строй выходного предложения должен в основном отражать структуру входного предложения (этот фактор не играет решающей роли, когда с текстом знакомится эксперт, так как для него стилистические погрешности могут быть восполнены глубоким знанием рассматриваемой предметной области).

Системы МП кроме лингвистической компоненты состоят также из программного обеспечения. В связи с этим были разработаны специальные стандарты. Для определения качества программного обеспечения систем созданы стандарты ISO 9126, затем ISO 14000. В их основу положены шесть специально выявленных характеристик качества программного обеспечения: функциональность, надежность, практичность, эффективность, открытость и портативность. Эти критерии применимы также ко всем остальным типам программного обеспечения.

4.4. Типология ошибок при работе систем машинного перевода

Проблемы оценки качества перевода неизменно присутствуют как в традиционном, так и в машинном переводе. Несмотря на все многообразие рассмотренных нами теоретических и практических подходов к решению этих проблем, их объединяет общая тенденция: какой бы ни была методика оценки качества переводов, она так или иначе исследует, классифицирует ошибки, допущенные в процессе перевода, и пытается их устранить.

Из всего обилия методов оценки перевода нам представляется целесообразным использовать принцип «черного ящика», так как практически все рассматриваемые системы являются коммерческими и специфика их внутренней организации и функционирования закрыты для широкого пользователя, составляя коммерческую тайну.

Для типологии ошибок, выявляемых при машинном переводе, анализу будет подвергаться выходной текст, в рамках которого осуществляется поиск и классификация допущенных ошибок. Теоретическим обоснованием данной типологии служит модель переводных соответствий, предложенная Ю. Н. Марчуком. Помимо непосредственно лингвистических факторов, мы считаем необходимым рассмотреть и экстралингвистические параметры наиболее распространенных систем МП, включающие в себя стоимость, пользовательский интерфейс, простоту работы с системой и возможность ее обновления.

Если рассматривать интересующий нас вопрос в широком аспекте, следует выделить два основных подхода к созданию машинного перевода: дедуктивный и индуктивный. Дедуктивный подход получил свое развитие в 80-х гг. прошедшего столетия. Его суть заключается в идее создания универсального языка, воплотившего в себе основные характеристики всех языков мира. На универсальном понятии интерлингвы базируются системы семантического типа. Однако, как уже отмечалось, такие системы до настоящего времени не нашли практического применения. Господствующий в современных системах индуктивный подход предусматривает создание практических моделей для получения машинного перевода. Примером такой модели является концепция переводных соответствий Ю. Н. Марчука, послужившая теоретической основой для типологического анализа ошибок при работе систем МП.

Теоретические гипотезы о наличии регулярных соответствий в языках перевода были высказаны еще классиками «традиционного перевода». Так, Я. И. Рецкер предложил деление соответствий на эквиваленты, аналоги и адекватные соответствия, что перекликается с типологией, выдвинутой В. Н. Комиссаровым, согласно которой соответствия подразделяются на постоянные, варианты и окказиональные. Эти разработки получили определенное отражение в модели переводных соответствий, созданной Ю. Н. Марчуком.

Данную модель можно представить в двух основных состояниях: генерации и трансляции. Этап генерации необходим во время разработки систем МП и систематизации элементов описания. Мы не будем подробно останавливаться на этом состоянии, поскольку избранный нами прием «черного ящика» в принципе не позволяет его рассматривать, а кроме того, этот этап подробно описан автором.

С учетом специфики нашей работы основными объектами исследования являются состояние трансляции и его соотношенность в рамках модели переводных соответствий с современными системами МП. Именно этот этап представляет собой непосредственное осуществление перевода на основе языкового материала.

К состоянию трансляции в качестве его вторичного признака будут отнесены элементы пополнения словарей и инструментарий для частичной предметной корректировки алгоритмов анализируемых систем.

Модель переводных соответствий состоит из двух компонентов: предметного и динамического. Сразу отметим, что динамический компонент в рамках данной работы рассматриваться не будет ввиду невозможности исследования лингвистического обеспечения и алгоритмов изучаемых систем. Подавляющее большинство исследуемых программ представляет собой коммерческие продукты, программная и лингвистическая компоненты которых защищены законом об авторских правах и имеют закрытый характер.

Предметный компонент содержит два элемента: собственно предметный и переводной. В состав первого входят лексика, грамматика и семантика, которые во взаимодействии образуют текст.

Переводной компонент состоит из разных типов переводных соответствий, с учетом которых будет строиться типология ошибок в работе системы МП. Как уже отмечалось, мы выбрали для рассмотрения три типа соответствий — эквивалентные, варианты и трансформационные.

Приведем определение переводного соответствия, предложенное Ю.Н. Марчуком Переводным соответствием (актуализацией единицы перевода) автор модели называет пару, образованную отрезком входного текста и отрезком выходного текста, между которыми существует отношение коммуникативной эквивалентности¹. Из сказанного следует, что результат переводческих действий является основополагающим фактором для выявления и классификации переводных соответствий.

Эквивалентные соответствия на уровне слов могут выступать в процессе перевода в следующих основных комбинациях:

- слово — слово;
- слово — словосочетание;
- словосочетание — слово;
- словосочетание — словосочетание.

Для вариантных соответствий на уровне слов и словосочетаний используются наиболее типичные комбинации сочетаний. Если выявлять соответствия на уровне текста, то их вариативность будет наблюдаться в подавляющем большинстве случаев.

При трансформационных соответствиях на уровне слов и словосочетаний характерны достаточно сложные преобразования. На уровне предложений трансформационные преобразования применяются при:

¹ См.: Марчук Ю.Н. Проблемы машинного перевода. — М., 1983. — С. 61.

- анализе и переводе предложений с отрицанием и отрицательными наречиями (*no, nor, nowhere, nothing, nobody*¹ и т.д.);
- перемене порядка слов для цепочки существительных, из которых первое является определением к последующим (*information retrieval*);
- анализе и переводе глагольных конструкций со вторичной предикацией (*I saw her dance, I heard him playing piano, I saw her entering the room*);
- анализе и переводе конструкций, использующих такие глаголы, как *to like, to be reported to* и т.д.;
- анализе и переводе специальных конструкций (*there is / are, was, were*).

Развивая предложенную Ю.Н. Марчуком модель переводных соответствий, некоторые лингвисты предлагают уточнить типологию и рассматривать их в качестве некоторых заданных соответствий, выявленных в результате контекстологического анализа и взятых в динамике реального переводческого процесса с использованием систем МП. Согласно уточненной типологии выделяются следующие типы переводных соответствий на различных уровнях:

- отдельных слов
 - однозначные (табличные),
 - многозначные (схемные);
- словосочетаний
 - табличные,
 - схемные;
- грамматических конструкций
 - схемные;
- грамматических категорий
 - схемные.

Под табличными понимаются независимые от контекста соответствия, которые задаются в системе в виде жестко взаимосвязанных таблиц. Схемными считаются соответствия, выбор которых, несмотря на конечную заданность всех возможных вариантов, зависит от контекста.

Основываясь на модели переводных соответствий и результатах эмпирических исследований современных коммерческих систем МП, проведенных по принципу «черного ящика», П.Н. Хроменков выделил две группы ошибок²:

1) ошибки автоматического анализа (лексические, морфологические, синтаксические, текстовые);

¹ В примерах мы используем наиболее распространенные в машинных переводах англоязычные слова.

² См.: Хроменков П.Н. Анализ и оценка эффективности современных систем машинного перевода: автореф. дис. ... канд. филол. наук. — М., 2000. — С. 15—16.

2) ошибки автоматического синтеза (текстовые, семантические, синтаксические, морфологические, лексические). Рассмотрим более подробно ошибки каждой группы.

4.4.1. Ошибки автоматического анализа

При автоматическом анализе, как мы уже указали, выделяются ошибки следующих типов: лексические, морфологические, синтаксические и текстовые.

Ошибки лексического типа наиболее транспарентны при оценке качества машинного перевода. К ним относится появление в выходном тексте непереведенных или неправильно переведенных слов, что может быть обусловлено неполнотой автоматического словаря, или орфографическими ошибками во входном тексте, или неправильным решением задач лексической омонимии и полисемии.

Наличие таких ошибок отражает наиболее характерные проблемы лексического анализа, вызванные неполнотой словаря и неправильным разрешением омонимии и полисемии. Словарные ошибки, относящиеся к разряду лексикографических, могут быть достаточно легко исправлены путем дополнения словарей и создания при необходимости алгоритмов разрешения омонимии и полисемии. В настоящее время многозначность достигается путем разработки и введения в систему узкоспециализированных словарей, позволяющих пользователю наиболее точно задать контекст переводимого текста.

К ошибкам, сопряженным с конкретной работой автоматического анализа, относятся нарушения согласования и управления. Они могут быть вызваны неправильной работой не только морфологического, но и синтаксического, а также семантического блоков. Примером может служить неправильно заданное падежное управление.

Морфологический анализ выявляет ошибки, связанные с неправильным определением грамматических планов.

Синтаксический анализ обнаруживает ошибки в определении (неправильном или неполном) синтаксической структуры входного предложения, что в свою очередь ведет к неверному синтезу выходного предложения. Несмотря на общность ошибок синтаксического анализа и синтеза, нам представляется целесообразным рассматривать их независимо друг от друга.

Среди ошибок, полученных в результате текстового анализа, самым распространенным является неправильное распознавание антецедентов местоимений, анафорических отношений и эллиптических структур. К сожалению, эта проблема до настоящего времени не имеет кардинального решения, так как распознавание и понимание названных структур человеком основывает-

ся на использовании понятийно-категориального аппарата и на осмыслении контекста. На нынешнем уровне развития вычислительной техники полное разрешение данной неоднозначности алгоритмическим путем невозможно. Таким образом, ошибки текстового анализа могут быть устранены в основном на этапе редакторской правки, выполняемой непосредственно человеком.

4.4.2. Ошибки автоматического синтеза

По завершении автоматического анализа и необходимых трансферных преобразований наступает этап автоматического синтеза. Синтез выходного предложения для систем МП трансферного типа ведется в категориях выходного языка. При такого рода преобразованиях, равно как и при автоматическом анализе, системы допускают целый ряд ошибок на уровнях текстового, синтаксического, морфологического и лексического синтеза. Рассмотрим эти ошибки более подробно, проиллюстрировав их на примере переводов.

Большинство ошибок текстового синтеза заключается в неправильном синтезе связочных конструкций, антецедентов и т. д.

При синтаксическом синтезе наиболее распространенной ошибкой является неправильный порядок слов в выходном тексте. Особенно остро данная проблема возникает при переводах с немецкого языка: порядок слов в немецком предложении в некоторых случаях кардинально отличается от русского или английского языков.

В процессе морфологического синтеза выявляются такие ошибки, как неправильное употребление форм глаголов, неверное согласование составного сказуемого, некорректный синтез слов, принадлежащих к разным частям речи.

К ошибкам лексического синтеза мы относим не только собственно лексические ошибки, но и нарушения в управлении. Это мы считаем уместным в связи с тем, что в автоматическом словаре функции сильноуправляющих предлогов в подавляющем большинстве случаев включаются в информационную ячейку слова.

Завершая классификацию ошибок перевода по этапам работы системы МП, следует отметить, что подобное их выделение во многом формально, поскольку изолированное рассмотрение данных проблем в принципе неперспективно. Ошибки, допущенные на этапе лексического анализа, оказывают влияние на морфологический, синтаксический и текстовый этапы анализа и могут повлечь за собой появление новых ошибок на указанных этапах. Неправильная работа системы на этапе анализа и некорректное разрешение возникающих неоднозначностей в свою очередь ведут к появлению ошибок на этапе синтеза. Исходя из этого, мы полагаем, что ошибки, проявляющиеся в результате работы системы МП, следует рассматривать системно, не только с учетом инфор-

мации, имеющейся на выходе, но и с привлечением дополнительных сведений о работе тех или иных алгоритмов перевода на каждом языке системы.

В процессе работы над классификацией ошибок, выявленных в результате функционирования системы МП, была обнаружена зависимость между типом ошибок и частотой их появления при работе системы и переводными соответствиями.

Переводные соответствия лексического типа подробно рассмотрены Ю. Н. Марчуком и положены в основу его модели переводных соответствий. Мы же предполагаем рассмотреть переводные соответствия лексического типа с точки зрения оценки качества работы современных систем МП и выявления ошибок перевода.

Эквивалентные соответствия лексического типа представляют собой диалектическое противоречие, с одной стороны, являясь наиболее простыми элементами для перевода, а с другой — составляя едва ли не самый обширный класс переводческих ошибок. Ввиду однозначного соответствия лексических единиц входного и выходного языков их перевод осуществляется по установленной схеме, т. е. внесением в автоматический словарь значений соответствующих единиц. Несмотря на достаточно несложную процедуру представления эквивалентных соответствий лексического типа в системах МП, наиболее частые и заметные ошибки связаны именно с отсутствием того или иного слова в автоматическом словаре системы. Такая ситуация характерна не только для эквивалентных, но и для вариантных и трансформационных лексических соответствий. Ошибки, относящиеся к эквивалентным соответствиям этого типа и вызванные неполнотой словаря входного языка, могут быть оперативно исправлены без привлечения специалистов. Пользователь, обладающий минимальными знаниями в области языка, в состоянии самостоятельно пополнить словарь без ущерба для общего качества перевода. Ошибки, связанные с невозможностью перевода таких соответствий, легко устраняются при отладке системы путем добавления в словарь новых терминов и не влекут за собой новых ошибок.

Вариантные соответствия также составляют один из наиболее крупных классов лексических ошибок при машинном переводе. В отличие от эквивалентных, вариантные соответствия лексического типа и ошибки, вызванные их неправильным переводом, не могут быть столь успешно и быстро устранены. Разрешение многозначности в процессе машинного перевода решается не только словарными, но и алгоритмическими методами. В настоящее время разработчики систем МП пошли по пути создания максимального числа узкоспециальных словарей, в которых были бы отражены все известные значения перевода вариантных соответствий лексического типа. Тем не менее данная проблема пока не полу-

чила однозначного решения, и пополнение словарей пользователей может вызвать некорректную работу системы в других массивах текстов. Оптимальным было бы решение проблемы многозначности самими разработчиками системы, т.е. конструктивными способами, но это не всегда осуществимо.

Трансформационные соответствия лексического типа — наиболее сложный вид соответствий для перевода посредством систем МП. И в большинстве случаев он выполняется некорректно, так как требует значительного преобразования структуры словосочетаний. Наибольшее распространение ошибки в трансформационных соответствиях получили на синтаксическом и текстовом уровнях, причем на последнем они играют ведущую роль. Трансформационные соответствия будут рассмотрены нами более детально на синтаксическом уровне.

Разберем особенности функционирования переводных соответствий на морфологическом уровне. Эквивалентные соответствия морфологического типа, как правило, остаются неизменными и образуют переводное единство в рамках взятого текста. Под такими структурами понимаются слова, морфологические характеристики которых сохраняются при переводе на иностранный язык и не требуют кардинальных преобразований. При переводе существует лишь один правильный вариант преобразования морфологической структуры, что позволяет отнести ее к разряду эквивалентных соответствий и выявить устойчивые закономерности функционирования сходных структур. Суть этих закономерностей в том, что сходные структуры, вне зависимости от контекста и прочих факторов, будут иметь одинаковый перевод на уровне морфологических характеристик.

К вариантным соответствиям морфологического типа относятся конструкции, при машинном переводе которых имеется несколько взаимоприемлемых вариантов отображения морфологических характеристик. При выборе любого из предложенных системой вариантов общий смысл перевода кардинально не меняется, иногда можно лишь отметить некоторую потерю стилистически ориентированной информации. Причем данные системой варианты считаются приемлемыми для машинного перевода с последующим выбором из них доминирующего при постредактировании, выполняемом человеком.

Трансформационные соответствия морфологического типа встречаются довольно редко, если говорить о них в чистом виде. Однако, принимая во внимание специфику работы системы МП, представляется возможным отнести к этому типу те грамматические формы, полный эквивалент которых отсутствует в другом языке. В таком случае для перевода надо произвести трансформации и синтезировать новые структуры, эквиваленты которых отсутствуют в языке перевода.

Минимальной единицей переводных соответствий синтаксического типа является предложение. Целесообразно также выделить эквивалентные, вариантные и трансформационные соответствия на уровне предложений.

К эквивалентным соответствиям синтаксического типа мы относим предложения, структура которых на входном языке полностью соответствует структуре предложений на выходном языке. В этом случае перевод может выполняться не только с привлечением эквивалентных соответствий синтаксического типа, но и с использованием в отдельных случаях эквивалентных соответствий лексического и морфологического типов.

Кроме синтаксических признаков вариативности перевода необходимо учитывать и структурные признаки, которые используются при алгоритмическом обеспечении перевода.

Трансформационные соответствия синтаксического типа подразумевают полную перестройку структуры выходного предложения по сравнению со структурой входного предложения. Ввиду сложной трансформации предложения выполнение машиной правильного перевода такой структуры крайне затруднено и влечет неизбежные ошибки. Для иллюстрации приведем пример, заимствованный у В. Н. Комиссарова (из Ч. Диккенса)¹.

If the funeral had been yesterday, I could not recollect it better.

Правильный перевод на русский язык выглядит следующим образом:

Я помню эти похороны так, как будто они были вчера.

А вот какой перевод был получен П. А. Хроменковым с помощью системы МП PROMT 98:

Если похороны были вчера, я не мог бы вспоминать их лучше².

Точно так же это предложение переводит PROMT XP (см. рис. 11).

Оба примера машинных переводов свидетельствуют о том, что структуры, в которых необходима трансформация, не всегда переводятся правильно, а потому требуются знание контекста и редакторская правка.

Подводя итог проведенному типологическому обзору ошибок перевода, возникающих при работе систем МП, подчеркнем, что данную проблему обязательно следует рассматривать в комплексе.

Предложенная нами классификация по этапам работы анализа и синтеза представляет картину функционирования системы, которая становится полной только после рассмотрения всех состав-

¹ Комиссаров В. Н. Слово о переводе. — М., 1973. — С. 198.

² Хроменков П. Н. Анализ и оценка эффективности современных систем машинного перевода: автореф. дис. ... канд. филол. наук. — М., 2000. — С. 21.

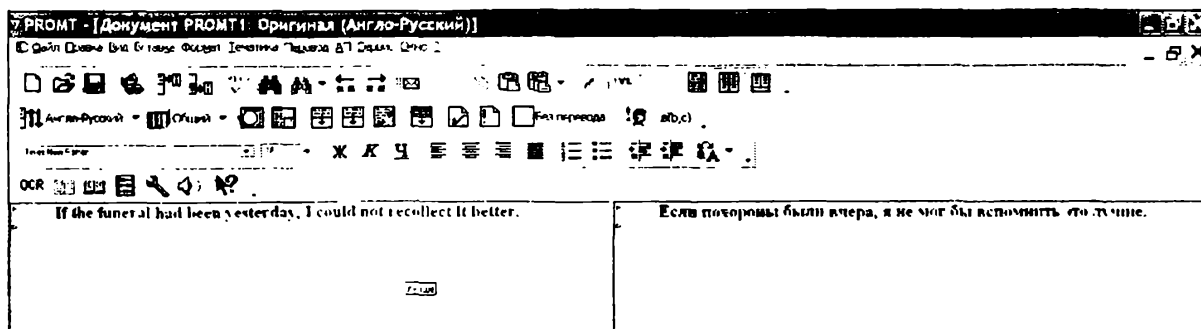


Рис. 11

ляющих ее элементов. При оценке систем МП на основе принципа «черного ящика» в ряде случаев невозможно однозначно отнести переводческую ошибку к тому или иному типу без обращения к изучению алгоритмов системы. Несмотря на значимость выделенной типологии, надо иметь в виду, что все ошибки перевода взаимосвязаны и часто одни из них порождают новые неточности. В связи с этим уместно говорить о комплексном подходе к оценке эффективности систем МП. Только четкое и подробное описание программы позволяет правильно классифицировать системы МП и характерные для них ошибки.

* * *

В результате типологического исследования ошибок машинного перевода стало возможным уточнение модели переводных соответствий с вычленением новых типов соответствий, свойственных функционированию современных систем МП. Предложенная нами классификация требует отдельного исследования, выполнение которого в рамках данной работы не представляется возможным.

Необходимо также отметить, что на основе рассмотренной классификации возможна разработка и дальнейшая эксплуатация систем МП, основанных на использовании примеров. Расширенная классификация позволяет определять соответствия и с учетом их характеристик исправлять ошибки на всех языковых уровнях, используемых в современных системах. Предложенная типология положена нами в основу оценки работы современных коммерческих систем МП.

4.5. Оценка эффективности основных действующих систем машинного перевода

Главными критериями оценки действующих систем МП правильнее всего будет избрать качество выполненного перевода, его понятность и адекватность оригиналу, а также функциональные возможности систем и их стоимость. При суждении об эффек-

тивности системы доминирующую роль, безусловно, играют лингвистические факторы, так как от них зависит оценка работы системы. Экстралингвистические факторы играют второстепенную роль, но тем не менее они важны для практической реализации системы, когда соотношение цена/качество имеет определенное значение и отражается на возможности ее применения в той или иной области.

4.5.1. Краткая характеристика систем машинного перевода: PROMT и SYSTRAN

Представление информации о современных коммерческих системах МП следует начать со сравнительного описания двух ведущих программ, разработанных в России и Соединенных Штатах Америки. Система PROMT занимает лидирующее положение среди российских систем МП, а система SYSTRAN возглавляет список зарубежных программ, осуществляющих перевод с английского языка на другие и обратно. Детальное описание двух этих систем позволяет раскрыть весь спектр функций и возможностей, которыми в той или иной степени располагают остальные системы МП.

Сравнение этих программ хотелось бы предварить краткой исторической справкой об их развитии и становлении в качестве коммерческих систем.

SYSTRAN относится к системам МП трансферного типа и подробно описана в связанной с этими проблемами зарубежной литературе. Ее разработка началась в 1970 г., а с 1976 г. она используется в Комиссии Европейского сообщества. В результате такой широкой эксплуатации ее состав постоянно пополняется новыми рабочими языками. В структурах обработки информационных потоков в нашей стране на официальном уровне эта система используется очень редко.

Система обладает несколькими языковыми парами для перевода, как правило, в обоих направлениях, а в отдельных случаях — в одном направлении, преимущественно на английский язык. В комплект поставки также входит SYSTRAN WebTrans, позволяющий осуществлять перевод Web-страниц. Кроме того, эта система с начала 1998 г. интегрирована в состав сервера AltaVista, что позволяет осуществлять автоматический перевод найденных Web-страниц.

В состав SYSTRAN входит несколько десятков специализированных словарей, представляющих собой самостоятельные подключаемые по желанию пользователя блоки.

У истоков отечественной системы МП Stylus стояла группа «Статистика речи», возглавляемая в 70-х гг. прошедшего столетия Р. Г. Пиотровским, которая работала в стенах Ленинградского

(ныне Российского) государственного педагогического университета им. А. И. Герцена (Санкт-Петербург) и занималась пословно-пооборотным переводом. На базе этой исследовательской группы в 1991 г. была основана компания «ПРОМТ», которая стала заниматься исследованиями в области прикладной лингвистики, разработкой технологий и систем МП для европейских языков. В 1993—1999 гг. эта компания была признана независимыми аналитическими агентствами «Dator» и «ЛамИнфо» одной из 100 наиболее влиятельных компаний в компьютерном бизнесе России. В 1995 и 1998 гг. ей присваивались награды за выдающийся вклад в развитие информационных технологий в стране.

В 2001 г. был значительно расширен список предлагаемых направлений перевода (созданы системы с испанско-русским, англо-немецким, испанско-английским и французско-испанским направлениями перевода) и по их числу компания вышла в мировые лидеры. В XXI век компания «ПРОМТ» вошла с многофункциональной системой МП PROMT XT Family. Это система трансферного типа, на базе которой компания выпускает достаточно широкий ассортимент коммерческих продуктов: PROMT Internet XT, PROMT Internet XT Premium, EXPRESS, PROMT XT Standard, PROMT XT Office, PROMT XT Professional и PROMT XT Premium. В различных сочетаниях в эти коммерческие системы входит более десятка основных модулей с самостоятельными функциями.

Наибольшее количество модулей имеется в составе комплекта PROMT XT Premium. Основной его модуль PROMT — главный компонент всей системы — представляет собой профессиональную среду для перевода и настроек его режимов. Модуль Dictionary Editor Pro позволяет пользователю создавать и редактировать собственные словари, подключаемые к основной системе и используемые в процессе перевода. Модуль PromtX — редактор-переводчик с возможностью автоматического перевода буфера обмена Windows и текстов, введенных с клавиатуры в любом приложении Windows, работающем с текстом. Программа PromtE может работать совместно с Microsoft Internet Explorer для перевода Web-страниц. Модуль PROMT Plug-In for Acrobat выполняет функции перевода документов формата PDF в среде Adobe Acrobat. Для реализации функции перевода в приложениях Microsoft Office 2000 и Microsoft Office XP (Word, Excel, PowerPoint, FrontPage, Outlook) используется программа Smartool. Модуль Electronic Dictionary — это электронный словарь, содержащий всю словарную базу специализированных словарей компании «ПРОМТ». Для интеграции специализированных словарей в систему включается программа TDSetup. Функции менеджера баз данных ассоциированной памяти осуществляет программа Associated Memory Manager. В состав поставки PROMT XT Premium входит утилита резервного сохранения/восстановления данных системы PROMT Backup. Для па-

кетного перевода файлов предназначена программа File Translator. Модуль Dictionary Services представляет собой набор средств для профессиональной работы со словарями (статистика, сравнение, слияние словарей). В систему входят также программа для формирования дистрибутивов специализированных словарей и шаблонов тематик SdcMaker и средство для автоматизации процессов извлечения и управления терминологией PROMT Terminology Manager. Система МП компании «ПРОМТ» может работать совместно с системой памяти перевода TRADOS, для чего в состав PROMT XT Premium включен модуль PROMT for TRADOS.

Комплект PROMT XT Office отличается от PROMT XT Premium тем, что в нем отсутствуют модули File Translator, Dictionary Services, SdcMaker, PROMT Terminology Manager и PROMT for TRADOS. А в комплекты PROMT Internet XT и PROMT Internet XT Premium входят лишь блоки PromtE и TDSetup. Наиболее распространенный комплект PROMT XT Standard содержит PromtE, PROMT Plug-In for Acrobat, Electronic Dictionary и TDSetup.

Требования системы МП PROMT XT к конфигурации компьютера более чем скромные, и даже в самой полной комплектации она может быть установлена на любой существующий в настоящее время компьютер.

Кроме собственно системы с тем или иным количеством модулей для успешного машинного перевода необходимы словари. Система PROMT XT Professional может включать практически любое количество словарей. Полная коллекция словарей, которые выпустила компания «ПРОМТ», называется «Гигант» и включает следующие словари:

1. Авиационный (англо-русский и русско-английский).
2. Аэрокосмический (англо-русский и русско-английский).
3. Банк (англо-русский и русско-английский).
4. Биология (англо-русский и русско-английский).
5. Бытовая техника (англо-русский).
6. Военный (англо-русский и русско-английский).
7. Горно-технический (англо-русский и русско-английский).
8. Деревообрабатывающая промышленность (англо-русский и русско-английский).
9. Добыча нефти и газа (англо-русский и русско-английский).
10. Информатика и вычислительная техника (англо-русский и немецко-русский).
11. Кино (англо-русский).
12. Коммерция (англо-русский и русско-английский, немецко-русский и русско-немецкий, французско-русский и русско-французский).
13. Контракты (англо-русский и русско-английский).
14. Косметическая промышленность (англо-русский).
15. Кулинария (англо-русский).

16. Логистика (англо-русский и русско-английский).
17. Математика (англо-русский и русско-английский).
18. Машиностроение (англо-русский и русско-английский).
19. Медицина (англо-русский и русско-английский, немецко-русский и русско-немецкий).
20. Менеджмент (англо-русский и русско-английский).
21. Металлургия (англо-русский и русско-английский).
22. Микроэлектроника (англо-русский и русско-английский).
23. Морской (англо-русский и русско-английский).
24. Музыка (англо-русский).
25. Парфюмерия и косметика (англо-русский и французско-русский).
26. Пищевая промышленность (англо-русский и русско-английский).
27. Полиграфия и издательское дело (англо-русский и немецко-русский).
28. Политехнический (англо-русский и русско-английский, немецко-русский).
29. Путешествия (англо-русский и русско-английский).
30. Религия (англо-русский и русско-английский).
31. Сельское хозяйство (англо-русский и русско-английский).
32. Спорт (англо-русский).
33. Строительство (англо-русский и русско-английский, немецко-русский).
34. Таможня (англо-русский и русско-английский).
35. Текстильная промышленность (англо-русский и русско-английский).
36. Телевидение и видео (русско-английский).
37. Телекоммуникации и связь (англо-русский и русско-английский).
38. Технология нефтедобычи (англо-русский и русско-английский).
39. Физика (англо-русский и русско-английский).
40. Финансы (англо-русский и русско-английский).
41. Химия (англо-русский и русско-английский).
42. Экология (англо-русский и русско-английский, французско-русский и русско-французский).
43. Электроэнергетика (англо-русский и русско-английский).
44. Юридический (англо-русский и русско-английский, немецко-русский и русско-немецкий, французско-русский и русско-французский).

В исследовательских целях к системе была подключена вся коллекция словарей. Но довольно трудно представить себе, что все они действительно могут понадобиться одновременно. В связи с этим компания сгруппировала эти словари и предлагает их в виде шести коллекций: «Коммерция», «Наука», «Тяжелая про-

мышленность», «Легкая промышленность», «Техника» и «Домашняя».

Для успешной работы современной системы МП словарь, пожалуй, имеет самое большое значение. Один из основоположников системы PROMT Р. Г. Пиотровский считает, что до 80 % информации научно-технического текста содержится именно в его лексическом составе¹. Лексика несет в себе основную смысловую нагрузку текста, а синтаксис только уточняет нюансы смысла.

Вот почему большую часть беспрецедентного конкурентного преимущества системе PROMT приносят очень полные по своим словникам и совершенные по лингвистическому обеспечению системные словари. Система поставляется вместе с генеральным словарем объемом более 100 тыс. общеупотребительных слов и словосочетаний. Подключение специализированных тематических словарей дает возможность расширить объем лексической базы системы вдвое.

Синтаксическая структура предложения в переводе прямо связана с адекватностью перевода его лексического состава. Поэтому специализированные словари содержат не только переводные эквиваленты, выстроенные в необходимой иерархии, но и наиболее частотные синтаксические конструкции, характерные для употребления этих эквивалентов в текстах определенной тематической направленности. Например, словарь по информатике и вычислительной технике ориентирован на частотное употребление в текстах сказуемого в повелительном направлении или инфинитива в функции подлежащего.

Специализированные словари системы представляют собой самостоятельные коммерческие продукты, поэтому закрыты для пользователей. В них невозможно вносить какую-либо правку самостоятельно. Но разработчики постоянно корректируют и пополняют словари, регулярно выпуская их новые версии. Кроме того, система предоставляет возможность создавать свои собственные пользовательские словари, которые могут подключаться к работе системы наравне с остальными специализированными словарями.

Подключение специализированных словарей достаточно эффективно решает одну из самых сложных проблем перевода — проблему нахождения адекватных эквивалентов для многозначной лексики. Например, в общем смысле существительное *run* означает *бег, пробег, пробежка* (*We took a run around the track. — Мы побегали по беговой дорожке.*), а в компьютерном словаре за этим словом закреплены эквиваленты *выполнение, запуск* [программы]. При чтении бытового текста потребуется генеральный словарь, а

¹ См.: Пиотровский Р. Г. Предисловие // Нелюбин Л. Л. Перевод и прикладная лингвистика. — М., 1983. — С. 4.

компьютерный словарь может ввести в заблуждение. И наоборот, при переводе текста, связанного с работой какого-либо программного средства, генеральный словарь будет почти бесполезен.

Одновременно могут использоваться сразу несколько словарей или даже все существующие в системе. Но они будут обладать различным приоритетом: первый словарь в списке используемых для перевода данного текста будет иметь наибольший приоритет. В процессе работы список словарей (а следовательно, и приоритеты) можно изменять. Однако необходимо учитывать и тот факт, что в процессе перевода система обязана просмотреть все словари, подключенные в данный момент. Это требует определенного времени, поэтому нецелесообразно использовать все словари сразу. Правильный подбор тематических словарей для выполнения конкретного перевода и логичное распределение их в списке подключения эффективно способствует повышению качества полученного машинного перевода.

В документации к коллекции словарей для системы PROMT XT Family¹ разработчики приводят интересный пример.

Сопоставляются переводы некоторых слов, данных в трех англо-русских словарях различной направленности — по коммерции, информатике и в генеральном (табл. 1).

Таблица 1

Слово	Словари		
	Коммерция	Информатика и вычислительная техника	Генеральный
Store	<i>запас</i>	<i>память</i>	<i>склад (магазин)</i>
Capitalized	<i>превратил в капитал</i>	<i>печатал прописными буквами</i>	<i>превратил в капитал</i>
File	<i>картотека</i>	<i>файл</i>	<i>файл</i>

Предположим, что нам нужно перевести предложение в тексте, содержащем лишь общеупотребительную лексику: *He is in the store.* (Правильный перевод: *Он находится в магазине.*)

Допустим, что к системе подключены два специализированных словаря и приоритетным является словарь по коммерции. Тогда перевод будет таким: *Он находится в запасе.* Если изменить расположение словарей, отдав приоритетное место словарю по информатике, перевод будет: *Он находится в памяти.* Только если отключим оба специализированных словаря и оставим один генеральный, мы получим правильный перевод: *Он находится в складе (магазина).*

¹ Коллекция специализированных словарей. Руководство пользователя. — СПб., 1991—2002. — С. 11—13.

Теперь посмотрим, как система переведет предложение из текста по программному обеспечению: *We capitalized the names of files.* (Правильный перевод: *Мы печатали имена файлов прописными буквами.*)

Предположим, что подключены оба специализированных словаря и приоритетом обладает словарь по коммерции. В этом случае мы получим следующий перевод: *Мы превратили в капитал имена карточек.*

Чтобы добиться правильного перевода, нужно либо вообще отключить словарь по коммерции, либо сделать его менее приоритетным, чем словарь по информатике. Тогда перевод будет правильным: *Мы печатали прописными буквами имена файлов.*

Из приведенных примеров ясно, что для целенаправленной работы не следует подключать лишние, не относящиеся к тематике переводимого текста специализированные словари. Если текст содержит слова из нескольких предметных областей, надо определить, какая из них преобладает, и сделать соответствующий словарь приоритетным.

Преимуществом версии PROMT XT Office по сравнению с предыдущей версией PROMT Translation Office 2000 являются¹:

1. Новый модуль, повышающий качество перевода для 55—65 % предложений.

2. Реализация революционной технологии ассоциированной памяти, обеспечивающей самообучение системы и дополнительное улучшение качества перевода. Ассоциированная память работает по принципу накопления «опыта»: в процессе перевода сохраняется исходный сегмент (предложение или его часть) и его перевод. При подключении базы ассоциированной памяти система сравнивает каждое предложение с содержащимися в базе сегментами. Если сегмент, идентичный исходному, найден, он может быть использован для перевода. Пример записей для ассоциированной памяти приведен на рис. 12.

3. Новая структура словарной базы данных и лингвистические алгоритмы, использующие ее преимущества и обеспечивающие высокое качество перевода. Лингвистические алгоритмы PROMT XT Office характеризуются следующими принципиальными изменениями:

– переходом на открытую архитектуру словарных баз данных, что позволило разработать и иметь возможность создавать в будущем более сложные алгоритмы. В предыдущей версии она имела жесткую структуру и ограничивала объем грамматической информации, который можно было хранить в словаре. В линии XT это ограничение снято;

¹ Далее используется техническая документация фирмы «ПРОМТ», поставляемая пользователям вместе с программными продуктами, а также материалы сайта этой фирмы.

– переходом к технологии «синтеза выходного языка», при которой происходит не только правильный анализ структур, но и более точная их передача при переводе, что в целом улучшило понимание перевода и требует меньшей корректировки при редактировании;

– появлением новых возможностей для пользователя в работе со словарем, система стала более «открытой».

4. Эффективный лингвистический редактор, обладающий возможностью интерактивного управления переводом (на уровне синтеза):

– синхронная подсветка связанных фрагментов в оригинале и переводе;

– работа с вариантами перевода;

– режимы специального поиска: поиск всех форм искомого слова с использованием подстановочных символов и указанием направления поиска.

5. Электронные словари для всех направлений перевода.

6. Генеральный словарь увеличился на 35 тыс. единиц.

7. Встроенная поддержка макросов (PROMT Scripting Tools) и готовые примеры макросов:

– поиск и выделение BOLD по всему тексту строки;

– обработка текста, содержащего примеры исходного кода программ на языке Visual Basic;

– замена части исходного текста на заранее подготовленный перевод (некоторый аналог Translation Memory);

– преобработка и постобработка текста для перевода разреженного текста (типа «Hello World»);



Оригинал	Перевод
The long-term investments we make today will determine the value of our business tomorrow.	Долгосрочные капиталовложения, которые мы делаем сегодня, определяют стоимость
In order to make long-term investments in new product lines, new equipment and other assets, managers must	Чтобы делать долгосрочные капиталовложения в новых производственных
The cost associated with different sources of funds is called the cost of capital.	Стоимость, связанную с различными источниками финансирования называют стоимостью капитала.
Cost of Capital represents the rate a business must pay for each source of funds -	Стоимость капитала представляет норму {разряд}, которую бизнес

Сохранить все Отмена Справка

Рис. 12

Типовой документ 1



Типовой документ 2

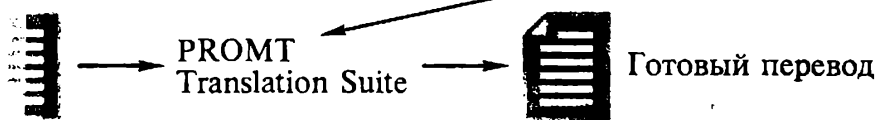


Рис. 13

– резервирование слов с большой буквы, не являющихся первыми словами в предложении;

– резервирование всех слов, начинающихся с большой буквы и имеющих большую букву в своем составе (типа *TextEditor*), вне зависимости от места слова в предложении.

8. Интеграция в приложения Microsoft Office XP (Word, Excel, PowerPoint, FrontPage и Outlook).

9. Перевод PDF-документов непосредственно в Adobe Acrobat 4.x/5.x и Adobe Acrobat Reader 4.x/5.x.

10. Поддержка интеграции с новыми версиями FineReader, CuneiForm, Lingvo, MultiLex, Contex.

Весь этот набор функций работает безукоризненно. Важно отметить, что соотношение цена—качество систем МП фирмы «ПРОМТ» всегда было одним из главных ее преимуществ по сравнению со всеми предложениями сегодняшнего рынка. При правильном подходе к оценке качества и функциональных возможностей системы она вполне работоспособна в современном деловом мире.

В настоящее время в продаже имеется новая версия PROMT Translation Suite 7.0¹ — высокоэффективный инструмент для максимально качественного перевода типовых документов, сходных по своей тематике и структуре (контракты, инструкции, статьи, описания программных продуктов, законы и т.д.). Система обеспечивает максимальное сохранение структуры и форматирования исходных документов. Новая технология обладает возможностью одновременного использования функции системы МП и ТМ-инструментов (ТМ Trados). Интеграция двух функций в одной технологии позволила добиться синергетического эффекта — усилить преимущества обеих систем, одновременно минимизировав их недостатки.

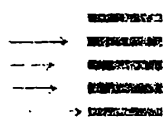
На рис. 13 показана общая схема работы с PROMT Translation Suite 7.0.

¹ Информация о PROMT Translation Suite 7.0, включая иллюстрации, заимствована из документации фирмы, представленной на ее сайте в Интернете.

Результаты перевода можно сохранять в базе и использовать в будущем, что значительно увеличивает скорость работы. По мере заполнения базы эффективность работы с системой повышается — основная часть переводов будет подставляться из базы, а от пользователя потребуются лишь корректировка небольших фрагментов текста, сделанных с помощью машинного перевода.

Технология машинного перевода используется в программе для получения автоматического перевода тех фрагментов текста, аналогии которых в базе отсутствуют.

Этапы перевода документа в PROMT Translation Suite 7.0



1. Сегментация текста. Исходный текст разбивается на сегменты (фрагменты) в соответствии с правилами сегментации.

2. Перевод сегментов текста.

А. База переводов.



75%

При обработке нового текста, поступившего на перевод, система сравнивает его с имеющимися в базе Translation Memory сегментами. При обнаружении идентичного или схожего сегмента его перевод отображается с указанием совпадения в процентах. Сегменты, в которых процент совпадения ниже указанного пользователем, автоматически отправляются на машинный перевод.

Б. Машинный перевод.



Любой сегмент исходного текста можно перевести с помощью технологии машинного перевода. Это может потребоваться в двух случаях:

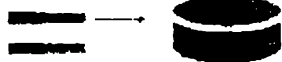
- если процент совпадения между сегментом исходного текста и сегментом из базы меньше заданного;
- если перевод сегмента отсутствует в базе.

В. Ручной перевод.



Перевод любого сегмента текста в программе может быть выполнен пользователем вручную.

3. Пополнение базы переводов.



Проверенные пользователем пары сегментов (т. е. оригинальный фрагмент текста и его перевод) рекомендуется сохранять в базе для дальнейшего использования.

4. Результат перевода.



Проверенные и отредактированные пользователем сегменты объединяются в связный текст перевода с полным сохранением исходного форматирования.

4.5.2. Практикум перевода с использованием системы PROMT

Предлагаемый практикум имеет характер испытания возможностей системы МП¹.

Для эксперимента выбраны четыре текста различной тематической направленности. При отборе учитывались принципиальные возможности системы МП. Одна из главных задач грамотного пользователя состоит в том, чтобы отобрать только те тексты, при переводе которых использование этой технологии оправдано и рентабельно.

Оригинальные английские тексты предварительно не упрощались, и полученный русский машинный перевод не редактировался.

Нередактированные машинные переводы с английского языка на русский были проанализированы по фразам несколькими экспертами и распределены по следующим четырем категориям:

1. Полностью понятные по содержанию и не требующие коррекции либо требующие только легкой стилистической правки.

2. Понятные по содержанию, но требующие синтаксической и другой правки.

3. Понимаемые с большим трудом и требующие большой синтаксической, семантической и другой правки.

4. Совершенно не понятные и не поддающиеся никакой правке.

После установки системы МП PROMT следует выяснить, можно ли ее использовать как самостоятельный лингвистический процессор с интерфейсом, представленным на рис. 14, или надо совместить с текстовым редактором Word. На рис. 15 показано, как выглядит интерфейс текстового редактора Word с установленной системой МП PROMT (обратите внимание на нижний ряд пиктограмм).

При автономном использовании системы перевода рабочий экран автоматически разбивается на два окна — для оригинала и для перевода. У пользователя есть возможность редактировать тексты и в окне перевода, и в окне оригинала. Рабочий экран по желанию может быть разбит на параллельные окна (вертикальное

¹ Идеология и методика этого практикума разработана директором Всероссийского центра переводов научно-технической литературы и документации, профессором И. И. Убиным.

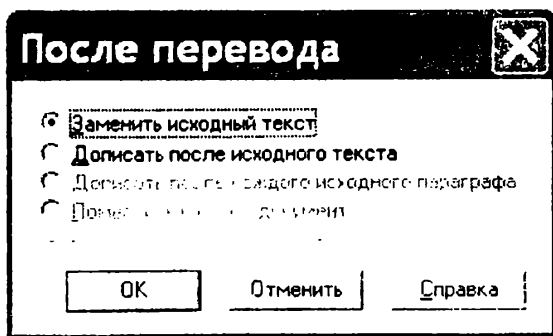


Рис. 18

способы вывода результатов перевода. На рис. 18 показано меню для выбора варианта оформления результатов перевода.

Режим фрагментарного перевода с заменой отрезков оригинала соответствующими абзацами перевода в среде текстового редактора Word показан на рис. 19.

В любом случае пользователь получает обычный текстовый файл, пригодный для всех операций редактирования. Следует подчеркнуть, что в переводе полностью сохраняется формат оригинала: абзацы, шрифты, выделения и другие особенности.

Система PROMT — сложнейший интеллектуальный лингвистический инструмент, выполняющий многочисленные операции. Для правильного выбора назначения этих операций, конечно, требуется навык. Для развития такого навыка в комплект поставки системы входит очень подробный хорошо иллюстрированный учебник объемом более 300 страниц.

Установив систему на компьютер, мы продолжаем эксперимент по оценке принципиальных возможностей современных систем МП.

На экране (рис. 20) выведен большой фрагмент письма и его перевод на русский язык без какого-либо редактирования. Такое

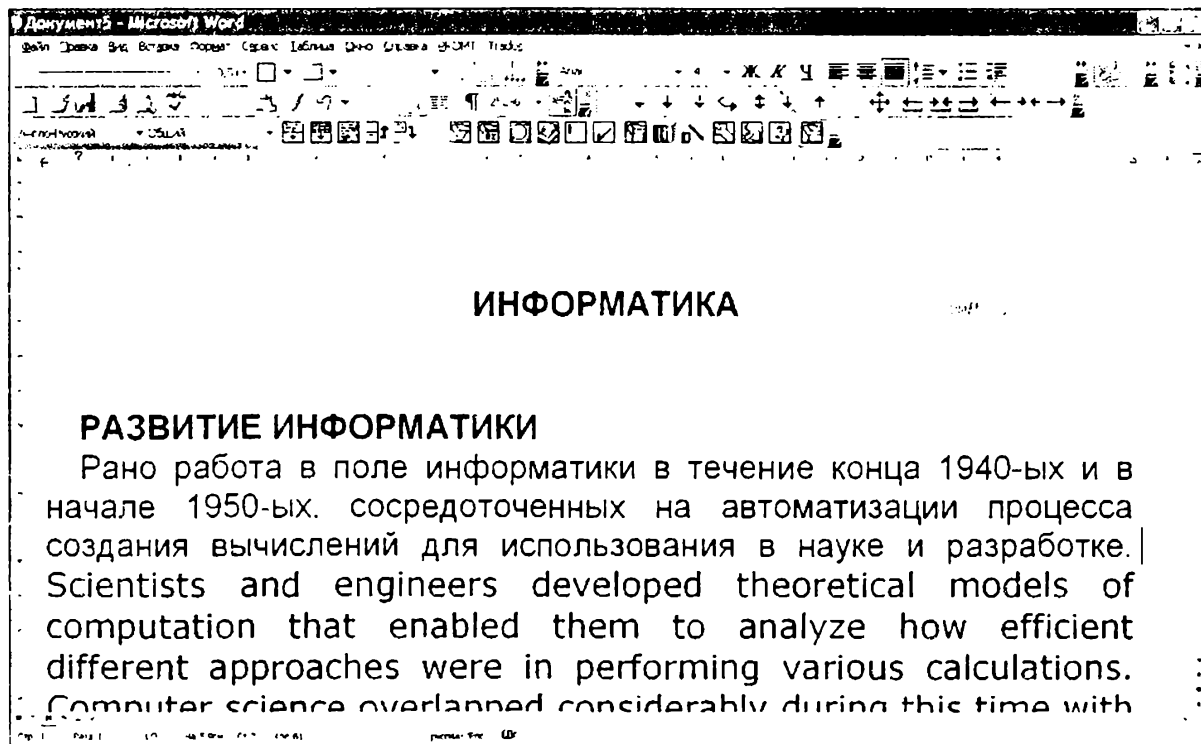


Рис. 19

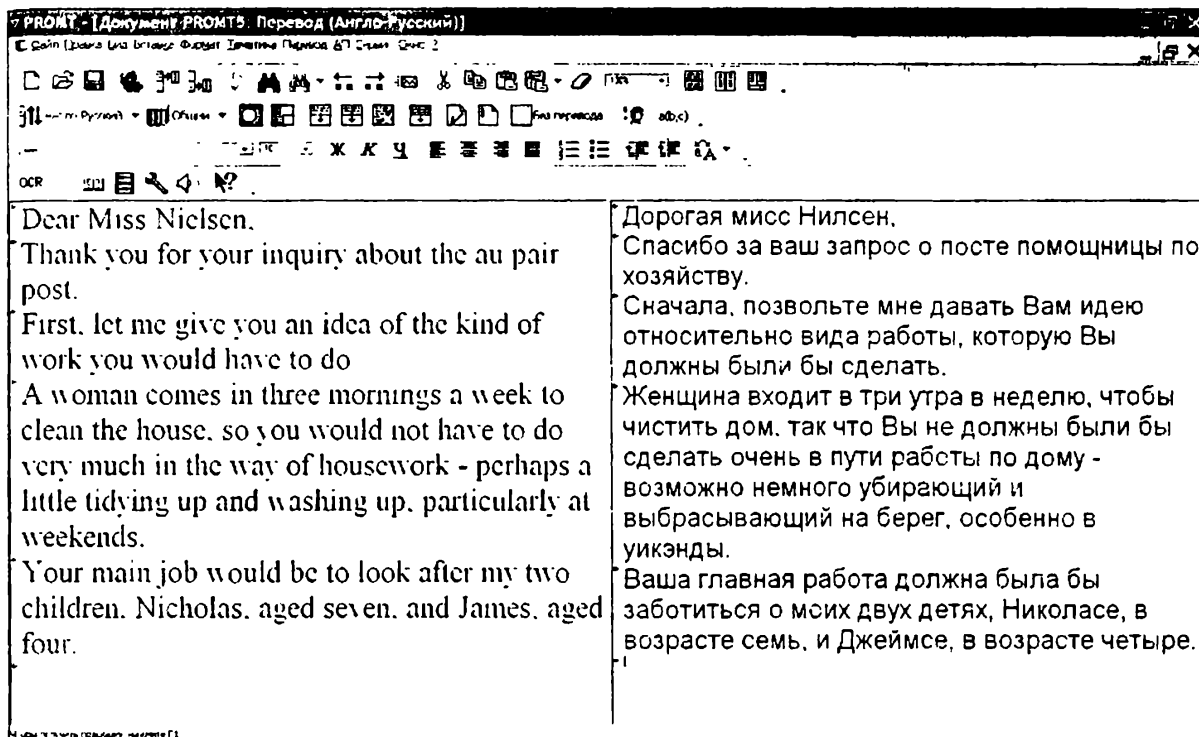


Рис. 20

письмо вполне могло быть отправлено по электронной почте и его перевод удобнее всего было бы выполнить с помощью системы МП. Безусловно, неточности можно заметить в каждом предложении, но общий смысл передан полностью. Следовательно, уже на основе этого перевода вполне может состояться адекватная коммуникация. Мы считаем возможным отнести этот перевод

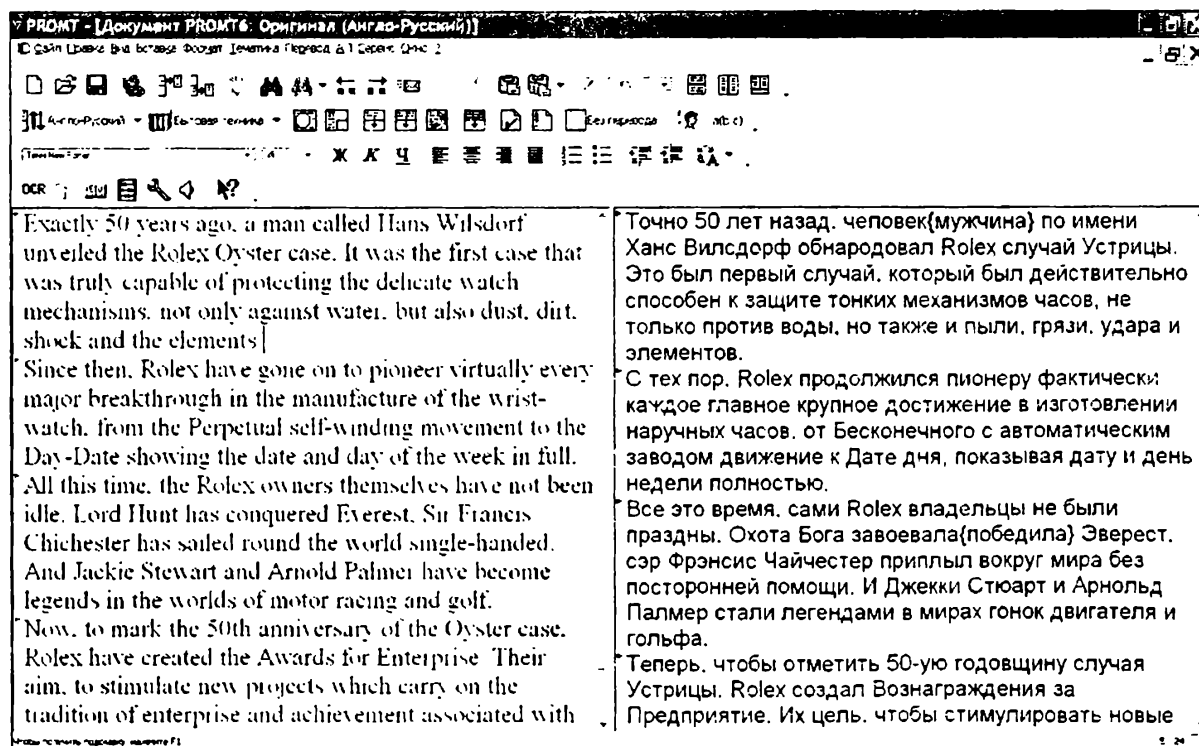


Рис. 21

к первой категории качества, при которой довольно часто требуется уточняющее стилистическое редактирование.

Второй текст (рис. 21) взят из журнала. Это статья «на правах рекламы» и в ней приводятся факты из 50-летней истории знаменитой часовой фирмы «Ролекс». Для перевода мы использовали подборку словарей по бытовой технике.

Получившийся перевод по качеству большинство экспертов отнесли ко второй категории. Смысл понятен, но имеются нарушения синтаксиса и досадная лексическая ошибка, связанная с многозначностью слова *case*. Если бы вместо одного из его значений «случай» было использовано значение «корпус», качество перевода было бы значительно выше. Переводы этой категории в основном требуют лексико-стилистической правки.

Третью категорию качества можно присвоить переводу фрагмента текста из области информатики, представленному на рис. 22 (при его выполнении была использована подборка словарей по информатике). Сразу же видно, что перевод в определенном смысле выхолащивает и упрощает текст оригинала и, безусловно, может быть использован только в информационных целях после значительной лексико-грамматической и стилистической правки.

Тематику текста, представленного на рис. 23, определить довольно сложно: на равных основаниях он может принадлежать и к области медицины, и к общей бытовой проблематике. Его перевод был выполнен с использованием коллекции словарей по медицине. Результаты оказались явно неудовлетворительными. Возникло предположение, что лексические проблемы могли привести к оши-

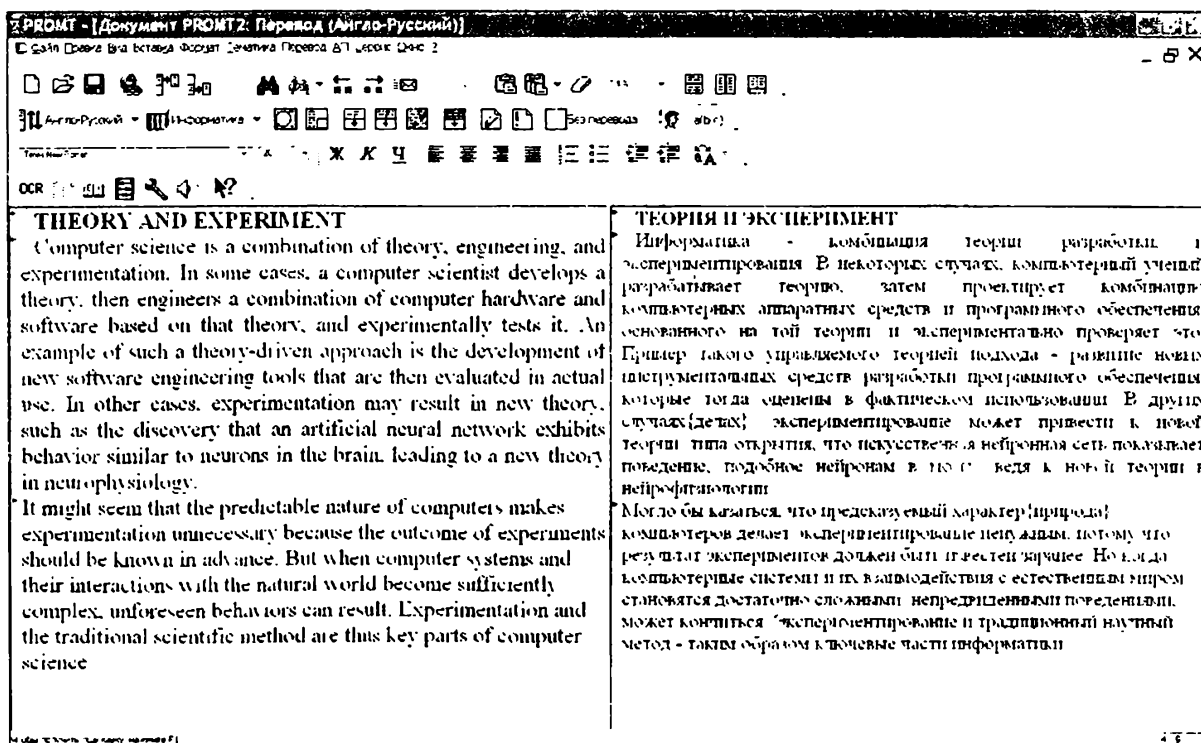


Рис. 22

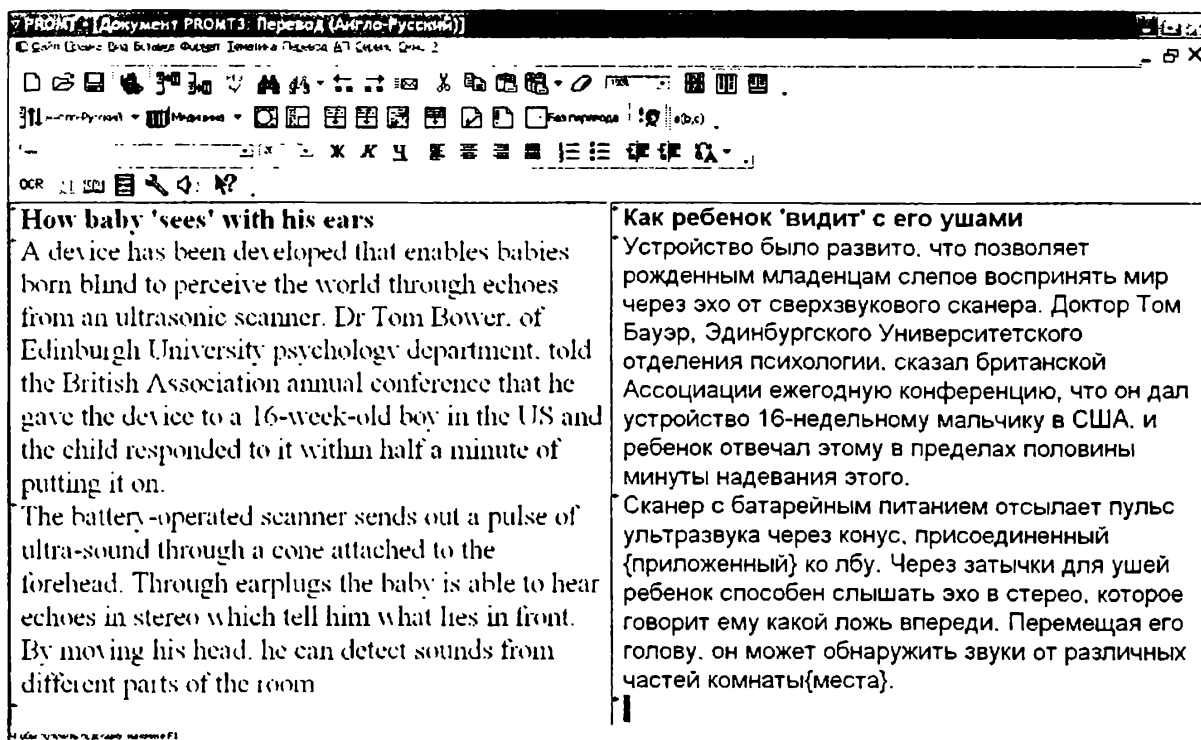


Рис. 23

бочному анализу синтаксиса оригинала и, следовательно, к синтаксическим ошибкам в переводе. Но, к сожалению, повторный перевод с другим комплектом словарей (рис. 24) не привел к сколько-нибудь заметному улучшению качества текста.

Оба перевода приходится считать неудовлетворительными. Но главная ошибка, скорее всего, была допущена не системой, а

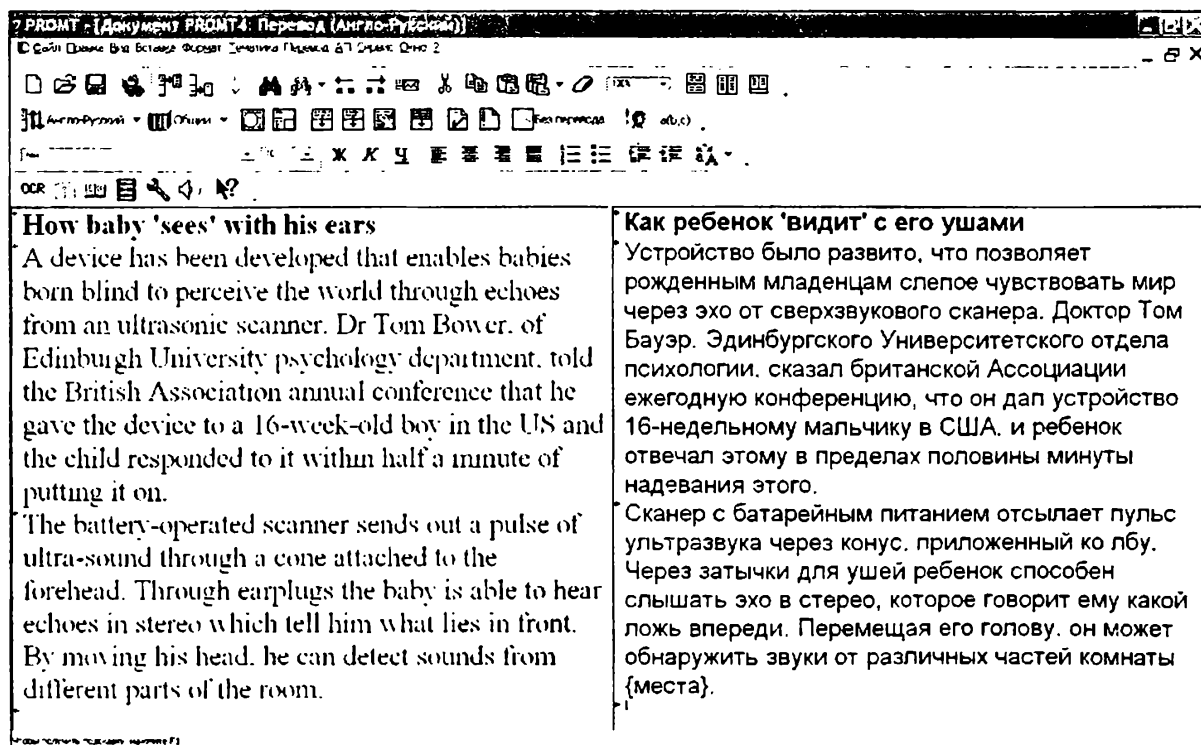


Рис. 24

пользователями. Вряд ли стоило поручать системе МП работу над текстом, касающимся не просто медицинских проблем, а проблем здоровья ребенка.

В общей теории перевода мы обычно исходим из аксиомы, гласящей, что непереводаемых текстов не существует. Это убеждение вытекает из признания интеллектуальной независимости речемыслительного процесса от языка выражения информации. Поскольку современные системы МП еще не обладают искусственным интеллектом, аксиома переводимости не может распространяться на их переводческую деятельность.

Учитывая то, что системы постоянно развиваются, было бы большой ошибкой давать окончательную оценку их возможностей по качеству перевода. Ее словари постоянно пополняются, и нет уверенности в том, что уже сейчас не существует обновленная версия, которая могла бы решить ту или иную из обнаруженных нами проблем.

Тем не менее приведенные примеры все же позволяют высказать некоторые оценочные суждения. Неработавший машинный перевод можно считать рентабельным для информационных целей, если доля фраз, понятных без редактирования (первая и вторая категории качества), составляет не менее 60 %. Перевод с постредктированием считается рентабельным, если суммарная скорость выполнения машинного перевода и его «ручного» постредктирования выше скорости «ручного» перевода этого же текста. Уровень восприятия экспертами качества машинного перевода всех четырех текстов оказался достаточно высоким. Следовательно, даже небольшой эксперимент ясно определяет наиболее рациональную сферу применения машинного перевода — в среде Интернета для информационно-познавательных целей.

Контрольные вопросы и задания

1. Выберите наиболее заинтересовавшую вас онлайн-систему МП и попытайтесь перевести с ее помощью несколько предложений.

2. Назовите три основных типа систем МП и дайте краткую характеристику каждого из них. Попробуйте определить, к какому типу относится онлайн-система МП, которой вы иногда можете пользоваться в Интернете.

3. Каковы особенности оценки машинного перевода по сравнению с традиционным?

4. Назовите основные типы ошибок машинного перевода.

5. Вернитесь к первому заданию и дайте оценку сделанных вами переводов и объясните причины замеченных ошибок.

6. Дайте краткую характеристику системы МП PROMT и назовите ее преимущества по сравнению с системой SYSTRAN.

7. Попробуйте выполнить онлайн-тестовый перевод на web-сайте фирмы «ПРОМТ» и составьте его лингвистическую оценку.

АВТОМАТИЧЕСКИЕ ПЕРЕВОДНЫЕ СЛОВАРИ¹**5.1. Принципы номинативной деятельности**

Человечество решает проблемы номинации предметов и явлений окружающей его действительности уже многие тысячелетия. Эти проблемы сопровождают человека буквально с его первых шагов и решаются им скорее или медленнее на разных этапах развития общественной культуры и взаимоотношений между людьми. И их масштаб постоянно увеличивается, пожалуй, быстрее, чем в геометрической прогрессии.

Лексический состав языка, сформированный и постоянно уточняемый самим человеком, представляет собой наиболее динамичный из всех уровней языка. Под лексическим составом понимается вся совокупность слов и устойчивых словосочетаний, образующих единую лексическую систему языка, обладающую упорядоченной структурой с устойчивыми связями между своими подсистемами и их отдельными элементами. Именно взаимосвязанность всех процессов, происходящих в языке, требует информационно-технологического системного подхода при решении проблем межъязыковой коммуникации.

К основным функциям языка относятся: коммуникативная, благодаря которой он становится средством общения людей; конструктивная, поставляющая способы формирования мыслей; экспрессивная, лежащая в основе выражения и выразительности мыслей, и аккумулятивная, благодаря которой реализуется возможность накопления и хранения знаний и опыта. Выполнение каждой из этих функций связано с использованием лексического состава языка. Количественно и качественно этот состав постоянно меняется и зависит от многих условий, среди которых важное место занимает уровень развития общества и общественных взаимоотношений.

И содержательные, и количественные характеристики лексического состава, необходимого для установления различных отношений, зависящие от особенностей языковой коммуникации, определяются прежде всего языковыми возможностями участников коммуникации, а также содержанием побудительных мотивов коммуникации — задачами коммуникативного взаимодействия людей в процессе их совместной деятельности. Это выводит ком-

¹ Эта глава написана в соавторстве с И. И. Убиным.

муникативную функцию на первый план среди других функций языка.

С технологической точки зрения межъязыковой коммуникативный процесс можно в общих чертах рассматривать как прямое воздействие на лексико-грамматические трансформации. Как любой аспект естественного процесса мотивированного взаимодействия людей, лексическое обеспечение языковой коммуникации может объективно оцениваться категориями количества и качества. Мы попытаемся проанализировать количественные и качественные характеристики лексического состава, обеспечивающие коммуникацию в сфере науки и техники.

Исследования по этому вопросу в настоящее время ведутся очень интенсивно, но нерешенных проблем не становится меньше. Даже не проводя какого-либо статистического анализа, можно быть абсолютно уверенным, что из трех аспектов языка — звукового состава, лексического состава и грамматических средств — объектом обсуждений и публикаций чаще бывают лексические аспекты, среди которых особенно выделяются терминологические проблемы. Поэтому любые новые технологии в языке касаются прежде всего его лексического уровня. Первые лингвистические труды в истории человечества были посвящены лексике. Имеются сведения, что словари имелись уже в XV в. до н. э., а основы восточной лексикографии были заложены словарем китайских иероглифов «Эръя» в III в. до н. э.

На фоне традиционной лексикографической деятельности общества в XX в. наиболее динамичное развитие получила терминологическая лексикография. Этот процесс начался сравнительно недавно и в его становлении имеется еще немало резервов.

Современная терминологическая школа в нашей стране была основана первыми трудами в этой области С. А. Чаплыгина и Д. С. Лотте и продолжается вплотную приближенными к прагматическим аспектам лексико- и терминографическими работами И. И. Убина (соавтора этой главы) и К. Я. Авербуха. За рубежом терминологическая школа обозначилась трудами Э. Вюстера. Важным организационно-методическим отличием русской терминологической школы является привлечение к систематологической работе в равной мере терминологов-специалистов в систематизируемой предметной области и терминологов-методистов, организующих и создающих информационные технологии лингвистического обеспечения терминологической работы. И в нашей стране и за рубежом этой работе придается большое социально-экономическое значение. Языковая информация о продукции, выпускаемой на рынок, уже стала неотъемлемой комплектующей частью этой продукции.

Прежде всего информационные технологии, используемые в терминологической работе, направлены на обслуживание пере-

вода — важнейшего аспекта координирования деятельности различных специалистов, от которой во многом зависит развитие предметной области. Одним из параметров информационно-технологического аспекта перевода является сопряжение двуязычной терминологической номинации соответствующих референтов. Такое сопряжение, да и сама по себе номинация являются результатом сближения информации о действующем предмете, которой владеют переводчики и разноязычные специалисты, вступающие в коммуникацию.

Наиболее заметные изменения в языке особенно часто происходят в периоды активизации общественной деятельности людей¹. Общественная деятельность человека всегда направлена на видоизменение к лучшему его экономического положения, на совершенствование сферы производства. Активизация предметных областей, естественно, ведет к активизации коммуникативной деятельности. И наконец, такая активизация достигает столь высокой мощности, что требует новых технологий коммуникации. Отставание развития некоторых предметных областей в той или иной стране по сравнению с развитием мировой технологии стимулирует межъязыковые процессы, связанные с появлением новой терминологии, а следовательно, и с необходимостью ее упорядочения. Изменение основ экономического положения общества также вводит новые понятия, номинация которых — процесс актуальный и неотложный.

Наряду с разработкой частных вопросов анализа и упорядочения имеющейся и вводимой терминологии научная жизнь общества требует решения обобщенных задач универсального характера. Но и последние могут быть решены лишь в тесной взаимосвязи с проблемами конкретных предметных областей. От информационных работников требуется не только выработка информационно-технологических рекомендаций, но и их практическая реализация. В естественных языках образование и принятие в «свою семью» новых лексических единиц — процесс исключительно медленный. Вообще такие явления, как возникновение новых слов, крайне редки. Примерами словообразования в английском языке могут быть слова *iff* (математический термин «тогда и только тогда», синоним *if and only if*) или *beatle* (робототехнический термин «такт выпуска», синоним *throughput rate*). Наибольший интерес вызывает расширение многозначности — внутренний процесс, происходящий в лексическом составе языка и представляющий большую проблему для всех современных информационных технологий.

¹ По этому вопросу советуем почитать научные, но увлекательно написанные книги А.Н.Селищева (Язык революционной эпохи. — М., 1928) и П.Лафарга (Язык и революция. — М., 1930).

Переход к новой идеологии использования компьютера в текстовой деятельности обусловлен целым комплексом социальных, лингвистических и технических факторов, которые сложились в середине 60-х гг. прошедшего столетия. Первое — быстрый рост объема научно-технической и деловой документации на разных языках и необходимость оперативного перевода значительной ее части; второе — невозможность выполнить эту работу традиционными способами, т. е. силами человека-переводчика в сжатые сроки; третье — неспособность созданных к тому времени систем МП обработать этот огромный поток информации; четвертое — возросшие возможности вычислительной техники. Влияние последнего фактора на развитие машинного перевода и вычислительной лексикографии некоторыми лингвистами недооценивается. Но если при теоретических исследованиях лингвисты действительно редко учитывают реальные возможности вычислительной техники, то выбор оптимальных решений конкретных лингвистических задач в рамках промышленных систем МП и автоматических словарей в значительной степени зависит от уровня развития вычислительной техники. Понимание этого проявилось в середине 60-х гг., когда возросшие возможности ЭВМ позволили создать автоматические переводные словари объемом в десятки, а затем и в сотни тысяч лексических единиц, что было абсолютно невозможно в 50-х гг., когда только начинались работы по машинному переводу.

Развитие вычислительной техники оказывало постоянное влияние на развитие вычислительной лексикографии. Так, создание разнообразных, простых в обращении и дешевых микроЭВМ с достаточно большими возможностями привело к появлению нового направления в разработке автоматических переводных словарей в помощь человеку — словарей индивидуального пользования, так называемых автоматических рабочих мест переводчика.

Созданию автоматических переводных словарей предшествовали исследования по определению вида и количества ошибок, совершаемых переводчиком при переводе текстов различной трудности, и времени, которое он затратит на поиск в словарях и справочниках незнакомых ему слов. Для анализа ошибок были взяты 72 перевода по научно-технической тематике с английского языка на немецкий различной степени сложности, общим объемом 80 страниц, выполненные 55 переводчиками с использованием печатных словарей и справочников. Переводы были разделены по трем степеням трудности (каждой степени соответствовала треть всего исследуемого материала):

1 — относительно простые тексты — популярные научные газетные статьи или несложные технические инструкции по эксплуатации;

2 — технические тексты средней сложности;

3 — сложные научно-технические тексты о новых теориях и направлениях, о которых в немецкой литературе имеется мало сведений и которые содержат терминологию, не зафиксированную в существующих специальных словарях.

Тексты переводов были отредактированы 26 опытными редакторами. По результатам их работы исследователи выделили три основные категории ошибок:

- ошибки в немецком языке (орфографические, грамматические, знаки препинания);
- переводческие ошибки (лексические ошибки, пропуски, контекстуальные неточности);
- стилистические ошибки.

Опустим подробный анализ различных видов ошибок и приведем только данные по лексическим ошибкам, т. е. в выборе переводного эквивалента. В табл. 2 показано процентное соотношение лексических ошибок в рассматриваемых примерах научно-технического перевода.

Итак, лексические ошибки составили более 50 % всех ошибок, при этом их количество возрастало с повышением сложности текста оригинала. Следует учитывать и тот факт, что терминология по современным отраслям науки и техники расширяется и обновляется столь быстрыми темпами, что ни один словарь, издаваемый традиционным способом, не в состоянии фиксировать ее оперативно. При таком положении дел переводчик вынужден все больше времени затрачивать на поиск незнакомой ему лексики в различных справочниках, пособиях и номенклатурных списках.

Для того чтобы определить среднее время, которое переводчик тратит на поиск незнакомой ему технической лексики, был проведен специальный эксперимент. Двум переводчикам был дан один и тот же научный текст для перевода с английского языка на немецкий. Оба они не были специалистами в той области, к которой относился этот текст. Переводчик А работал традиционным способом, т. е. пользовался обычными переводными словарями и справочниками, которые мог найти в библиотеках, а также консультировался со специалистами. Переводчик Б получил специальный словарь, в который были включены все технические тер-

Таблица 2

Степень трудности перевода	Общепотребительные слова	Термины	Всего (от общего числа ошибок разного типа)
1	39,5	9,9	49,4
2	26,0	28,3	54,3
3	21,9	40,2	62,1

мины, встретившиеся в предложенном для перевода тексте. Английские термины с их немецкими эквивалентами были даны в этом словаре в том порядке, в котором они появлялись в тексте. Чтобы уменьшить влияние фактора квалификации переводчика, при повторном эксперименте возможность пользоваться специальным словарем была предоставлена переводчику А, а переводчик Б работал традиционным способом. Эксперимент был повторен с участием двух других переводчиков и новым текстом. Результаты показали, что переводчик, работающий традиционным способом, тратит в среднем на 66 % больше времени, чем переводчик, использующий специально подготовленный словарь.

Исследователи нередко обращают внимание и на то, что количество ошибок в переводах, выполненных с помощью списков терминов с их переводными эквивалентами, на треть меньше, чем в переводах, выполненных традиционным методом.

Для объективности картины необходимо изложить и другую точку зрения. Некоторые профессиональные редакторы и переводчики считают, что опытный и высококвалифицированный переводчик, работая с текстами хорошо знакомой тематики, весьма редко обращается к словарям, и поэтому автоматический переводной словарь не будет представлять для него практической ценности. Не отвергая мысль о том, что чем выше квалификация переводчика, тем реже он обращается к словарям, мы должны указать, что, во-первых, таких переводчиков немного и не они выполняют основную часть научно-технических переводов. Большая часть научно-технических переводов делается переводчиками средней квалификации либо техническими специалистами, знающими иностранный язык. Во-вторых, даже переводчики высокой квалификации испытывают трудности, встречая в работе новые, еще не зафиксированные в традиционных словарях термины. В силу своей динамичности автоматические словари имеют возможность значительно раньше, чем традиционные «бумажные» словари, фиксировать новую терминологию, и именно в этом аспекте они будут или, по крайней мере, могут быть полезны и профессионалам перевода высокой квалификации.

5.2. Формирование словаря

Основным источником пополнения лексического состава языка являются устойчивые словосочетания. Одну из главных составляющих процесса формирования устойчивых словосочетаний как раз и представляет бурно развивающееся терминообразование. Многие исследователи доказывают в своих работах, что терминообразование и терминологизация общей лексики служат мощным источником пополнения словарного состава языка.

При подходе к пониманию сути проблемы терминологизации коммуникативных процессов воспользуемся интерпретацией понятия «термин» как особого слова, имеющего специальное значение и выполняющего целевую функцию. В плане выражения термин имеет вид лексической единицы, сочетающейся со всеми аспектами языка: обладает звуковым составом, продиктованным традициями конкретного языка, подчиняется его грамматическим закономерностям. Вместе с тем термин действительно является особым словом: в терминологической среде почти исключена полисемия, ослаблена синонимия; термин не может иметь экспрессивно-стилистической окраски. Все эти свойства оказывают большое влияние на технологию анализа и сопоставления терминов в различных языках.

Доказанность двойственности положения терминов очень важна для методологии исследования процессов терминообразования, для системности терминологии и функционирования терминосистем. Понимание термина как особого слова требует и особого подхода к его оценке, интерпретации и информационной систематизации.

Из всех отличий наиболее значимым является признак системности, которым термин обладает в наибольшей степени, чем любые другие слои лексики. Вследствие системности окружающего мира нет и не может быть такого понятия и выражающего его термина, которое не было бы системно связано с другими понятиями данной предметной области. Каждый термин существует лишь как член определенной системы терминов.

Одновременно он имеет свое место в лексико-семантической системе языка и с позиций лексикологии должен интерпретироваться как единица логоса и как единица лексиса, т.е. занимая свое единственно возможное место в терминологической системе конкретной предметной области, термин интерпретируется как единица логоса.

Приведенные выше положения отражают суть исследований многих ученых, ведущих работы по изучению проблем соотношения и сосуществования общей лексики и терминологии. Всеми этими работами создана прочная основа для доказательства объективности существования самостоятельной науки — терминоведения, имеющей свой предмет, структуру и методологию.

В настоящее время для упорядочения терминологий разных областей знаний задействован научный потенциал многих исследовательских центров. Диалектический подход к пониманию базиса терминологических процессов в языке требует построения логической последовательности восприятия действительности от системы областей знаний к системам понятий, а от систем понятий через номинативные процессы к системам терминов.

Факт существования терминосистем является результатом естественного процесса, который можно было бы описать в терминах информационных технологий. Развитие логической последовательности, характерной для информационной технологии, определяется коммуникативным процессом, который таким образом становится одной из форм существования технологично построенной терминосистемы.

До принятия технологического подхода положение лингвистов-терминологов было слабо ориентировано в информационно-логической последовательности, и им приходилось чаще иметь дело с уже существующим результатом — естественным образом сложившейся терминосистемой.

Что же касается системы области знаний, которая имеет ярко выраженный динамический характер, то исследование ее терминологии явно отстает от развития самой этой системы. Объективной причиной этого является сильная интерференция предметных областей. Субъективная причина кроется в явном отставании в развитии потенциала науковедения от его методологии.

Исследовательским аргументом для лингвиста-терминолога остается и широко используется сложившееся условное разделение наук на теоретические (фундаментальные) и технические (прикладные). Слабая научная дифференциация предметных областей и отсутствие системного описания их границ приводят к неразличению научной и технической терминологии. Вместе с тем объективное отдельное наличие и той, и другой в пределах одной предметной области не отрицает понятия «научно-техническая терминология». Оно реально существует, но, являясь комплексным, имеет два плана содержания.

На фоне известных из литературных источников исследований, связанных с терминологией и терминосистемами, проводимых и в нашей стране, и за рубежом, сравнительно мало работ, построенных на изучении объективно сформировавшихся терминосистем с позиции информационно-технологического системного подхода. Проведено много успешных исследований структур терминосистем конкретных предметных областей, исследований функционирования их элементов — терминов. Но следовало бы ожидать бóльших результатов именно от системного подхода к изучению функционирования терминосистемы как единого целого, наделенного внутренними процессами.

Исследование терминосистемы с позиций понимания ее системных свойств одновременно дает возможность изучения терминов (терминологии) как части в целом и как элемента в среде. Влияние терминосистемы как целого на свойства термина как ее элемента достаточно сильно проявляется при номинации новых явлений. Системный подход открывает возможности активно использовать в практических целях синергетические процессы, про-

исходящие внутри терминосистемы. Большое значение для продуктивного употребления терминологии имеет также учет четко определенных относительных границ всей этой системы.

Системный подход к исследованию функционирования терминосистемы позволяет результативно экстраполировать на терминологические аспекты языковой коммуникации опыт лингвистической интерпретации соотношения системы и среды, накопленный в современных информационных структурах. Изучение терминосистемы без определения особенностей среды, в которой она функционирует, сужает рамки обзора и понимания ее влияния, маскирует многие ее возможности, приводит к ошибочным выводам о процессах интерференции, происходящих с особой интенсивностью в современных развивающихся предметных областях. Определение и знание особенностей среды функционирования терминологии и терминосистемы открывают возможность наиболее полно представлять функциональные связи на разных уровнях функционирования терминологии в различных предметных областях.

К сожалению, хотя уже имеются фундаментальные исследования в области терминоведения, остается неразработанным понятие границ терминосистемы при представлении основ этой науки, ее предмета и структуры.

Отсутствие четко очерченного значения этого понятия приводит к потере понимания и реализации значительной части функциональных возможностей, которыми потенциально располагает терминосистема, как вообще любая система в ее целостности. Было бы целесообразно исходить из общепризнанного представления об относительной границе как о признаке целостности. Именно такая позиция действительно дает основания для системного подхода при определении свойств, связей, среды существования, внутренней структуры и синергетических процессов.

Востребованность терминосистемы при коммуникации происходит в двух ситуациях, определяемых задачами коммуникативного процесса. Обращение к терминосистеме осуществляется в условиях одноэтапного синтеза нового текста и сложной аналитико-синтетической обработки уже имеющегося текста. Ситуация аналитико-синтетической обработки текста изучена меньше и этим она привлекает к себе особое внимание. О необходимости исследовать связи терминосистемы и процесса аналитико-синтетической обработки текста свидетельствуют некоторые исследовательские материалы, накопленные в переводческой и редакторской деятельности, имеющей самое непосредственное отношение к аналитико-синтетическому процессу.

Описывая историю появления первых автоматических словарей, Ф. Кроллманн¹ упоминает, что автоматический переводной

¹ Цит. по: Убин И.И. ЭВМ и словарь. — М., 1992. — С. 16.

словарь ЛЕКСИС был задуман сразу как многоязычный англо-французско-русско-немецкий словарь. В качестве лексической базы использовался терминологический фонд Службы переводов ФРГ, который к началу 60-х гг. XX в. насчитывал уже около 300 тыс. терминов по нескольким тематическим областям с эквивалентами на разных языках. Первая версия словаря предназначалась для эксплуатации только в режиме пакетной обработки информации. В дальнейшем словарь значительно расширил свои функции и получил возможность работать также в режиме диалога через дисплей. Этот режим чаще используется при выполнении срочных переводов, в обычных условиях переводчики стараются его избегать из-за высокой стоимости. Наибольший интерес режим диалога представляет для терминологов и лексикографов, которые с помощью дисплея могут оперативно пополнять и перестраивать основные массивы словаря.

Пакетная обработка информации бывает необходима при составлении текстоориентированных микрословарей на бумажном носителе. Словарные статьи в таких словарях могут размещаться либо в порядке появления запрошенных терминов в тексте, либо в алфавитном порядке. Пакетный режим обработки информации значительно дешевле режима диалога.

Текстоориентированные словари на бумажном носителе позволяют прибегать к услугам автоматического словаря как отдельным переводчикам, не имеющим возможности пользоваться режимом диалога, так и небольшим переводческим коллективам, прикрепленным к Федеральному бюро языков. Другое преимущество этих словарей заключалось в том, что непосредственно на их полях можно записывать исправления, добавления и другие комментарии к словарной статье. Подобная информация весьма полезна при модернизации основных массивов автоматического словаря, что происходит постоянно. Недостаток текстоориентированных словарей состоит в некотором разрыве по времени между составлением запроса на него и его получением.

Большая часть массива статей словаря ЛЕКСИС содержит однозначные термины и терминологические словосочетания. При этом разработчики считают, что основное внимание следует уделять словосочетаниям, так как именно к ним относится примерно 80 % запросов переводчиков. Иноязычные слова и словосочетания и их немецкие переводные эквиваленты даются в словарной статье в исходной форме. Максимальная длина каждой иноязычной части словарной статьи и немецкого переводного эквивалента равна 70 символам (по 35 символов на каждый язык). Для экономии памяти ЭВМ и ускорения поиска осуществлено сжатие информации. Омографы внутри одной части речи записываются один раз, но каждый немецкий переводной эквивалент сопровождается индексом той отрасли науки и техники, в которой он

употребляется. Подобным же образом решается проблема многозначности. Разные значения слова получают код той области техники, к которой они относятся. В словаре фиксируются все орфографические варианты иностранной, в частности английской, лексики. Дефисы в словосочетаниях заменены пробелами.

В первой версии этого автоматического словаря использовался метод ручного формирования запроса. Пользователь сам составлял запрос, подчеркивая в иностранном тексте термины, эквиваленты которых он желал получить. Он же производил разделение словосочетаний, осмысливал, что в выражении является постоянным, а что может быть заменено. Переход от словоформ к исходным формам на этом этапе осуществлялся также вручную. В дальнейшем разработчики ввели элементы частичной автоматизации.

По желанию пользователя ему выдаются либо переводные эквиваленты только указанных им областей науки и техники, либо вся словарная статья и он сам выбирает нужный эквивалент. Не обнаруженные в словаре термины поступают в блок накопления, затем передаются терминологам для работы над ними.

Автоматический словарь ЛЕКСИС показал себя надежным и рентабельным средством решения многих проблем, которые возникают в работе переводчиков и лексикографов. За время, прошедшее с момента его создания, словарь естественно развивался в направлении как увеличения количества языков и объемов лексических массивов, так и повышения оперативности и улучшения эксплуатационных характеристик, особенно в части диалога с человеком.

5.3. Становление компьютерной лексикографии

Изучение проблем терминологии в нашей стране ведется уже более полувека. Однако нельзя не отметить сильную разобщенность научной и практической деятельности видных терминологов и отсутствие единого центра, координирующего всю эту работу. В результате не найдены общие подходы даже в толковании основополагающих понятий этого научного направления.

Рассматривая период становления компьютерной лексикографии, И. И. Убин¹ обращает внимание на тот факт, что вслед за крупными государственными и международными переводческими организациями на автоматические словари обратили внимание и частные национальные и межнациональные фирмы. Так, например, фирмой «Сименс» уже в течение многих лет эксплуатируется автоматический переводной словарь TEAM (Terminologie-Erfassungs- und Auswertungs-Methode), главное назначение кото-

¹ См.: Убин И. И. ЭВМ и словарь. — М., 1992. — С. 21.

рого состоит в том, чтобы освободить переводчика от утомительной работы по поиску слов в обычных словарях. Основной частью ТЕАМ является банк терминологических данных, который позволяет получать:

- технические словари на магнитных дисках;
- технические словари по определенным темам или текстоориентированные словари, необходимые для переводов больших текстовых объемов;
- отдельные словарные статьи на экране дисплея в режиме диалога «человек — ЭВМ».

В принципе словарная статья может содержать термины на восьми языках — немецком, английском, французском, испанском, русском, итальянском, португальском и датском. Однако практически немногие статьи имеют эквиваленты на всех восьми языках. В тех случаях, когда в каком-либо языке термин не используется, то вместо точного переводного эквивалента дается толкование его значения. Термин, включенный в банк данных, может состоять как из одного, так и из нескольких слов. Помимо переводных эквивалентов термина на восьми языках словарная статья может иметь дополнительную информацию, которая содержит толкование термина, контекст, источник, дату поступления в банк и индекс качества (надежности). Для некоторых слов указаны также грамматический класс и род. Для решения проблемы омонимии вводится код тематики. Дополнительная информация и полный набор переводных эквивалентов выдаются только при специальном запросе пользователя. Максимальная длина словарной статьи — 3 тыс. символов. Все входные единицы заносятся в словарь в своей исходной форме. Для терминов, имеющих сокращения, фиксируются обе формы — полная и сокращенная. Различные орфографические варианты либо нормируются и сводятся к одному, либо все заносятся в словарь, как, например, в случае с английским и американским вариантами написания некоторых терминов. Русская часть словаря транскрибируется латинскими буквами на основе чешского алфавита (с учетом дополнительных символов).

Помимо одно- и многословных терминов в автоматический переводной словарь включаются также различные свободные и фразеологические словосочетания, например такие глагольные словосочетания, как *to load a register*, *to link a programme*, *to set a switch*, или даже целые предложения: *a programme generates an error report*. Обработка и поиск подобных сложных единиц и их иноязычных переводных эквивалентов осуществляются на общих основаниях. В составе такого сложного терминологического или фразеологического сочетания выделяется ключевое слово, по которому и осуществляется поиск этой единицы. Ключевые слова в исходной форме сведены в отдельные списки. Как показала прак-

тика, выделение ключевых слов и другие преобразования исходного языкового материала выгоднее осуществлять вручную, а не автоматически.

Запрос представляет собой отдельное слово или словосочетание. В режиме диалога перед вводом в ЭВМ оно высвечивается на экране дисплея, чтобы пользователь мог устранить возможные ошибки в запросе. При этом надо указать язык запроса и язык (один или несколько) ответа. Выдав на запрос первый ответ (переводной эквивалент), система спрашивает, нужно ли выдавать второй и т. д. Предусмотрена также пакетная обработка запросов и выдача ответов в виде микрословарей либо на бумажном, либо на микроносителе.

* * *

Одним из основных тезисов в теоретической дискуссии по вопросу «что есть термин?» является понимание и трактовка соотношения «лексиса» и «логоса». Факт принадлежности термина к «лексису» очевиден, достаточно хорошо изучен и обсужден в лингвистической литературе.

Определение термина как «логоса» исходит из его связи с понятием, но многими теоретиками эта связь определяется неоднозначно.

В трактовке вопроса о связи термина с понятием сложились два основных направления. Одно базируется на номинативной тенденции термина, т. е. на его связи с «вещью». Другое исходит из связанности термина со смысловым содержанием понятия.

При достаточно высокой условности разграничения понятий «наука» и «техника» оно все же существует и отражено в трудах теоретиков и практиков перевода, что не может не влиять на дальнейшие исследования в области перевода и терминологии.

Совершенно очевидно, что в терминосистемах подязыков техники связь термина с «вещью» доминирует. Определения и понятия техники и технологий более овеществлены, формализованы и стабильны. Перечень составляющих их «вещей» конечен, относительно мало изменяем даже в развивающихся предметных областях.

Овеществленность не может существовать абстрактно, она должна иметь конкретное воплощение в каком-либо объекте, через который она осознается. Говоря об овеществленности предметных областей техники, обычно подразумевают всю совокупность материальных объектов, каждый из которых в его взаимосвязи с другими представляет собой часть материального мира. С точки зрения терминологии материальный объект — это понятие, соответствующее предмету действительности («вещи»), отражаемому в сознании, представленному в языке и имеющему ту или иную семантическую и синтаксическую функцию.

Реальность предлагает разнообразные классификации материальных объектов. Однако эти объекты и их классификации отражаются не непосредственно в языке, а в сознании, которое фиксирует это отражение в конвенциональных знаках, т. е. между реальной вещью и ее языковым представлением находится отраженный в сознании образ этой вещи, некое усредненное ее представление. В связи с этим используется понятие интериоризации — «переход, в результате которого внешние по своей форме процессы с внешними же, вещественными предметами преобразуются в процессы, протекающие в умственном плане, в плане сознания»¹. В подъязыках науки, и особенно в подъязыках науки социальной и экономической, определяющую роль играет связь термина с абстрактным понятием. Здесь осуществляется постоянное активное терминообразование, и ведущая роль в коммуникативном процессе отводится термину с глубинным планом содержания, расширенным универсализированным толкованием и заметной долей абстракции.

С точки зрения «возможных миров» даже нематериальные понятия можно представить вполне конкретными овеществленными объектами. Уже почти около 100 лет назад в одной из своих основополагающих работ Г. О. Винокур писал, что «специальные языки отличаются один от другого... преимущественно — если не исключительно — своим словарным составом»². А это значит, что все специфические особенности подъязыка — это особенности его терминологии.

Дальнейшее акцентирование внимания на разграничении связей термина с «вещью» или с абстрактным понятием принадлежит А. А. Реформатскому: «терминология, прежде всего, связана с системой понятий данной науки, номенклатура же лишь этикеттирует ее объекты»³. Но в последующих работах А. А. Реформатский уже не ставил вопрос о разграничении номенклатуры и терминологии столь категорично, хотя многие его последователи продолжают серьезно изучать эту проблему.

При практическом подходе к этому вопросу можно заметить некоторые невыраженные закономерности деления. Например, терминология подъязыка предметной области робототехники обладает более номинативной направленностью. Здесь преобладают детали конструкций, узлы машин, объекты обработки, средства автоматизации и другие материальные объекты. Статистический анализ, проведенный на материалах русского и английского языков,

¹ Лсонтьев А. А. Деятельность, сознание, личность. — М., 1975. — С. 95.

² Винокур Г. О. О некоторых явлениях словообразования в русской технической терминологии // Труды Моск. ин-та истории, философии и литературы. — М., 1939. — Т. 5. — С. 8.

³ Реформатский А. А. Что такое термин и терминология. — М., 1959. — С. 2.

показывает, что среди однословных терминов терминосистемы робототехники 70,81 % составляют существительные. В вычислительной технике количество существительных среди однословных терминов снижается приблизительно до 64 %. В большей части это происходит за счет увеличения количества глаголов, также активно участвующих в терминообразовании, например команд. В терминосистеме подъязыка маркетинга среди однословных терминов зарегистрировано 60,12 % существительных также за счет еще более активного вовлечения в терминологию глаголов. Это подтверждает наличие у глаголов номинативных функций и оправдывает значительное число терминов-существительных с семантикой действия. Сопоставление данных относительно доли существительных в терминологии конкретных предметных областей, которые были получены как для русского, так и для английского языков, привело к выводу, что язык, на котором представлена терминология, не имеет значения. Отклонения в сотых долях не позволяют вывести какие-либо лингвистические закономерности. Для сравнения приведем данные о процентном соотношении основных частей речи, полученные при статистическом обследовании русского литературного языка: существительные составили 44 %, глаголы — 24, прилагательные — 13, наречия — 9.

Неубедительность формулировок сторонников разграничения терминологии и номенклатуры, практическая неактуальность такого разграничения и, напротив, их полная совместимость на основе вхождения номенклатуры в терминологию — все это достаточное основание для того, чтобы не усложнять информационные технологии дифференциацией этих понятий и рассматривать термины и терминосистемы как «совокупности словесных средств различения понятий или объектов специфического вида деятельности»¹.

5.4. Основные признаки системности терминологии

На терминосистему распространяются все основные признаки систем, но они наполнены специфическим содержанием. В связи с этим целесообразно проанализировать эти признаки и сформулировать основные представления о системности терминологии, раскрыв их как лингвистическое, так и, в необходимой мере, экстралингвистическое содержание.

К числу основных признаков системности, характеризующих условия существования и функционирования терминосистемы как системного объекта, участвующего в информационной технологии,

¹ Подробнее см.: Шайкевич А.Я. Проблемы терминологической лексикографии: Обзорная информация. — М., 1983. — Вып. 8. — С. 9.

относятся: целостность, структурность, взаимозависимость со средой, иерархичность, допустимость множественности ее описаний.

Наибольший интерес для практического исследования и прикладного использования, имея в виду особенности функционирования терминосистем предметных областей науки и техники, представляют три первых признака — целостность, структурность и взаимосвязь системы и среды.

5.4.1. Целостность

Под целостностью терминосистемы понимается невозможность сведения ее свойств к сумме свойств ее элементов и также невозможность выведения из последних свойств целого; т. е. зависимость каждого элемента, его свойств и отношений в системе от его места, его функций и от многих других взаимодействий внутри целого.

Объективность существования терминосистемы диктует необходимость познания ее как целостного объекта. И это познание достижимо с позиций исследования структурных связей целого и получения существующего в нем искомого элемента. Целостность терминосистемы — это условие ее существования, так как внешним силам может устойчиво противостоять только целостная система, а не отдельные подсистемы, ее элементы или связи между ними.

Признание целостности терминосистемы имеет большое значение для ее технологического существования, так как позволяет сделать вывод о невозможности внедрения в ее структуру принципиально инородных компонентов. Терминосистема каждой конкретной предметной области обладает иммунитетной жесткостью. Под иммунитетной жесткостью понимается способность противостоять проникновению в ее состав терминов, не отвечающих лингвистическим свойствам терминосистемы данной предметной области. Таким термином может быть, например, номинативная единица, принятая при появлении нового понятия. При отсутствии в сфере терминотворчества координации между терминосистемой и предметной областью в условиях экстралингвистической и лингвистической адаптации нового понятия, инициатива в процессе номинации принадлежит не систематизирующим свойствам терминосистемы, а скорее субъектам предметной области, которые могут не учитывать конкретных лингвистических условий. В полевых условиях действия какой-либо информационной технологии ситуация усугубляется сужением рамок интерпретации понятия до частной его реализации без учета возможного развития или любой формы его эволюции, при которой план содержания может вступить в противоречие с планом выражения.

Под каждой конкретной информационной технологией в предметной области понимается сложное социальное формирование,

представленное совокупностями материализуемых понятий (объектов предметной области), конкретных субъектов (языкового коллектива, объединенного коммуникативной деятельностью в предметной области) и социальных отношений. При таком понимании предметной области легко объяснима сложная двойственность процесса номинации новых понятий: на него оказывают воздействие, с одной стороны, субъекты предметной области, с другой — систематизирующие функции систем понятий и терминов.

При исследовании терминосистемы предметной области робототехники такие процессы можно наблюдать на конкретных примерах. Около 30 лет назад в США было разработано устройство в виде металлической пластины, имеющей форму замочной скважины. На ее расширенной площадке закреплялась обрабатываемая деталь, а удлиненная площадка в свою очередь устанавливалась в транспортере, который подавал деталь от одного исполнительного механизма робота к другому без необходимости вновь укреплять ее перед каждым исполнительным механизмом. Англоязычный вариант терминосистемы робототехники отразил это понятие термином «*pallet*». Вскоре это конструктивно-технологическое решение было освоено в нашей стране. Интернационализовавшееся слово *палета* было легко узнаваемо, мотивировано внешним видом и производимым им внешним впечатлением. Но с позиций лингвистических основ русского варианта терминосистемы робототехники произошла инверсия, т.е. не номинация понятия, а присвоение имевшейся языковой единицы сложившемуся новому понятию. По существу этой ситуации Х.Ортега-и-Гассет пишет: «Когда ученый открывает неизвестное явление, то есть когда он создает новое понятие, он должен его назвать. Поскольку совершенно новое слово ничего не говорило бы носителям языка, он вынужден пользоваться существующим лексиконом, в котором за каждым словом уже закреплено значение... Термин приобретает новое значение через посредство и при помощи старого, которое за ним сохраняется»¹.

В течение длительного периода специалисты активно пользовались термином «палета». Но он не был адаптирован русской терминосистемой (она не приняла его). В настоящее время в отечественной робототехнике существует соответствующее понятие, которое выражено русским термином «спутник». Замена произошла закономерно и естественно под действием иммунитетной жесткости терминосистемы.

Лексическая система любого языка всегда тяжело переживает внедрение иноязычного заимствования. Приведенный пример лиш-

¹ Ортега-и-Гассет Х. Две великие метафоры // Теория метафоры. — М., 1990. — С. 69.

ний раз доказывает языковую сущность терминологии и ее большее тяготение к лексису, чем к логосу. Кроме того, термин «палета» отражал только незначительный объем понятия, связанный с конкретной конструкцией. Термин не обладал свойством универсальности, логически вытекающим из признака целостности терминосистемы. Под влиянием французского источника (*palette*) происходило сужение семантики термина (*палета* — небольшая пластина, используемая в функции поддона), и в результате он лишь частично соотносился с понятием. Термин «спутник» имеет более широкую семантику и полностью соответствует новому понятию в робототехнике, постоянно конструктивно развивающемуся. Необходимо подчеркнуть, что терминосистема робототехники, отражая предметную область высочайшей универсальности по своему социальному назначению, должна и сама в полной мере соответствовать этому признаку.

Аналогичные примеры существуют и в маркетинге. Например, термин немецкого происхождения «банковский переводной вексель» довольно долго сохраняется в отечественной терминосистеме по причине, в свое время хорошо сформулированной в связи с ним в словаре иностранных слов: «В СССР сохранился только в торговых операциях с границей»¹. Само понятие было на продолжительное время исключено и его редкие употребления отражали содержание, определяемое другими общественными отношениями. В связи с изменением ситуации это понятие вновь активизировалось и потребовало более удобной номинации, что и было удовлетворено термином «тратта». Этот термин, итальянский по происхождению, имеет связь с более глубокими русскими традициями и обладает экономичным планом выражения, что очень важно для динамичного подъязыка маркетинга.

В терминосистеме всегда объективно существует свободная ячейка с направленными в нее валентными связями других терминов. Она готова принять новый термин только с валентными связями, способными замкнуть соответствующие связи других терминов, направленных в эту же ячейку. Упомянутый выше термин «спутник» соответствует этим условиям, универсален и обладает дополнительными потенциальными возможностями. На первый взгляд может показаться, что изменение плана содержания привело к изменению плана выражения. Но в происшедшем есть более широкий смысл: терминосистема обладает целостностью, характеризующейся иммунитетной жесткостью.

Оба эти примера иллюстрируют естественное противодействие терминосистемы проникновению инородной лексики. Наличие в системе потенциально свободных связей может в некоторые моменты под влиянием внешних сил стимулировать заполнение их

¹ Словарь иностранных слов. — М., 1980. — С. 101.

лексическими единицами с высокой валентностью. Такими единицами обычно бывают элементы общенаучной лексики. Высокая их продуктивность в терминообразовании может послужить причиной временного насыщения потенциально свободных связей и искусственного размывания границ системы.

Понятие границ — важнейшая составляющая целостности терминосистемы. Четкие представления о границах создают основу для понимания структуры, функций и назначения системы. Такие границы в большой степени относительны, хотя и четко обозначаемы, так как создают вполне определенные условия, необходимые для функционирования терминосистемы во взаимодействии со средой и определяют ее целостность.

5.4.2. Структурность

В информационно-технологическом плане понятие структуры было определено В. М. Солнцевым как совокупность связей и отношений между составляющими элементами¹. Структура это атрибут некоей системы, но и она не равна объекту в целом. Можно сказать, что структура есть объект минус элементы, его составляющие, или система минус ее наполнение. Для конкретной специфической системы, каковой является терминосистема, этот признак имеет особый смысл, так как выводит на понимание процессов адаптации новых терминов — результатов терминотворчества, активно происходящего в развивающихся предметных областях.

Во-первых, структурность терминосистемы такова, что позволяет присоединять к ней новые термины, не входя в противоречие со свойством ее целостности. Это значит, что структурные границы терминосистемы не являются лимитирующим инструментом, накладывающим какие-либо ограничения на коммуникацию как организацию действий в предметной области, а наоборот, они представляют собой фактор, содействующий развитию, поскольку, уточняя выразительные средства, способствует совершенствованию логических процессов, происходящих в предметной области.

Структурность терминосистемы обеспечивает не хаотичное накопление, а упорядоченное присоединение новой терминологии. Этот вывод имеет особое практическое значение, поскольку определение места, занимаемого термином, дает дополнительную информацию для его дефиниции, через связи в уже существующей структуре системы.

При противопоставлении терминологии общей лексике в качестве основного аргумента в пользу такого противопоставления

¹ См.: Солнцев В. М. Язык как системно-структурное образование. — М., 1977. — С. 5.

используют отсутствие в терминологии полисемии и синонимии. Однако наличие компонента лексиса в собственно термине оставляет за терминологией право на полисемию, которая вместе с тем в значительной степени подавляется строго упорядоченной структурой терминосистемы. Влияние структурности, обладающей сильными связями, позволяет дифференцировать терминопотребления в коммуникативном процессе.

Например, в робототехнике имеются два родовых понятия со следующими дефинициями:

1 — устройство, как правило, дистанционно управляемое оператором и/или программным средством, содержащее рабочий орган, который предназначен для имитации перемещений и рабочих функций кисти руки оператора;

2 — механическая часть робота.

В русской терминосистеме оба эти понятия именуется одинаково: «манипулятор». В англоязычном варианте терминосистемы робототехники ситуация полностью совпадает.

Если этот пример может показаться в некоторой степени отягощенным интернационализацией, то следующий пример, на наш взгляд, более объективен:

1 — устройство балансировки рабочих плечей исполнительного механизма для уравнивания нагрузки, создаваемой объектом обработки;

2 — накопительное устройство в виде опрокидывающегося бункера при смещении его центра тяжести по мере накопления, например, готовых изделий.

В русской терминосистеме этим дефинициям соответствует один термин «противовес», и в английской также один — «counterweight».

В основных терминах по маркетингу можно найти аналогичные примеры. Термин «стояночное время» может быть соотнесен с понятием срока для разгрузки и погрузки и в равной мере с понятиями: а) срока регламентных работ; б) срока санитарно-эпидемиологического контроля. Причем в английской терминологии во всех ситуациях также используется один термин «lay days».

Факты синонимии в терминосистемах признаются многими исследователями. Одни рассматривают их как обычные явления, другие придают им характер исключительности. Признавая явления синонимии в терминологии и имея доказывающие примеры во всех трех сравниваемых языках, мы в основном соглашаемся с точкой зрения К. Я. Авербуха, установившего причины терминологической синонимии, основная из которых — активное заимствование из других языков¹.

¹ См.: Авербух К. Я. Терминологическая вариантность: теоретический и прикладной аспекты // Вопросы языкознания. — 1986. — № 6. — С. 39.

Во-вторых, структурность терминосистемы обуславливает системоприобретенные свойства новых терминов. Под системоприобретенными свойствами понимаются те свойства, которыми термин наделяется, вступая в валентные связи и отношения внутри терминосистемы. В отличие от обычных единиц в лексической системе языка, где системоприобретенные свойства не столь сильны. для термина, вступающего в систему, и для самой терминосистемы, эти свойства весьма важны. Большинство терминов представляют собой вторичное использование лексической единицы в результате членения и трансформации ее семантического состава, при котором неизбежно образование валентностей особого характера — валентностей «пустоты» (т. е. отсеченного участка семантического состава), что восполняется системоприобретенными понятийно-смысловыми компонентами.

Этот аспект имеет существенное значение для предметной области робототехники, где достаточно много номинативных единиц с семантикой действия. Например¹:

Захватить 1. кого-что. Хватая, взять, забрать, схватить.

2. кого-что. Силой овладеть кем-чем-н.

3. кого-что. Взять с собой.

4. перен., кого (что). Сильно заинтересовать, увлечь, поглотить все внимание, все силы.

5. кого (что). То же, что застать (прост.).

6. что. Вовремя принять меры против чего-нибудь.

Для номинации в робототехнике используется только первое значение, и оставшийся семантический потенциал этого глагола (валентность «пустоты») требует нейтрализации по одной из самых распространенных в терминосистеме робототехники модели:

$$V \rightarrow A + S$$

захватить → *захватное* + *устройство*

У глагола *захватить* при переходе на терминологический уровень остается не использованной часть его семантического состава, вместо которой у него образуется валентность. При установлении семантических связей внутри терминосистемы робототехники эта валентность обеспечивает его системоприобретенными понятийно-смысловыми компонентами: *захватное устройство*, *захватный механизм*.

Структурность терминосистемы является ее важнейшим свойством в связи с еще одной спецификой ее элементов — терминов: они могут быть как простыми — однолексемными, так и сложными — синтаксическими конструкциями. Такая специфика элемента системы в данном случае не только не нарушает, но даже подчеркивает признак структурности. Дискретность как признак по

¹ Англо-русский словарь по вычислительной технике. — М., 1990. — С. 201.

отношению к терминосистеме в целом никак не связана с относительной делимостью сложного термина, оформленного в виде синтаксической конструкции. Делимый термин участвует в терминосистеме как единый элемент, а его свойство относительной делимости касается реализации его синтагматических связей.

Как показали лексико-статистические исследования, проведенные на текстах двух предметных областей (робототехники и маркетинга), для русского языка наиболее характерны сложные в синтаксическом отношении термины, вполне, однако, усваиваемые и систематизируемые в структуре терминосистемы.

Например, среди более 330 основных английских терминов по маркетингу только один содержал пять лексем; бóльшая же их часть (более 60 %) состояла из двух-трех лексем. В то же время в русской терминосистеме 24 термина содержали пять и более лексем. Большинство «длинных» русских терминов представляли собой сложные синтаксические конструкции: *товар, определенный общими родовыми признаками (fungible goods)*; *торговый аппарат, построенный в разбивке по клиентам (customer-structured sales force)*; *установление цен с принятием на себя расходов по доставке (freight absorption pricing)* и т. п.

Признак структурности терминосистемы не входит в противоречие с тезисом о наличии границы как необходимом условии существования терминосистемы и о влиянии границы на ее целостность и принципы функционирования. Связи и отношения, составляющие структуру терминосистемы, не замыкаются ее границами. Природа связи состоит в нарушении границ, и связь в таком качестве может характеризоваться видом нарушения.

Степень структурированности терминосистемы в определенном смысле служит мерой ее познаваемости. Постигание свойств терминосистемы через разложение ее структуры имеет практический смысл в условиях межъязыковой коммуникации при переводе текста с одного языка на другой. Принцип взаимообусловленности языка и его терминосистемы дает нам основание констатировать и объяснять факты несовпадений конкретных языковых вариантов терминосистемы, например вариантов терминосистемы предметной области робототехники в русском и английском языках. Но разложение структур, позволяющее выделить в них подсистемы и микросистемы, создает возможности для сопоставительного анализа в целях повышения качества перевода на уровне терминологического обеспечения.

5.4.3. Взаимосвязь системы и среды

Языковая система и ее подсистемы функционируют во взаимодействии между собой и с окружающим миром. Взаимодействие с окружающей средой — это условие существования системы, и

оно является одним из основных системных признаков. Понимание механизмов функционирования терминосистемы возможно только при анализе особенностей ее соотношений со средой. Ни о каком функционировании в принципе не может идти речь, если не иметь в виду среду, в которой это функционирование осуществляется.

Исследование среды, в которой реализуется терминосистема, представляет интерес не с позиций оценки окружающих условий, а с позиций понимания процессов, происходящих в терминосистеме. Здесь важна закономерность того, что множество образует устойчивую структуру, если связи определенного вида между элементами этого множества (внутренние связи) преобладают над аналогичного вида связями между элементами множества и окружающей средой (внешние связи). В настоящее время мало изучен аспект взаимоотношений системы и среды на переходе «действительность — возможность». В аспекте такого перехода для существования терминосистемы необходимо наличие другого объекта — среды. Только при этом условии можно экспериментально или теоретически указать валентности системы и среды, которыми они будут обладать, если при заданных внешних связях окажутся во взаимодействии.

Что же составляет среду функционирования терминосистемы?

Прежде всего ее можно характеризовать как сложноструктурную системную среду, состоящую из ряда отдельных подсистем, каждая из которых обладает независимыми или малозависимыми между собой связями с терминосистемой. Поскольку отношения между системой и средой формируются под воздействием системности реального мира, они имеют функциональный характер, т. е. эти отношения направлены на выполнение какой-либо функции. В конкретном случае терминосистемы предметной области робототехники требуется реализация функции обеспечения коммуникации в общепринятых терминах. Это создает необходимость взаимодействия терминосистемы робототехники с системой понятий в этой предметной области, с системами технической политики и социальных отношений, и, наконец, с лексической системой языка. Все эти составляющие окружающего мира образуют среду функционирования терминосистемы.

Дихотомия «система — среда» приобретает смысл только при наличии валентных связей, которыми располагают и терминосистема, и среда. По этим связям в любой информационной технологии осуществляются обменные процессы, суть которых на одном из уровней заключается для терминосистемы — в терминологическом обеспечении коммуникативного процесса, а для системы понятий — в формировании потенциального энергетического состояния терминосистемы как инициатора номинативного процесса.

По характеру обменных процессов можно судить о динамических возможностях терминосистемы и об уровне развития предметной области.

Если по отношению к терминосистеме обменные процессы строятся на входящих потоках, можно говорить о развитии предметной области и активном номинативном процессе. Если потоки уравновешены, то такая терминосистема соотнесена с системой понятий неразвивающейся предметной области.

Наблюдения за терминосистемой и системой понятий развивающейся предметной области робототехники дают возможность утверждать, что возможно и третье состояние, когда обменные процессы терминосистемы и системы понятий строятся на исходящих из терминосистемы потоках, т. е. терминосистема приобретает продуцирующую функцию, инициированную косвенными факторами.

Это глубоко специфический аспект функционирования терминосистемы, поскольку отражает характер связей терминосистемы с системой понятий конкретной предметной области и последней — с системами понятий смежных предметных областей.

Во-первых, в условиях сильной интерференции предметных областей, связанных с усложнением технологических процессов и системным подходом к решению проблем в науке и технике, имеет место не просто взаимопроникновение терминов, а их взаимообусловленность. Построение термина робототехники идет по моделям терминов той предметной области, где робототехника находит применение.

В английском языке специфический термин предметной области микроэлектроники *clean-room electronics* («интегральные схемы, изготавливаемые в особо чистых условиях») вызвал появление термина *clean-room robotics* («роботы для работы в особо чистых условиях»). У этих двух терминов совпадает только синтаксическая конструкция, а объем понятий сильно различается: в первом случае понятие касается объекта, во втором — субъекта, но окружающая среда — тексты по применению робототехники для производства интегральных схем — у обоих терминов одинаковая и сходные подчинительные связи.

В данном случае коммуникативный процесс в терминологическом плане обеспечивают две терминосистемы, и, если в одной из них имеется устойчивый термин, а в другой его нет, но в системе понятий возникает нечто аналогичное, среда функционирования может становиться регламентирующим фактором номинативного процесса.

Под воздействием опыта терминосистемы предметной области микроэлектроники и правила, выведенного средой функционирования, роль англоязычного варианта терминосистемы робототехники оказалась активнее роли системы понятий предметной

области. И существующие между терминосистемами языковые связи, осуществляемые в единой среде, обобщили терминообразовательную модель.

Во-вторых, глубокое проникновение робототехники в сферу производственных отношений создало условия для построения ее терминов по моделям, соотнесенным с понятиями служебной иерархии:

blue collar robot («рабочий робот»);

white collar robot («робот-координатор»).

В-третьих, терминосистема робототехники порождает термины, соотносимые с избыточной детализацией, которые, впрочем, при определенных условиях активно используются. К ним относятся термины, происходящие из аналогий действий робота и действий человека. Такова серия антропоморфизмов в робототехнике: *рука, запястье, предплечье, плечо, палец* и др.

Все это характеризует окружающую среду как совокупность обобщающих каналов связи, по которым осуществляются обменные процессы. Обобщающий характер каналов связи, поддерживаемых средой, наполняет обменные процессы опытом реализации качества, накопленного в различных элементах среды и системы.

И. В. Арнольд предложила семиотический подход к исследованию взаимодействия языковых систем и среды¹. В узкоприкладном понимании семиотического членения на синтактику, семантику и прагматику можно наблюдать три аспекта взаимодействия терминосистемы со средой. Принято считать, что синтаксический аспект раскрывает только внутренние связи. Однако с учетом многоуровневой структуры терминосистемы и возможностей самостоятельного, а зачастую автономного функционирования ее уровней в виде подсистем смысл синтаксических связей как межэлементных значительно расширяется из-за комплексности и структурной сложности элементов.

Без учета окружающих условий процессы упорядочения терминологии в терминосистеме, конечно, можно было бы объяснить внутрисистемными отношениями, т. е. на уровне синтактики. Понимание внутренних связей, безусловно, дает возможность описать принципы функционирования терминосистемы. Но такое описание было бы статичным и представляло бы идеализированную модель, которая действительно в состоянии отобразить принципы, но не результаты функционирования терминосистемы или даже ее подсистем, что лишает исследование практической ценности.

С учетом среды появляется возможность продуктивного анализа семантических связей терминосистемы с предметной областью

¹ См.: Арнольд И. В. Основы научных исследований в лингвистике. — М., 1991. — С. 125.

и прагматических связей со своими пользователями и с системой понятий.

Кроме общего представления о семантических связях как о соотношении лексического состава языка и предметного мира при системном подходе к пониманию принципов и результатов функционирования терминосистемы в соотносимой с нею среде важен учет влияния диалектики количественно-качественных превращений. Накопление знаний о предметной области активизирует номинативный процесс. С появлением новых терминов происходит перераспределение внутренних связей. Такое перераспределение под воздействием количественных изменений соотносимо с прогрессом предметной области. Но в терминосистеме они имеют более конкретный характер по сравнению с прогрессивными процессами в предметной области, поскольку классификация в системе терминов всегда конкретнее и нагляднее, чем в системе понятий. Качественные изменения терминосистемы отражаются на изменении ее связей со средой. Подобный естественно существующий дисбаланс становится инициатором динамического обмена между терминосистемой и средой.

В терминосистеме количественные изменения происходят в двух плоскостях: 1) расширении синтагматических связей, соотносимых с углублением исследовательских процессов в предметной области, т.е. с детализацией имеющихся знаний; 2) углублении парадигматических связей, соотносимых с расширением исследований, т.е. появлении новых направлений в исследованиях. Здесь трудно ожидать ярко выраженных результатов, но изменения в первой плоскости больше связаны с усилением синонимии, во второй — с полисемией и омонимией.

Исследование прагматического компонента системы внешних связей даст возможность оценить роль участников коммуникативных действий в обменных процессах с терминосистемой. Этот аспект представляет интерес своей непосредственной связью с кругом социолингвистических проблем. Определяя направления и проблематику научных исследований, социальные условия влияют на формирование объема понятий и на развитие их системы.

Роботизация бытового обслуживания в США и Японии послужила поводом для изменения отношения к роботам. Роботы приобретают элементы искусственного интеллекта и способны выполнять некоторые логически сложные последовательности операций; более того — стало возможным речевое управление ими. С этими процессами связано появление в англоязычном варианте терминосистемы робототехники таких терминов, как *ear robot* («робот с акустическими сенсорами») или *oral robot* («робот с голосовым управлением»). Семантика этих терминов ясна и они легко адаптируются в вариантах терминосистем других языков.

Экономия выразительных средств, свойственная языку вообще и имеющая тенденцию к усилению с развитием индустриализации, наиболее заметна в устной речи. В терминосистеме робототехники существует немало примеров нормализации устных терминов. Но вопрос о существовании устных и письменных речевых терминов — это вопрос синонимии в терминосистеме.

В прагматическом аспекте взаимосвязей терминосистемы и среды важен учет экстралингвистических проблем. Их значение и влияние находятся в прямой зависимости от особенностей каждой конкретной предметной области. В робототехнике значение экстралингвистических проблем определяется спецификой взаимоотношений в системе «человек — машина» и особенностями функционирования «интеллектуальных систем» в антропоцентрической среде. Экстралингвистический подход к оценке робототехнической терминологии может объяснить объективность существования среды терминов некоторых алогизмов.

Таким образом, функционирование терминосистемы конкретной предметной области не может быть понято и объяснено в практических целях без достаточно глубокого анализа ее взаимодействия с окружающей средой.

Терминосистема и ее среда соподчинены функционально. Их отношения и связи служат обеспечению процесса общения и организации целенаправленного взаимодействия в конкретных регламентированных социальных условиях.

5.5. Внешние и внутренние признаки типологизации терминологии

На фоне бурного развития вычислительной техники трудно поверить в то, что в области идеологии и технологии создания компьютерных словарей не происходит сколько-нибудь значительных изменений. Но, к сожалению, это так.

Работа с текстовой информацией требует от компьютера не только высокой скорости выполнения рутинных операций как и в большинстве задач естественных наук, но еще и огромных объемов оперативной и дисковой памяти. Поэтому основные изменения в области компьютеризации словарного обеспечения переводческого процесса происходят не в усовершенствовании программного обеспечения, а в простом наращивании возможных переводных эквивалентов внутри словарной статьи. Но идеология автоматического переводного словаря, по крайней мере, за последние 20 лет практически не изменилась.

Терминологии разных предметных областей могут значительно отличаться по своей структуре и принципам соотношения с системой понятий предметной области. Так, например, анализ распре-

деления терминов по частям речи указывает на существенные различия лингвистических параметров терминологий разнотипных предметных областей.

На фоне того, что уже доказано существование терминоведения как самостоятельного раздела языкознания и определены его предмет, методы и структура, еще совершенно не разработана его типология. Вместе с тем достаточно четко сформулированы теоретические принципы типологии терминосистем по внешним и внутренним признакам¹.

Внешним признаком, обусловленным связью языка с внеязыковыми системами, является соотнесенность терминологии:

– с различными областями трудовой деятельности людей — производством, техникой, наукой, управлением;

– с профессиональным членением каждой из областей трудовой деятельности: в машиностроении — станкостроение, робототехника и т. д.;

– с научными направлениями и техническими школами и методами, с личностными системами научных взглядов, технологических и конструктивных концепций.

К наиболее существенным внутрискруктурным свойствам терминологии относят:

а) понятийную сторону семантики термина (связь термина с понятием), одновременно сближающую термины и удаляющую их друг от друга, формирующую терминологические пласты, ряды, поля и др.;

б) словообразование и дериватологию терминов, устанавливающую семантические (понятийные) связи, влияющие на формирование терминосистемы и отношения между ними;

в) логико-семантическую, семантико-морфологическую и семантико-дериватологическую парадигматику терминов;

г) внутреннюю (т. е. возникающую в пределах одного и того же терминосочетания) и внешнюю (в цепи высказывания между терминами, а также терминами и нетерминами) синтагматику;

д) лексико-семантические свойства терминов, обуславливающие терминологические синонимии, антонимии, семантические сближения и удаления терминов, пересечения и подчинения их значений.

В общую структуру, или макроструктуру, автоматического переводного словаря входят основные компоненты словарной системы, объем и границы словника, принцип расположения лексических единиц.

В самом общем виде автоматический переводной словарь состоит из трех основных компонентов — блока обработки запроса,

¹ Более подробно об этом см.: Головин Б. Н. Типы терминосистем и основания их различия // Термин и слово. — Горький. 1981. — С. 3—10.

лексических массивов, блока выдачи ответа. Здесь, бесспорно, просматривается сходство со структурой систем МП, которые в общем виде также состоят из трех компонентов: блока анализа входного текста, словаря и блока лексических трансформаций, блока синтеза выходного текста. Однако при сопоставлении содержания и функций составных частей автоматического переводного словаря и систем МП обнаруживается больше различий, чем сходств.

Блок анализа входного текста в системе МП осуществляет морфологический, синтаксический анализ лексических единиц входного языка с целью получения максимально возможной информации о грамматической форме слова, его роли в предложении и о его связях с другими словами. В некоторых продвинутых системах предусматриваются также семантический анализ входного текста и семантический синтез выходного текста.

Задачи блока обработки запроса в автоматическом словаре существенно проще и чаще всего заключаются в сведении словоформ к их словарным (исходным) формам. В словарях с исходной формой запроса или с поисковыми массивами словоформ блок обработки запроса в лингвистическом отношении фактически отсутствует, так как в таких системах на этапе запроса осуществляется побуквенное сопоставление запроса с лексическими единицами из состава словаря без какой-либо лингвистической обработки самого запроса. Вместе с тем такой упрощенный подход к задачам и функциям этого блока уменьшает ценность словаря, поэтому обычно создатели автоматических словарей уделяют особое внимание увеличению их разрешающей способности на этапе обработки запроса, что позволяет повысить гибкость этого блока, упрощает работу со словарем, расширяет его возможности.

Сравнение принципов построения словарей систем МП и лексических массивов автоматических словарей также выявляет немалые различия. Лексические единицы входного языка в системах МП хранятся в форме машинных основ. Каждая основа имеет набор данных о грамматической принадлежности слова, его однозначности или многозначности, о парадигме словоизменений, моделях синтаксического управления и другой информации, необходимой при анализе входного текста. Специально указывается номер переводного эквивалента, который является относительным адресом соответствующей лексической единицы выходного словаря в памяти ЭВМ. Лексика выходного языка также хранится в виде машинных основ и получает информацию, необходимую для синтеза выходного текста. В автоматическом переводном словаре лексические единицы могут храниться как в виде основ, так и в виде словоформ или словарных форм — в зависимости от наличия или отсутствия алгоритма лемматизации и принципов его построения. Словарная статья по составу информации и по

форме ее представления значительно ближе к словарным статьям традиционных словарей, чем к словарям систем МП.

Блок выдачи ответа — наиболее простая часть автоматического переводного словаря. Его основное назначение состоит в выдаче в словарной форме переводного эквивалента запрошенной лексической единицы. В отдельных случаях этапу выдачи ответа предшествует этап морфологического синтеза словарной формы слова из его основы и соответствующего окончания. Еще одной задачей этого блока является варьирование объема ответа в зависимости от пожеланий пользователя. Обе эти задачи с лингвистической точки зрения значительно проще, чем задачи, стоящие перед блоком синтеза выходного текста систем МП, а именно: морфологический синтез соответствующей словоформы и синтаксический синтез предложения выходного языка, являющегося переводом соответствующего предложения входного языка.

Простота назначения блока выдачи ответа в автоматическом переводном словаре — это одновременно и сознательный, и вынужденный шаг. Его осознанность обусловлена тем, что автоматический словарь в своей основной функции ориентирован на диалог с человеком, в котором словарь берет на себя всю работу с лексикой, а решение трудноформализуемых задач, таких, например, как выбор правильного переводного эквивалента из нескольких, предложенных словарем, и синтез выходного текста возлагаются на человека. Вынужденность упрощенного построения блока выдачи ответа определяется ограниченными возможностями блока обработки запроса. На вход автоматического словаря поступают отдельные слова и словосочетания, с помощью анализа которых можно получить лишь сведения о грамматическом классе слова и его грамматической форме. Решить на основании этой информации проблемы омонимии, многозначности, определить грамматическую форму переводного эквивалента и его синтаксическую функцию в тексте невозможно, поэтому разработчики автоматических словарей таких задач перед собой не ставят.

В настоящее время в языкознании накоплено немало лингвистических исследований терминологий отдельных предметных областей. Их материалы могут предоставить достаточные сведения для сопоставительного анализа.

Исследуя терминологии подъязыков робототехники и маркетинга и, кроме того, имея основные данные из других терминологий, удастся подтвердить универсальность теоретической методики типологизации терминосистем, предложенной Б. Н. Головиным. Соотнесенность терминологии с областями трудовой деятельности, безусловно, является типологизирующим признаком. Терминологию робототехники, обладающую высокой овеществленностью, следует отнести к терминологиям техники. По этому параметру она близка к терминосистемам таких областей, как гибкие

производственные системы или механическая обработка материалов и др. Менее овеществлена терминология подъязыка вычислительной техники и программирования. И еще дальше от нее по степени овеществленности отстоит терминология маркетинга.

Не делая никаких различий в оценке характеристик функционирования между номенклатурой и терминологией, мы использовали такой параметр, как относительное содержание номенов в качестве выразителя овеществленности при сравнительном статистическом исследовании терминосистем. Количество номенов в терминологии робототехники составило 49,3 %, вычислительной техники и программирования — 38,6 %, маркетинга — 14,7 %. Высокая дифференциация результатов убеждает в возможности применения такой методики типологизации и разграничения терминосистем.

Общие наблюдения за функционированием терминологии в научных, технических и рекламных текстах, которые постоянно ведутся в режиме аналитико-синтетического процесса по технологии Всероссийского центра переводов, дают достаточные основания считать, что в терминологиях технических областей деятельности половина терминов соотнесена с овеществленными денотатами. В областях, больше связанных с научной деятельностью, этот показатель падает до 30—35 %, а в экономике и управлении он не превышает 15 %. В последнем случае наблюдается явное преобладание отвлеченных понятий и категорий.

Оценка соотнесенности терминологии с профессиональным членением внутри трудовой деятельности имеет практический смысл для разграничения смежных терминосистем, а также для продуктивного исследования иерархии внутри терминосистемы. Как бы ни были близки между собой области робототехники и гибких производственных систем, они имеют свои разграниченные системы понятий и соотнесенные с ними самостоятельные терминосистемы. Близость этих предметных областей позволяет, например, издать общий словарь по гибким производственным системам и робототехнике. Наравне с терминами робототехники в терминосистеме гибких производственных систем используются термины автоматизированных систем управления и др. Выделяя главные признаки терминосистем, Б. Н. Головин отметил, что границы между терминосистемами могут быть более четкими или менее четкими («размытыми»), что терминосистемы могут накладываться друг на друга и в таких случаях образуются их пересечения.

Проблема соотнесенности терминологии с научными направлениями и техническими школами мало выражена в традиционных технических областях. Но эта проблема действительно существует, в некоторых случаях стоит достаточно остро и приобретает типологизирующий характер. Более всего она проявляется в

Американский вариант	Английский вариант	Русский перевод
Code	programme	программа
Font	fount	шрифт
Procedure	subroutine	процедура

терминологии научных направлений. По терминологическому составу английского текста часто представляется возможным атрибутировать его принадлежность к американской или европейской школе программирования и вычислительной техники. Сказанное иллюстрируют примеры, приведенные в табл. 3.

Для американского варианта подъязыка вычислительной техники и программирования характерно повышенное содержание сокращенных написаний и аббревиатур.

Что касается соотносительности терминологии с личностными системами научных взглядов, то экстралингвистические исследования приводят к выводу о реальном существовании такого признака. Однако современные информационные технологии перевода характеризуются широким применением текстовых процессоров со всевозможными сервисными «чекерами», которые в значительной степени выхолащивают какую-либо персонализацию терминологии.

С учетом тенденций к антропоцентризму и прагматике в науке и тенденций к изучению языка в связи с деятельностью и личностью человека в языкознании проблема соотносительности терминологии с личностными системами научных взглядов постепенно может приобрести большую актуальность.

Согласно определениям разработанной Б. Н. Головиным типологии терминосистем, например, терминосистему подъязыка предметной области робототехники можно представить как техническую, с узкой специализацией в пределах отрасли машиностроения и частичным пересечением с терминосистемами станкостроения и гибких производственных систем и с несущественным влиянием научных и технических школ и личностных научных взглядов.

Среди внутривидовых типологических признаков ведущую роль, конечно, играет оценка понятийной стороны семантики гермина. Все приведенные выше внутривидовые параметры типологизации, предложенные Б. Н. Головиным, вытекают из главной проблемы семантизации терминов. К сожалению, такая подробная градация параметров для реально функционирующей терминосистемы практически неосуществима. Процесс семантизации терминологии имеет типоразличительное значение.

Важнейшим средством семантизации является толкование термина. В ситуации развивающихся предметных областей не все тер-

мины одновременно имеют полные исчерпывающие толкования. По полноте толкования можно судить о новизне термина и степени его адаптированности в терминосистеме. Как правило, до 35 % новых терминов (пока они обладают статусом новых) не имеют определений, а 10 % из них требуют расширенного толкования, и их перевод на русский язык носит описательный характер. Проблемы семантизации в таком случае имеют несколько косвенных решений, которые укладываются в перечень внутрисистемных признаков типологизации, приведенный выше.

Наравне с традиционным способом семантизации по толкованиям новые информационные технологии используют еще несколько способов семантизации новых терминов: через контекст, через синонимы, через иллюстративный материал и через переводные эквиваленты. В табл. 4 дан пример семантизации через синонимы.

А вот пример семантизации одновременно через контекст и переводные эквиваленты: *Use the **cold site** to support the operation of a computer system in the event of fire.* (Для поддержания работоспособности вычислительной системы при пожаре используйте резервное устройство.) Таким образом определен термин «резервное устройство». По английскому эквиваленту была выяснена семантика термина, а по контексту сформирован план выражения.

Есть еще особый способ семантизации — это интуиция термиолога. Мы относим этот способ к вполне научным, поскольку он основывается на экстралингвистических научных знаниях термиолога. О требованиях к современному термиологу неоднократно писали многие ученые. Все сходятся на том, что ему необходимы знания предметной области и опыт лингвистического анализа и описания терминологии. Экстралингвистические знания, развивающие интуицию, могут быть инструментом решения контекстуальных задач.

Если в предметной области вычислительной техники проблемы семантизации — это чаще всего проблемы лингвистического описания термина, которые в основном включают анализ семантики и выработку достаточного толкования, то в предметной области робототехники они скорее опираются на интерпретацию синонимов. Высокую синонимичность можно считать типологическим признаком терминосистемы робототехники. Например, из

Таблица 4

Новый термин (ISO-IVP)	Синоним	Эквивалент
Administrative security	procedural security	защита диспетчерских данных (в вычислительных сетях)
Trailer label	end-of-file label	метка конца файла

830 новых терминов по вычислительной технике и программированию только около 5 % имеют по одному синониму. Большого количества синонимов не имел ни один термин. В робототехнике из 789 английских терминов по одному синониму имеют 13,28 % терминов, по два — 3,13 %, по три — 1,56 % и 0,78 % терминов имеют четыре синонима, т. е. почти 19 % терминов по робототехнике имеют синонимы. В то же время из 338 терминов по маркетингу только два имели синонимы (0,59 %).

Эти статистические данные демонстрируют типологизирующий характер внутрисистемных семантических исследований. Для бурно развивающейся технической предметной области характерна высокая синонимия.

5.6. Условия соотнесенности терминов и понятий

Соотнесенность терминологии с системой понятий предметной области является самой главной и многоаспектной проблемой в круге вопросов системного терминологического обеспечения аналитико-синтетических процессов. Основные аспекты этой проблемы создают фундамент для большинства направлений исследования терминосистемы. Собственно терминосистема не может существовать вне связи с предметной областью.

В практическом смысле наибольший интерес представляют два аспекта: безусловная соотнесенность терминологии с системой понятий и объективно-субъективная лингвистическая деформация структуры этой соотнесенности. Первый из них достаточно хорошо исследован и освещен в лингвистической литературе. Второй, пожалуй, служит источником основных неудач коммуникативного процесса.

Терминология связана с предметной областью непосредственно. Но если при ее формировании и исследовании исходить исключительно из ее взаимодействия с предметной областью (что составляло бы нереально идеальную ситуацию), то многие проблемы перевода терминов не были бы решены, а процесс перевода превратился бы в сумму двух слагаемых: подстановки эквивалентов языка перевода на место терминов текста оригинала и, если это необходимо, реорганизации первых в тексте перевода.

Более сложны терминологические проблемы перевода в реальных ситуациях. Исходя из первичности соотнесенности терминологии с системой понятий и того, что материальная предметная область существует объективно и развивается независимо от языка, который только впоследствии описывает процесс и результат этого развития, терминологию предметной области можно представить как определенное множество, имеющее структуру и связи, т. е. терминосистему. В понятие единой предметной области мы

вкладываем тот смысл, что, развиваясь независимо от языка, она по своей сути интернациональна и существует в неограниченном пространстве, а само пространство не дифференцируется никакими условиями за исключением структурной дифференциации предметной области. Неограниченное пространство едино по своей логической сути, а следовательно, и отображаемая в нем предметная область также едина.

Различия в степени развития одной предметной области, например, в разных странах, оказываются результатом неодинакового уровня освоения данной предметной области. Эти различия, как правило, имеют экстралингвистический характер.

Средства языка, и прежде всего терминологическая лексика, описывают текущее состояние предметной области. Такое описание и его полнота являются фактором и мерой понимания языковым коллективом уровня состояния предметной области. В любом языке средства для описания предметной области достаточны, но их различия, естественно, порождают многообразие описаний и субъективируют их на основе социальных доминант языковых средств в определенном языковом коллективе.

Субъективно-социальное описание предметной области, связанное с номинативным процессом, обусловленным конкретной языковой ситуацией, порождает вариант терминосистемы, отличающийся от других вариантов ее языковой интерпретацией.

В идеальном случае терминосистема — это объективно существующая модель, повторяющая своей структурой структуру системы понятий предметной области. Любое системное представление терминологии средствами конкретного языка является вариантом терминосистемы.

5.7. Вариативность терминосистемы

Процесс описания предметной области — это средство ее познания и создания базы знаний. База знаний представляет собой совокупность систематизированных основополагающих сведений, относящихся к определенной предметной области, объем которых необходим и достаточен для решения заданного круга теоретических или практических задач. Это определение можно экстраполировать на такую совокупность систематизированных сведений, как терминосистема, т. е. описание терминосистемы может и должно быть вариантом базы знаний о конкретной предметной области.

Описание терминосистемы как базы знаний основано на объективном развитии данной предметной области и представляет собой объективно-творческий процесс. Для познания самого процесса и его результатов в основном требуются экстралингвистические методы исследования.

Формирование языкового варианта терминосистемы как ее частной реализации средствами конкретного языка — это субъективно-творческий процесс в масштабах конкретного языка, поэтому для его реализации необходимы прежде всего лингвистические методы исследования.

Предметная область объединяет варианты терминологии и даже унифицирует их. Но этот процесс имеет свою меру и свои условия. Если бы это явление унификации существовало действительно объективно, не было бы никакой асимметрии в терминологии и термины появлялись бы на всех языках одновременно, исключая при этом всякие сложности их перевода.

Процессы, происходящие в системе понятий предметной области, определяют основные аспекты функционирования терминосистемы. Этот вывод особенно справедлив в отношении новых предметных областей и их терминосистем. Проблема функционирования новой терминологии развивающейся предметной области представляет наибольший интерес в практическом плане при исследовании коммуникативных процессов, обслуживаемых подъязыком данной области.

Понимание проблемы и решение практических задач терминологического обеспечения аналитико-синтетических процессов и особенно перевода весьма затруднено при исследовании только какой-либо одной терминосистемы. Экстралингвистические аспекты проблемы столь сильны сами по себе и так предопределены особенностями предметной области, что не могут субъективизировать лингвистические аспекты исследования и способны привести лишь к частным выводам, нередко вступающим в противоречие с реальной ситуацией аналитико-синтетического процесса, в частности перевода с одного языка на другой.

5.7.1. Асимметрия вариантов терминосистемы

В реальных условиях мы всегда имеем дело с сильной асимметрией вариантов терминосистемы на разных языках. Для сравнения выгодно выделить три предметные области: робототехнику, вычислительную технику и маркетинг.

Основной мерой несхожести вариантов терминосистем является асимметрия — комплексная характеристика, в составе которой выделяются различия:

- в восприятии объема понятий;
- в языковых средствах выражения;
- между формами изложения и восприятия.

Тесно связанные между собой, эти аспекты имеют определенную самостоятельность и в различных ситуациях проявляются неодинаково.

Различия восприятия объема понятия. Это качество асимметрии имеет выраженный социолингвистический характер, что легко понять, например, при сравнении выражения категории количества средствами различных языковых культур. Э.Сепир, рассматривая вопросы *How much?*, *How many?*, *How far?*, *How long?*, относит их к области имплицитного градуирования¹, поскольку в них присутствует точка отсчета — *much*, *many*, *far*, *long*. В русскоязычных вариантах этих вопросов такое градуирование отсутствует, они имеют нейтральную форму *сколько?*

Этому же посвящено исследование В. И. Хайруллина. Понятие антиципации связано с понятием эксплицитности — большей или меньшей детализации описания. Детализация может выражаться не только введением в смысловую структуру описания дополнительных деталей или характеристик материального объекта, но и за счет конкретизации, которая достигается при использовании таких единиц именованности, в значениях которых фиксируются разные стороны описываемого объекта. Стремлением к экономии выразительных средств языка определяется употребление более общих характеристик при описании материальных объектов².

На терминологическом уровне асимметрия объемов понятий вариантов терминов на разных языках может быть продемонстрирована на примере термина «манипулятор». В русском языке он воспринимается в широком смысле, включающем два понятия: исполнительное устройство автоматического механизма и собственно робот. Присвоение второго понятия имеет социальные корни. Анализ литературы по робототехнике в сопоставлении с технической политикой и методами социального воздействия на эту политику показывает, что русский язык научно-технической литературы до конца 70-х гг. XX в. избегал термина «робот» за счет расширения объема понятия «манипулятор». В английском языке между этими терминами, как и между самими понятиями, существуют определенные границы и семантические связи. Японский язык следует инициативе английского языка.

При переводе с русского языка на английский вследствие частичной искусственной синонимичности русских вариантов терминов «робот» и «манипулятор» часто допускается ошибка, связанная с расширением понятия «робот» по аналогии с английским термином. В японском языке социальный фактор проявляется в другом аспекте. Оба варианта терминов («робот» и «манипулятор») пишутся катаканой, чем подчеркивается отношение к этим словам как к заимствованным. Такое же отношение распростра-

¹ См.: Сепир Э. Градуирование: Семантическое исследование // Лингвистическая прагматика: Новое в зарубежной лингвистике. — М., 1985. — Вып. 16. — С. 56.

² См.: Хайруллин В.И. Перевод научного текста: Лингвокультурный аспект. — М., 1992. — С. 61.

ется и на оценку объема понятия, а потому ошибочных переводов этих терминов с японского языка не обнаружено.

Убедительным примером разности восприятия объемов понятий является термин «поколение» — «generation». В русском языке соблюдается традиция восприятия обязательной динамичности процессов и явлений, описываемых термином «поколение». Считается, что одно поколение обязательно и чаще всего безвозвратно сменяет другое. Некоторые словари даже дают временные рамки этого понятия — до 30 лет. Эта традиция, оправданная в социальном плане, ошибочно экстраполируется на терминологический уровень. В английском языке вариант «generation» в основном выступает в атрибутивной функции при установлении видовых отношений. В частности, в подязыке робототехники он служит смыслоразличительным элементом классификации роботов от механических устройств для автоматического выполнения простых операций (роботы первого поколения) до сложнейших логических устройств (роботы четвертого поколения). Эта ситуация может оказаться источником переводческих ошибок. Одновременно шести переводчикам была предложена для перевода на английский язык простейшая фраза: «Фирма “Фанук” представила на весеннюю ярмарку свою новейшую модель робота первого поколения F-V411». Разность восприятия объема понятия «поколения» создала проблему перевода термина «робот первого поколения». Три переводчика автоматически внесли ошибку в перевод, написав «робота четвертого поколения», полагая, что они исправили ошибку оригинала. Еще один переводчик, дав правильный перевод, все же сделал сноску («Видимо, ошибка в тексте — имеется в виду робот четвертого поколения»). На самом деле ошибки здесь нет, и в промышленности по сей день разрабатываются и эксплуатируются одновременно роботы всех поколений в зависимости от назначения и сложности выполняемых операций.

Различие языковых средств выражения. В основе разности языковых средств выражения лежит глубинное различие самих языков и языковых традиций. Важным фактором, обуславливающим асимметрию терминосистем, являются национально-государственные особенности специальной сферы. Вследствие интернационального характера современной научно-технической революции такая ситуация распространена в технических сферах человеческой деятельности. Например, причиной асимметрии терминологии военно-технической сферы оказываются национально-государственные особенности, определяемые, с одной стороны, вообще национально-государственной спецификой военного дела, и с другой — известным стремлением каждого государства скрывать от вероятных противников, а иногда и союзников, имеющиеся достижения или провалы.

В этой части исследования был проведен эксперимент на материалах британского и американского вариантов английского подъязыка маркетинга. Такой выбор материала был вызван стремлением как можно дальше, а следовательно, и убедительнее развести между собой лингвистические и экстралингвистические факторы. Было выбрано 720 терминов, функционирующих в параллельных текстах, подготовленных в разных языковых вариантах. Все термины имели дефиниции, на основе которых удавалось совмещать варианты. Фактически в пределах одного языка обнаружилось функционирование двух вариантов терминосистемы. Было зарегистрировано устойчивое несовпадение 26,39 % вариантов терминов, т. е. из 720 терминов, использованных в британском и американском вариантах английского языка, 190 терминов не совпали по форме, но полностью соответствовали друг другу по объему понятия, что было установлено на основе совпадения дефиниций. Например:

1. Дефиниция: установление одним предприятием контроля над другим посредством получения всего или большей части его акционерного капитала.

Использованные варианты термина:
американский — *acquisition*;
британский — *take-over*,
русский — *получение*.

2. Дефиниция: любой документ, представляющий права, которые обычно могут быть использованы только лицом, владеющим этими документами.

Использованные варианты термина:
американский — *negotiable document*;
британский — *negotiable instrument*;
русский — *оборотный документ*.

Около 2 % терминов исследованного массива составили термины, функционирующие в американском варианте английского языка и отсутствующие в британском. Например:

1. Дефиниция: финансовая уступка с учетом распределенных прибылей в соответствии с владением одной компанией более чем установленной процентной ставкой акций по сравнению с другой компанией.

Использованные варианты термина:
американский — *dividends received deduction*;
русский — *уступка по полученным дивидендам*.

2. Дефиниция: продажа ценных бумаг, векселей или долгосрочных займов, препятствующая перекупке продавцами и перепродаже покупателями в конце определенного периода.

Использованные варианты термина:
американский — *selling and repurchase agreement*;

русский — *соглашение о купле-продаже ценных бумаг.*

3. Дефиниция: документ, выпущенный владельцем товарного склада, лицензирующий торговцев, размещающих товары на этом складе.

Использованные варианты термина:

американский — *negotiable warehouse receipt*;

русский — *оборотное право владельца на депозит; право владельца депозита.*

Исследование, проведенное на материалах одного языка, функционирующего в различных социальных условиях, отличается достаточной чистотой для утверждения вывода о том, что социальные условия определяют вариантно-инвариантные особенности формирования терминосистемы.

Отсутствие определенного объема терминов в одних условиях по отношению к другим в данном случае связано с такими факторами, как консервативность кредитно-валютной системы в одной стране и быстрой сменой стиля финансовых отношений в другой. Но тем не менее это никак не связано с лингвистическими проблемами. И даже более значительные различия в терминологическом составе (более 25 %) так же могут быть объяснены экстралингвистическими факторами. При сохранении единой языковой среды отличия состоят в основном в средствах выражения понятия с использованием при терминообразовании разночастотной лексики и различных аффиксов. В проанализированном словарном материале не было обнаружено специфических вариантов синтаксиса; общие грамматические нормы в основном соблюдались.

Различие между формами изложения и восприятия. Текст представляет собой модель двух видов коммуникативной деятельности — первичной и вторичной. Вместе с тем их отражение в тексте, моделирование, осуществляется не отдельно и не параллельно, а вместе и слитно. Поэтому в соответствии с коммуникативной сущностью текста следует говорить о сопряженной модели коммуникативных деятельностей¹. Вместе с тем в ситуации двуязычной коммуникации, если ее рассматривать как макроструктуру, текст-модель имеет специфическую форму реализации — он существует в двух вариантах: как модель первичной коммуникативной деятельности (текст оригинала, т. е. текст адресанта) и как модель вторичной коммуникативной деятельности (текст перевода).

Факт физического существования двух текстов никак не может противоречить отношению к тексту как к сопряженной модели коммуникативных деятельностей, поскольку текст-оригинал и текст-перевод — это варианты единого и необходимого компо-

¹ Подробнее см.: Сидоров Е. В. Основы системной концепции текста: автореф. дис. ... д-ра филол. наук. — М., 1998.

нента коммуникации, а именно средства осуществления конкретной задачи человеческой деятельности. Вариантность текста оказывается единственной возможностью координации человеческой деятельности: непонятый текст не может корректно управлять деятельностью адресата.

Попытки рассматривать двуязычную коммуникацию на уровне микросистем или ее подсистем связаны с независимым подходом к тексту-оригиналу и тексту-переводу, что упрощает исследование каждого в отдельности, но исключает возможность системного подхода к пониманию коммуникативного процесса.

Как любое инструментальное средство, текст порождает различные результаты деятельности адресатов в зависимости от их собственных субъективных свойств. Содержание разных текстов никогда не будет одинаковым. Объективность этого вывода распространяет его на любые сравнимые тексты и подтверждает его, когда речь идет о текстах на разных языках.

Однако общность содержания оригинала и перевода обеспечивает тождественность текстов, достаточную для рассмотрения и сопоставления их как одного варианта относительно другого. Содержательная сущность вариантов текстов основана на том, что обязательным компонентом коммуникативного взаимодействия является действительность как объект отражения, познания и преобразования. В информационной технологии наибольший интерес для нас представляет та часть действительности, которая отражается конкретным набором таких единиц текста, как термины.

Подход к тексту как к инструментальному средству прикладного значения для коммуникативного взаимодействия позволяет обнаружить, объяснить и решить некоторые терминологические проблемы. Прикладной характер этого средства необходимо рассматривать в двух аспектах его соотнесенности:

- с объективными научно-культурными традициями изложения содержания;

- с объективными особенностями восприятия содержания.

В условиях двуязычной коммуникации для текста оригинала значение имеет в основном первый аспект, для текста перевода — второй. При переводе с английского языка на русский нас интересуют англоязычные научно-культурные традиции изложения содержания на терминологическом уровне и особенности восприятия заданного англоязычным оригиналом содержания на русском языке.

Научно-культурные традиции изложения содержания на терминологическом уровне определяются рядом параметров, из которых для нас важны те, которые отражают особенности языкового и научно-содержательного планов. Специфика таких параметров проявляется в организации смысловой структуры текста, описывающего материальный объект и его действия.

Материальность — одно из важнейших понятий действительности. Она воплощается в объекте, который в научном и техническом тексте может быть описан термином.

В связном тексте можно наблюдать все явления функционирования лексики. Практическая текстовая деятельность последнего времени все более характеризуется многоязычностью. Количество аутентичных текстов постоянно увеличивается. Самым простым примером таких тенденций являются сопроводительные инструкции для продукции каких-либо фирм, поступающих на многочисленные рынки всего мира. Этот уже не требующий доказательства факт повышает актуальность использования многоязычных словарей.

Различные виды практической текстовой деятельности (перевод, преподавание и изучение иностранного языка, составление дву- и многоязычных словарей) всегда основывались на значительной доле сопоставительной работы. Можно считать, что сопоставление — это метод обучения переводу как виду деятельности и как результату сопоставительной работы. Масштабные лексикографические исследования на материалах параллельных текстов ведутся во Всероссийском центре переводов научно-технической литературы и документации.

5.7.2. Типологизация параллельных текстов

Все сказанное приводит к выводу о том, что параллельные тексты могут служить для внутриязыковых сравнений на всех уровнях языка с использованием различных средств и способов. В параллельных текстах можно выделить три основных типа, что соотносится с тремя компонентами контрастивной текстологии — прагматикой, семантикой и синтагматикой. В типологизации параллельных текстов учитываются и особенности дискурса, и критерии эквивалентности, но тем не менее разделение на эти три типа остается весьма условным (участвующие в этом процессе взаимодействия показаны в табл. 5).

Тексты класса А обычно являются результатом полного профессионального перевода, полученного в процессе сознательной аппроксимации, в котором оригинальное сообщение на передающем языке превращается в ситуационно оправданный текст на языке перевода. Здесь обычно не принимаются в систематический учет различия переводческих традиций, существующих в каждом языке и в каждой культуре, хотя такие различия поддаются выявлению, так как письменный перевод — процесс однопользовательный.

Тексты класса Б, как правило, представляют собой результат предметной адаптации сообщения в соответствующих соотношениях других языков с целью передачи идентичного сообщения

Предпринятые действия	Типология параллельных текстов		
	А	Б	В
Достижение эквивалентности	Перевод	Адаптация	Равенство контекстов
Достижение согласованности (параллельности)	Апроксимирование	Соотнесенность с общим источником	Сравнение
Направленность действий	Передающий язык → язык перевода	Одновременное (согласованное) формулирование	Независимая генерация
Характерные жанры	Художественная литература, технические инструкции, устные переводы переговоров	Рекламные тексты одного товара для покупателей разных стран	Сообщения в газетах разных стран об одних и тех же событиях

адресату, находящемуся в весьма отдаленных лингвокультурных условиях. Соответствие между двумя текстами обусловлено не направлением процесса или формальными параметрами дискурса оригинала, а необходимостью достичь аналогичной реакции адресата.

Тексты класса В могут не иметь, вообще, или иметь условное соотношение на уровне экспертной оценки того, что оба они, порожденные одними обстоятельствами, имеют случайные совпадения, производящие одинаковый эффект.

Нельзя не заметить, что на такую классификацию, пожалуй, самое сильное влияние оказывает прагматический компонент как текста на передающем языке, так и текста на языке перевода. Особенно это отражается на текстах класса А, но в различной степени справедливо и для случаев Б и В.

Исследовательское преимущество параллельных текстов создается тем, что они документируют контрасты между типами дискурсов в пределах одного языка и между языками и дают фактический материал для систематизации эквивалентов в разных языках.

Изучение синтаксических параметров важно, когда мы хотим охарактеризовать линейную прогрессию дискурса в терминах связей между последовательными компонентами, например, развитие повествовательных структур в терминах событий, участников,

установок и связанных с ними лексико-грамматических реализаций. Контрастивный синтаксис текста распространяется и на уровень дискурса и показывает, как различаются языки по таким параметрам, как введение в основной текст (подзаголовки), последовательное развитие (анафорические ссылки, связи) и т. п. Все эти параметры можно просмотреть в сопоставительном режиме по параллельным текстам.

Семантические параметры текста играют значимую роль в обозначении пути, по которому референтная информация распределяется между последовательными элементами текста или локализуется в одном из элементов.

Все примеры исследований на параллельных текстах подтверждают тот факт, что значение и смысл отдельных элементов складываются в дискурсе по законам его прагматических, семантических и синтагматических компонентов.

Сопоставительный анализ является основным источником формирования многоязычных словарей.

5.8. Коммуникативные условия функционирования терминосистемы

5.8.1. Многоязычие в переводной лексикографии

В традиционной лексикографии проблема многоязычия переводных словарей решается в зависимости от их типа и назначения. В целом преобладает тенденция разработки двуязычных переводных словарей, поскольку словари общей лексики с тремя и более языками имеют меньшую описательную и порождающую мощь. Количество содержащихся в них словарных статей обычно невелико, внимание уделяется только отдельным словам, многозначные слова представлены лишь в одном, максимум двух значениях, крайне ограничен и выбор переводных эквивалентов; недостаточно внимания уделяется такой информации, как, например, фонетическая транскрипция, грамматические сведения, иллюстративный материал и т. п.

Можно сделать вывод, что чем больше традиционный переводной словарь объединяет разных языков, тем меньше он включает различных лексических единиц, тем беднее его словарные статьи. Весьма часто традиционные многоязычные словари общей лексики представляют собой лишь скелетные списки слов и мало пригодны для решения большинства задач, встающих при обработке текстов на иностранных языках. Такое положение с многоязычными словарями в общей лексикографии объясняется, на наш взгляд, прежде всего тем, что объемы значений слов, считающихся переводными эквивалентами друг друга, весьма часто в

разных языках различны, а также трудностями установления переводных соответствий между разноязычными единицами, меньшими чем слово, т. е. словозначениями, из-за жесткости структуры подобных словарей. В целом можно сказать, что многоязычие нехарактерно для традиционных словарей общей лексики, и, вероятно, поэтому оно не используется в качестве различительного признака при построении типологий нетерминологических словарей.

Иное положение с многоязычными словарями в терминологической лексикографии. Хотя критические замечания относительно бедности словарных статей, высказанные в адрес многоязычных словарей общей лексики, в равной степени относятся и к многоязычным терминологическим словарям, тем не менее такие словари — довольно частое явление в терминологической лексикографии.

Доля многоязычных словарей общей лексики значительно меньше доли многоязычных терминологических словарей. Причины этого следует искать в первую очередь на лингвистическом уровне. Высокая доля многоязычных терминологических словарей, по нашему мнению, объясняется тем, что разноязычные лексические единицы, входящие в их состав — термины или номены, — опираются на одни и те же понятия либо объекты. Этим снимается основная трудность создания многоязычных словарей общей лексики, а именно несоответствие объемов значений лексических единиц разных языков.

Объекты описания вычислительной лексикографии четко распадаются по многоязычию на две категории. С одной стороны, это — автоматические переводные словари, а с другой — системы МП. Одно из основных отличительных свойств крупных промышленных автоматических переводных словарей — многоязычие.

Объединяет автоматические словари и терминологические банки не только многоязычие само по себе, но и выбор конкретных языков. Все основные словари включают английский, русский, немецкий и французский языки, т. е. признанные в мире языки научно-технической литературы. Ведущую роль занимает английский язык, на который приходится основная доля научно-технических публикаций. Вместе с тем отдельные исследователи отмечают некоторое уменьшение в последнее время доли английского языка как главного мирового языка науки и техники и появление все большего количества публикаций на труднодоступных для широкого круга ученых и специалистов языках. Особенно быстро растет объем научно-технических публикаций на японском языке, однако сложности ввода в ЭВМ иероглифической информации пока сдерживают широкомасштабное включение этого языка в многоязычные автоматические терминологические банки и словари за пределами Японии.

Основополагающим принципом построения многоязычного автоматического переводного словаря является способ организации связей между разноязычными лексическими единицами. В промышленных словарях, вне зависимости от количества используемых языков, применяются три способа организации межлексических связей.

1. Любая лексическая единица описывается для каждого языка отдельно (сведения о заголовке, грамматическая информация, толкование, все виды лингвистической и другой информации) и только средствами данного языка. Словарная статья обычно состоит из многих зон, каждая из которых хранится независимо от других. При таком подходе появляется возможность организовать весьма обширную и гибкую сеть связей между различными видами информации в пределах описания лексики одного языка, что позволяет заметно увеличить круг задач, решаемых этим словарем.

2. Для двух разноязычных лексических единиц, являющихся переводными эквивалентами друг друга, создается общая словарная статья. Подобные двуязычные статьи затем могут объединяться в многоязычные. Данного принципа организации межлексических связей придерживаются разработчики автоматических переводных словарей ЕВРОДИКАТОМ и ЛЕКСИС.

3. Результатом такого объединения становится единая словарная статья, содержащая переводные эквиваленты на всех языках, включенных в словарь. Однако не все словарные статьи имеют полный набор переводных эквивалентов на всех языках. В некоторых словарях в случаях, когда в статье отсутствует точный переводной эквивалент на каком-либо языке, вместо него дается толкование его значения из общей словарной статьи.

Многоязычие, столь характерное для крупных автоматических переводных словарей, не свойственно промышленным системам МП. Однако следует отметить, что в последние годы становится все более заметной тенденция к созданию многоязычных систем МП. В силу ряда экономических и политических причин проблема многоязычия занимает все большее место в машинном переводе в широком понимании этого термина.

Система МП СИСТРАН построена как двуязычная система. В настоящее время эксплуатируются несколько версий этой системы для переводов с русского языка на английский, с французского на английский, с английского на немецкий и др., но все они являются разновидностями одной и той же системы, аналогичны друг другу по структуре и эксплуатируются независимо друг от друга, т. е. все версии СИСТРАН двуязычны.

В системе МП ПРОМТ используется модульный принцип построения. Имеющаяся в нашем распоряжении и анализируемая в данном учебнике версия ПРОМТ включает англо-русский, французско-русский, испанско-русский и русско-английский модули.

Программное обеспечение ориентировано на обработку входного языка аналитического типа, например английского. При переходе к входному языку флективного типа, например русскому, требуется модификация лингвистического и программного обеспечения. Программы, работающие с выходным словарем, предусматривают морфологический синтез единиц языка флективного типа. При переходе к обработке лексических единиц языков аналитического типа достаточно только модификации лингвистического обеспечения. Таким образом, данная система МП, как и СИСТ-РАН, является двуязычным модулем.

На фоне двуязычия неинтерактивных промышленных систем МП и традиционных словарей общей лексики, с одной стороны, и разнообразия многоязычных терминологических словарей — с другой, в развитии автоматических переводных словарей проявляется тенденция к увеличению в них числа языков. Эта тенденция развивается в двух направлениях:

- увеличение количества разных языков в крупных промышленных автоматических словарях, которое осуществляется либо путем формирования новых языковых пар по старым или новым тематикам, либо путем добавления лексики одного нового языка к уже имеющимся лексическим массивам;

- создание новых двуязычных сменных блоков в языковых процессорах и интерактивных системах МП.

Многоязычие автоматических переводных словарей можно объяснить экстралингвистическими и лингвистическими причинами. На экстралингвистическом уровне это явление мотивировано тем, что увеличение количества разных языков расширяет сферу применения словарей и круг пользователей и соответственно делает их более рентабельными. На лингвистическом уровне это может быть обусловлено несколькими причинами. Во-первых, такие словари ориентированы на диалог с человеком и поэтому не нуждаются в блоках синтаксического и семантического анализа входного языка и синтеза выходного языка, поскольку эти задачи остаются для человека, а ведь именно здесь кроются основные различия между языками. Во-вторых, по своему составу автоматические переводные словари приближаются к традиционным терминологическим словарям, поэтому проблема обработки разноязычной лексики для большей части состава автоматического переводного словаря сильно упрощается. В-третьих, материальной базой автоматических словарей является ЭВМ, которая позволяет преодолевать некоторые трудности общей лексикографии: например, она может не сохранять словарную статью многозначного слова целиком, а при необходимости синтезировать ее из отдельных словозначений этого слова, которые содержатся в ее памяти отдельно.

В целом можно предположить, что в ближайшем будущем тенденция к росту многоязычия крупных автоматических словарей

сохранится, хотя формы и способы ее реализации **в зависимости** от типа словаря могут и должны быть разными.

5.8.2. Обратимость переводного словаря

Под обратимостью переводного словаря мы понимаем возможность произвольно, по желанию пользователя, менять входной и выходной языки местами. Такой рабочий прием довольно часто используется в лексикографической практике для расширения целевой направленности словаря и круга его пользователей. Учитывая, что лексика как объект описания автоматических и традиционных словарей едина, мы рассмотрим пути достижения обратимости в традиционной лексикографии и в автоматических переводных словарях.

По определению, обратимый переводной словарь должен иметь два разноязычных входа. Если мы попытаемся смоделировать обратимый англо-русский и русско-английский словарь на базе какого-либо конкретного англо-русского словаря, то это означает, что в одном словаре два разноязычных входа, в данном случае английский и русский, будут иметь один выход — одну объяснительную часть, в нашем примере — описание содержания лексики английского языка. Однако этим будет нарушено требование семантической эквивалентности словника и объяснительной части словаря, ибо в общем виде не может быть семантической эквивалентности между русскими словами и описанием содержания английских слов, даже если между самими словами существуют отношения переводной эквивалентности.

Значение слова как объект описания в словаре представляет собой составную сущность, включающую в себя абсолютную ценность (информацию об обозначаемом внеязыковом факте), относительную ценность (информацию о способности слова вступать в парадигматические отношения с другими словами, о его стилистической и статистической значимости, внутренней форме и др.) и сочетательную ценность (информацию о способности слова сочетаться определенным образом с определенными словами).

Переводной эквивалент в словарной статье фиксирует относительное равенство абсолютных ценностей двух разноязычных слов. Вся остальная информация словарной статьи — это сведения об относительной и сочетательной ценностях заголовочного слова, которые чаще всего не идентичны соответствующим ценностям его переводного эквивалента, так как они глубоко индивидуальны и существенно варьируют от слова к слову в пределах даже одного языка.

Межъязыковые различия оказываются еще большими. При описании одноязычных слов или лексико-семантических вариантов

многоязычных слов они весьма существенны, но еще более увеличиваются при описании многозначных слов как единого целого. Поэтому столь несхожи друг с другом словарные статьи слов, связанных между собой отношением переводной эквивалентности, в разнонаправленных переводных словарях общей лексики, например в англо-русском и русско-английском словарях, которые в целом сопоставимы и по объему, и по степени подробности описания значений входных лексических единиц.

Рассмотрим конкретный пример. В словарной статье слова *долг* (в значении «взятое взаймы, одолженное») в рассматриваемом русско-английском словаре приведен его английский переводной эквивалент *debt*, а также 9 русских словосочетаний и для них 13 английских переводных эквивалентов, из которых только 3 зафиксированы в словарной статье *debt* англо-русского словаря Мюллера. В то же время в словарной статье *debt* этого словаря даны 10 английских словосочетаний и для них 8 русских переводных эквивалентов, из которых только 2 зафиксированы в словарной статье *долг* русско-английского словаря. На этом примере видно, что различия в описании сочетательных ценностей двух слов, являющихся переводными эквивалентами друг друга, в этих словарях весьма значительны. Подобные примеры можно было и продолжить, однако мы считаем его достаточным для того, чтобы показать основные трудности, которые возникают при построении традиционных обратимых переводных словарей общей лексики.

В переводных словарях наблюдается тенденция, которая проявляется в том, что чем подробнее описание, чем богаче словарная статья, тем больше расхождений между словарными статьями разноязычных слов, являющихся переводными эквивалентами друг друга. Особенно это заметно на примере словарных статей многозначных слов. Достичь обратимости словаря общей лексики можно за счет обеднения содержания словарной статьи и сведения его фактически к одним переводным эквивалентам, или использования принципиально новой структуры для построения всего словаря в целом.

В качестве одного из немногих примеров обратимого словаря с новой структурой можно привести «Словарь усилительных словосочетаний» И. И. Убина, который может использоваться как одноязычный словарь сочетаемости отдельно в пределах русского и английского языков или как переводной англо-русский и русско-английский словарь сочетаемости с учетом взаимных отсылок заглавных слов. Особенность этого словаря в том, что описание сочетаемости (в данном случае только усилительных словосочетаний) русских заглавных слов дается средствами русского языка, английских — средствами английского языка. Содержание словарных статей русских заглавных слов неидентично содержанию сло-

варных статей их английских переводных эквивалентов. Обратимость достигается благодаря тому, что словарные статьи двух независимых одноязычных словарей связаны между собой взаимными отсылками, причем отсылки адресованы не к отдельному словосочетанию, а к словарной статье в целом.

Стремление к экономии и расширению целевой направленности словаря приводит к тому, что лексикографы все больше внимания уделяют разработке различных методов и практических приемов для преобразования традиционных необратимых переводных словарей в обратимые. Наиболее широко принцип обратимости используется в терминологической лексикографии. Составители терминологических словарей используют различные способы обеспечения дополнительных входов в свои словари. В некоторых многоязычных словарях один из языков является основным, как бы стержневым, вокруг которого группируются другие языки. В словарной статье такого словаря заголовком служит лексическая единица стержневого языка, а лексические единицы остальных языков приводятся как ее переводные эквиваленты. В качестве приложений к основной части словаря имеются алфавитные списки лексических единиц всех языков, кроме стержневого, с отсылками к основным словарным статьям. В других многоязычных словарях материал представлен в виде двуязычных словарей для всех возможных языковых пар. В целях экономии в некоторых случаях в корпусе словаря каждой словарной статье присваивается цифровой код, на который и дается отсылка в указателе. Весьма рационально задача обратимости решается в словарях, в которых на каждой странице основного корпуса дается нумерация строк, а в указателе производится отсылка к нужной странице и строке.

Возможность обратимости в терминологической лексикографии можно объяснить, на наш взгляд, некоторыми особенностями терминов, которыми они обладают по сравнению с общеупотребительной лексикой.

Терминологические словари в основном ограничивают свою задачу описанием абсолютной ценности слова. На практике это проявляется в том, что такая словарная статья практически не содержит никакой другой информации, кроме переводных эквивалентов. Проблема многозначности и омонимии в терминологии (в особенности в пределах узких областей науки и техники) стоит не столь остро, как в общеупотребительной лексике. Составители предпочитают не включать в свои словари многозначные слова либо включают их в одном или двух значениях, которые считают основными для выбранной тематики.

При таких условиях построение обратимого «бумажного» словаря не представляет серьезных лексикографических и технических сложностей и не приводит к навязыванию законов выходного языка лексике языка входного. Вместе с тем терминологические

словари, построенные по принципу словарей интенсивного типа, т.е. с богатой словарной статьей и глубоким описанием омонимии и многозначности, как и крупные словари общей лексики, свойством обратимости не обладают.

Традиционные обратимые переводные словари имеют ряд качеств, которые существенно отличают их от однонаправленных переводных словарей. К нам можно отнести:

1) четко очерченную и узкофункциональную направленность (словари учебные, лексической сочетаемости, терминологические);

2) наличие в словарных статьях описания только одного компонента значения заглавной единицы, чаще всего представляющего абсолютную или сочетательную ценность;

3) составление описания значения заглавного слова или его компонента не с помощью другого языка (выходного), а его же собственными языковыми средствами (средствами входного языка), причем снимается или, точнее сказать, существенно уменьшается зависимость одного языка от другого.

Резюмируя сказанное о свойствах однонаправленных и обратимых переводных словарей, можно сделать вывод, что если описание значения лексической единицы выполнено достаточно подробно и отражает ее абсолютную и сочетательную ценность как единого целого, то практически невозможно добиться идентичности содержания словарных статей в традиционных «бумажных» разнонаправленных переводных словарях одной и той же пары языков. С общетеоретической точки зрения это можно объяснить тем, что описание в двуязычном словаре лексики одного (входного) языка осуществляется с учетом особенностей другого (выходного) языка, который налагает свои, свойственные только ему, ограничения и на описания значений входных единиц. С точки зрения межъязыкового общения, правила которого зафиксированы в двуязычном словаре, входной язык можно считать аргументом, а выходной язык — функцией. Если продолжить это математическое сравнение и поменять местами входной и выходной языки (что происходит в обратимых словарях, т.е. сделать выходной язык аргументом, а входной язык функцией), то мы получим другую функцию, отличающуюся от первой, что подтверждает вывод о принципиальном неравенстве содержания словарных статей в разнонаправленных переводных словарях одной и той же пары языков. Это означает, что в условиях традиционной лексикографии теоретически невозможно построить полностью обратимый переводной дву- или многоязычный словарь с подробным описанием всех аспектов значений лексических единиц входного и выходного языков. Реальное положение в традиционной переводной лексикографии подтверждает этот вывод.

Следует отметить, что обратимость словаря — понятие неоднородное, и реализуется оно по-разному в разных словарях. Обратимость автоматического переводного словаря может быть полной и частичной. При полной обратимости каждая лексическая единица любого входящего в состав словаря языка может выступать и в качестве входной единицы (при запросе), и в качестве выходной (при ответе) для любой теоретически возможной в рамках данного словаря пары языков. Однако на практике полная обратимость реализуется только в двуязычных автоматических переводных словарях и в немногих традиционных терминологических словарях малого объема. Крупные многоязычные автоматические словари, такие как ЕВРОДИКАТОМ и ЛЕКСИС, являются частично обратимыми. Это подтверждается тем, что объемы разноязычных лексических массивов в одном и том же словаре весьма различны. При создании, пополнении и коррекции больших лексических массивов, исчисляющихся сотнями тысяч словарных статей, практически невозможно добиться того, чтобы каждая статья имела переводные эквиваленты на всех включенных в переводной словарь языках.

Обратимость автоматического переводного словаря четко связана с уровнем его многоязычности. Чем больше в нем языков, тем сложнее достичь полной обратимости. Конкретные пути достижения обратимости автоматических переводных словарей следует искать в способах организации межлексических связей. В словаре с одноязычными словарными статьями обратимость достигается за счет установления жестких связей между заголовками разноязычных словарных статей, что теоретически не противоречит принципу полной обратимости. Для этого выделяется так называемый стержневой язык, а связи между лексическими единицами всех входящих в словарь языков осуществляются через лексику стержневого языка или языка-посредника.

В автоматических переводных словарях с двуязычными словарными статьями обратимость достигается тем, что такая словарная статья имеет двуединый заголовок, содержащий лексическую единицу одного языка и ее переводной эквивалент на другом. В автоматических переводных словарях с многоязычными словарными статьями обратимость происходит благодаря тому, что словарные статьи хранятся в памяти ЭВМ целиком и в таком же полном составе выдаются по запросу любой входящей в ее состав лексической единицы.

Все крупные промышленные автоматические словари и терминологические банки данных являются частично или полностью обратимыми. По этой характеристике они четко противостоят крупным переводным словарям общей лексики и промышленным системам МП, но имеют много общего с традиционными переводными терминологическими словарями.

5.8.3. Гибкость автоматического переводного словаря

Под гибкостью автоматического переводного словаря И. И. Убин понимает такие его свойства, которые упрощают обращение к словарю, расширяют возможности пользователя при работе с иноязычным текстом, снижают требования к уровню знания входного языка, а также позволяют использовать данный словарь не только для хранения лексических массивов, но и для решения других задач автоматической обработки текстов и автоматизации лексикографических работ.

Наиболее важным средством повышения гибкости словаря является упрощение формы запроса, т. е. обращения пользователя. В традиционных переводных словарях форма обращения совпадает со словарной формой, т. е. с формой хранения в них лексических единиц. Словарная, каноническая, исходная форма слова — одна из словоформ, традиционно выбираемая в качестве представителя данной лексемы в словаре, например инфинитив для глагола, именительный падеж единственного числа для существительного в русском языке.

Такой выбор формы хранения лексических единиц предъявляет достаточно высокие требования к уровню знания входного языка. Следовательно, для значительной части пользующихся двуязычным словарем для перевода как с иностранного, так и с родного языка необходима информация о грамматических особенностях отдельных конкретных слов. Установление индивидуальных особенностей слов по их грамматическому описанию — путь дедуктивный, требующий от читателя либо определенных предварительных знаний, либо переработки большого объема информации. Наиболее удобно было бы предоставлять в переводном словаре грамматическую информацию при слове, однако обычно все словоформы при словах не даются. Считается, что если словоформы образованы по общему правилу, то пользователю будет несложно представить всю парадигму. Но это все же иногда приводит к тому, что некоторые читатели оказываются не в состоянии эффективно работать с иноязычной частью словаря.

Трудности определения словарной формы слова возникают даже в тех случаях, когда входной язык является для пользователя родным, но их становится больше при работе с неродными языками. Особенно остро эта проблема стоит для языков синтетического типа. Для русского языка наибольшие сложности связаны с обработкой видовых форм глагола, причастий, супплетивных форм существительных, приставочных форм глагола. Для английского языка характерны трудности при обработке неправильных глаголов. Все это усугубляется еще и тем, что несловарные формы слова встречаются в текстах значительно чаще, чем словарные. Экспериментально было установлено, что русский и французский

глаголы реализуются в текстах в словарной форме в среднем в пять раз реже, чем в несловарных, русские существительные в словарной форме в текстах встречаются почти в три раза реже, чем в несловарных¹.

Составители традиционных переводных словарей предпринимают определенные шаги для увеличения возможностей входов в словари. Так, для английского языка обычным стало включение в состав словаря нерегулярных грамматических форм глаголов, существительных и прилагательных с отсылками к их словарным формам. Эту же цель в основном преследуют грамматические очерки входных языков, которые имеются во многих переводных словарях общей лексики. Однако все эти приемы только облегчают пользователю сведение возможных словоформ к словарной, но в основном задачу поиска он должен решать сам. Для языков с развитым словоизменением, например для русского, проблема определения словарной формы слова оказывается настолько серьезной, что делает практически невозможной работу с русско-иноязычным переводным словарем без основательного знания грамматической системы родного языка. В свою очередь, это приводит к тому, что тексты на русском языке остаются недоступными иностранному пользователю словаря, если он не имеет специальной языковой подготовки.

В целом можно сказать, что выбор в традиционных словарях только одной словоформы в качестве представителя всей лексемы является вынужденным. Такой путь для лексикографа достаточно удобен и представляет собой экономное решение многих лингвистических проблем; но вместе с тем он резко ограничивает количество входов в словарь, что в свою очередь снижает его ценность, так как качество словаря, помимо других его характеристик, напрямую зависит от количества входов.

Все промышленные системы МП в основном предназначены для перевода связных текстов и в своем составе имеют алгоритмы морфологического анализа, которые обрабатывают все встречающиеся в текстах словоформы. В связи с этим проблемы определения словарной формы слова для систем МП не существует. По этой характеристике они превосходят и традиционные, и автоматические словари.

А вот для автоматических словарей эта проблема весьма актуальна. Стремление разработчиков решить ее обусловлено несколькими причинами. Первое — желанием облегчить пользователю общение со словарем, снизить уровень требований к знанию входного языка — для этого необходимо приблизить форму работы

¹ См.: Убин И.И. Критерии выбора формы представления лексики в терминологических банках данных и автоматических словарях // Системный анализ научного текста. — Владивосток, 1984. — С. 169.

пользователя с автоматическим словарем к естественному языковому общению. Второе — потребностью преобразовать автоматические словари из «складов» в «фабрики», т. е. превратить их во многофункциональные автоматизированные лексикографические системы, — для этого требуются алгоритмы распознавания всех текстовых форм лексических единиц и сведения их к словарным формам.

Создатели всех крупных промышленных систем единодушно признают важность решения этой проблемы в полном объеме, т. е. для всех словоформ всех грамматических классов. Вместе с тем трудности создания таких алгоритмов вынуждают некоторых разработчиков промышленных автоматических словарей либо отказываться от поисков решения проблемы лемматизации, либо решать ее частично.

Выбор формы запроса и соответственно формы представления лексических единиц в автоматическом переводном словаре зависит от ряда организационных, технических и лингвистических факторов. К лингвистическим факторам относятся категории обрабатываемых текстов, уровни языковой компетенции пользователей и особенности грамматического строя языка.

Функционально-коммуникативные характеристики, например, подъязыка робототехники определяются его относительной «молодостью». Проблема хронологических рамок, которой придают большое значение многие исследователи подъязиков специальных областей, в данном случае существенной роли не играет: этот подъязык развивался в основном в последние десятилетия. Важно учитывать то, что зарождение предметной области, обслуживаемой подъязыком, приходится практически на начало периода новейшей истории. Как известно, сам термин «робот» впервые появился в 1920 г. Кроме того, подъязык робототехники активно расширился за последние 40—50 лет, когда влияние научно-технического прогресса модернизировало и язык повседневного общения. Конечно, максимальному воздействию подверглись стилистические регламентации, имеющие не самое решающее значение для формирования характеристик специальных подъязиков.

Основным объектом описания и обработки в автоматическом переводном словаре является лексика, поэтому тексты, которые подлежат обработке с помощью такого словаря, должны прежде всего рассматриваться с точки зрения их лексического состава. С этих позиций тексты в самом общем виде можно разделить на открытые и закрытые.

Открытыми в лексическом отношении И. И. Убин считает тексты, в которых могут встречаться слова всех грамматических классов во всех или в большинстве теоретически возможных словоформ. К ним относятся монографии, научные статьи, публицистические материалы, различные учебные пособия и т. п.

Закрытыми в лексическом отношении являются тексты, лексика которых ограничена как по форме (например, только словарной формой), так и по содержанию (например, одной тематикой). В таком понимании закрытые тексты представлены каталогами деталей, перечнями предметов поставки, различной импортно-экспортной документацией. Их лексический состав, как правило, ограничен существительными или номинативными словосочетаниями в словарной форме, а для русского языка — в основном в именительном и родительном падежах. При обработке таких текстов возможна и целесообразна словарная форма запроса. При работе с открытыми текстами ее использование оказывается неэффективным, в особенности для языков с развитым словоизменением, например для русского языка.

Если автоматический переводной словарь не предоставляет пользователю возможности осуществлять запрос в текстовой форме, то, работая с открытыми текстами, он вынужден сам сводить словоформы к словарному виду, что в общем случае предполагает достаточно хорошее знание языка запроса. Если язык запроса для пользователя является родным, то это требование практически не ограничивает ни его возможностей в составлении заявок, ни вообще круг пользователей. Когда же запрос делается на неродном языке, а именно такова наиболее распространенная ситуация чтения, перевода, реферирования иностранного текста, такое требование часто оказывается трудно выполнимым для многих технических специалистов, что, естественно, резко снижает ценность данного словаря для этой категории пользователей. Принципиальное многоязычие большинства автоматических переводных словарей еще более усложняет рассматриваемую проблему.

Таким образом, если автоматический переводной словарь ориентирован на широкий круг пользователей, имеющих различные уровни языковой подготовки, то он должен обеспечивать возможность запроса в текстовой форме.

В настоящее время повсеместно наблюдается активный двунаправленный процесс универсальной трансформации языковой нормы. Развитие подязыков специальных областей деятельности влияет на язык повседневного общения, трансформируя его лексические, синтаксические и стилистические нормы. Этому способствуют и события общественной жизни, в которые вовлечены и представители специальных областей, на естественных правах привносящие особенности своих подязыков в коммуникативный процесс. На это направление универсальной трансформации языковой нормы следует ответная реакция, средствами которой трансформирующее воздействие по тем же каналам передается подязыкам специальных предметных областей.

Например, к моменту зарождения подязыка робототехники сформировались подязыки смежных предметных областей, уча-

ствующих в развитии отдельных направлений робототехники. На понятийном уровне большое значение для становления подязыка робототехники имела сильная интерференция таких областей, как вычислительная техника, автоматика и телемеханика. Их подязыки отличались сильной формализацией и органическим вживлением в их структуру элементов лексики и синтаксиса искусственных формальных языков. Лавинное воздействие этой терминологии оказало большое влияние на особенности коммуникативных процессов в подязыке робототехники. Коммуникативная направленность текстов по робототехнике обусловлена необходимостью высокой интенсивности обмена информацией (естественной для периода развития), динамичностью плана выражения, существенной долей экстралингвистических факторов.

Если под динамичностью плана выражения понимать краткость изложения при передаче адекватной информации и информационную емкость высказывания, то по этому параметру характеристики текстов по робототехнике уступают только аналогичным характеристикам математических текстов. Принято считать, что средняя длина научной или технической журнальной статьи (жанр, на который приходится более 75 % выполняемых переводов) составляет 5 000 слов, при этом средняя длина математических статей — 3 800 слов; статей по робототехнике — 4 100 слов.

В информационных технологиях имеется способ определения информационной емкости через терминологическую насыщенность, т. е. через отношение количества лексических единиц, составляющих термины предметной области, к которой принадлежит данный текст, к общему количеству входящих в него значащих лексических единиц. Средняя терминологическая насыщенность научных и технических статей по результатам анализа случайной выборки без учета предметной области составила 0,42; математических статей — 0,58; статей по робототехнике — 0,54.

Коммуникативная направленность текстов по робототехнике находится в функциональной зависимости от регламентирующих экстралингвистических факторов формирования ее подязыка — территориальной неравномерности развития данной предметной области и интеграции научной мысли для решения общих интернациональных проблем. Изучение влияния экстралингвистического фактора территориальной неравномерности развития предметной области обнаруживает важные закономерности. Дисбаланс уровней развития предметной области в стране исходного языка и в стране языка перевода инициирует и определяет направления взаимопроникновения лексических и синтаксических элементов подязыков.

В самых общих чертах этот процесс характеризуется большим количеством заимствований терминологии и сокращением периода безэквивалентности новой терминологии.

Грамматический строй языка, в частности соотношение словарных и несловарных форм слова в открытых текстах, безусловно должен учитываться при выборе оптимальной формы запроса и хранения лексической единицы в автоматическом переводном словаре.

И. И. Убиным было проведено исследование по определению соотношения словарных и несловарных форм слов в оригинальных (непереводных) текстах по вычислительной технике в английском, русском и французском языках. Для анализа были выбраны существительные, глаголы и прилагательные из всех рассмотренных текстов.

Исследование показало, что для русского языка характерно преобладание в открытом тексте несловарных форм слов для всех указанных частей речи, что естественно предопределяет выбор текстовой формы запроса для русского языка в словаре. Такое же положение с глаголами в английском и во французском языках. У существительных и прилагательных в выбранных текстах явно преобладала словарная форма. Полученные данные позволяют сделать вывод: выбор формы запроса и хранения лексических единиц в автоматическом переводном словаре должны определяться отдельно для каждого языка и даже для разных частей речи в пределах одного языка.

Весьма удобным и простым для языков аналитического типа оказалось использование метода отсекающих от ненайденных словоформ наиболее частотных окончаний, например, для английского языка *-s*, *-ed*, *-ing*. Для языков с развитым словоизменением, например для русского, этот метод неприемлем.

Для решения проблемы лемматизации в промышленных автоматических переводных словарях и автоматизированных одноязычных словарных системах используются разные методы, среди которых наиболее перспективно создание поискового массива словоформ и словаря основ слов с блоком морфологической обработки. Эти способы могут по-разному сочетаться для различных языков и даже отдельных частей речи в рамках одного и того же словаря.

Ссылаясь на разработку Всероссийского института научной и технической информации, И. И. Убин указывает на одно из наиболее интересных технологических решений задачи нормализации одиночных русских слов и именных словосочетаний. Лингвистические алгоритмы построены на использовании таблиц нормализующих окончаний и словаря, содержащего буквенные коды слов, словообразовательные и словоизменительные основы, номера флективных классов, модели управления. Принцип нормализации заключается в следующем. Существительные и прилагательные, не имеющие чередующихся согласных и гласных в словоизменительной основе, нормализуются с помощью таблицы нор-

мирующих окончаний. Остальные классы слов нормализуются с помощью таблицы подстановок, в которой указываются правила замены текстовых форм на канонические. Именные словосочетания видоизменяются с помощью тех же алгоритмов, которые используются для нормализации одиночных существительных и прилагательных.

Следует отметить, что и метод поискового массива словоформ и метод словаря основ с блоком морфологической обработки имеют свои положительные и отрицательные стороны. При использовании первого в память ЭВМ заносятся все теоретически возможные словоформы лексем, включенных в словарь, с отсылками к словарной форме, которая также включена в поисковый массив. Распознавание запроса осуществляется путем сравнения реализованного в тексте варианта со словоформами поискового массива. Положительной стороной данного подхода является то, что запрашиваемая словоформа не подвергается никакому членению на составные части. Отсутствие алгоритма морфологической обработки уменьшает количество возможных сбоев и делает работу автоматического словаря в целом более надежной. Отрицательной же стороной являются повышенные требования к объему памяти ЭВМ. Так, поисковый массив словоформ английского языка будет примерно в два раза превышать по объему соответствующий массив основ, а для русского языка это соотношение увеличивается в 10—12 раз.

Метод словаря основ предполагает распознавание запрашиваемой словоформы путем членения ее на составные части — основу и словоизменительную флексию. Затем следует отсылка к словарной форме запрошенной словоформы, которая хранится в памяти ЭВМ, либо синтез словарной формы из основы и соответствующей флексии. Использование этого метода предполагает создание алгоритма морфологической обработки со словарем основ и таблицами словоизменительных флексий. Отрицательной стороной данного подхода является более высокая по сравнению с методом поискового массива сложность и трудоемкость лингвистических работ. Вместе с тем этот подход позволяет решать большее количество задач и экономит память ЭВМ.

Подъязык любой предметной области вместе с ее терминологией функционирует как система в среде данной предметной области. Вполне понятно, что среда в значительной степени определяет многие параметры действующей в ней системы. Робототехника отличается строгой иерархией структуры понятий, высокой скоростью реакции исполнительных структур, динамичностью рабочих процессов. Все это накладывает дополнительные условия на язык описания ее понятий и определяет особенности коммуникативных действий. Эти особенности, нередко имеющие экстралингвистические предпосылки, составляют основу комплекса

факторов, влияющих на коммуникативные функции подъязыка робототехники. Основное своеобразие функционально-коммуникативных характеристик подъязыка робототехники заключается в его реакции на лингвистическую и экстралингвистическую среду, в которой он функционирует. Эта среда требует от обслуживающего средства информационного обмена высокой скорости реагирования, а следовательно, экономичности синтаксических конструкций, повышенной определенности семантических границ.

Из-за повышенной полисемии учет иерархии структурных уровней предметной области робототехники особенно актуален. Экстралингвистические факторы, связанные с характером исполнительных действий в данной области, проявляются в повышенной частотности использования в терминосистеме повелительного наклонения в высказываниях и заметного увеличения количества глаголов.

Динамические особенности подъязыка робототехники сказываются в компрессии плана выражения на различных уровнях: в терминологии робототехники существенная доля приходится на сложносокращенные термины; тексты отличаются преобладанием предложений с простым синтаксисом; активное использование форм повелительного наклонения ведет к сжатию синтаксической конструкции предложения. Необходимо также отметить высокую степень формализации текстов по робототехнике. В подъязыке робототехники находят отражение элементы всех типов формализации — структурной, условной, произвольной и семантической.

Коммуникативная функция подъязыка робототехники в значительной степени регламентирована сильной территориальной неравномерностью развития этой области. Высокие темпы ее развития инициализируют функционально-коммуникативные аспекты процессов трансформации, адаптации или нейтрализации лексических и синтаксических элементов, отражающих особенности подъязыка данной предметной области, его связи с языком повседневного общения и его участие в обмене информацией на различных языках.

Проблема обращения к автоматическому переводному словарю в естественной языковой форме является одной из основных при его построении. Без ее решения в максимально полном объеме словарь не может претендовать на успешное решение своих основных задач — помощи человеку при работе с иноязычными текстами и автоматизации лексикографических работ. Конкретные лингвистические и программные приемы построения блока «общения» человека со словарем на естественном языке должны предусматривать достижение оптимальных эксплуатационных характеристик словаря, и в первую очередь по таким параметрам, как скорость получения ответа, объем занимаемой памяти ЭВМ,

надежность работы алгоритма лемматизации, и все это с учетом общей трудоемкости и стоимости указанных приемов в зависимости от выбранного пути.

Анализ отказов автоматического словаря показывает, что многие из них объясняются не только отсутствием в составе словаря запрошенных лексических единиц или орфографическими ошибками, но и «некорректным», т. е. неточным, составлением запросов. «Некорректными» признаются все запросы, по форме отличающиеся от лексических единиц в составе словаря.

«Некорректное» обращение пользователя к автоматизированным информационным системам — явление обычное в условиях промышленной их эксплуатации. Невозможно требовать от пользователя безошибочной работы и полного знания лексики и структуры информации, содержащейся в системе. Незнание особенностей ее функционирования, учитывающего естественный язык, не позволяет пользователю понять и проанализировать причины неудач. Поэтому при неточном запросе подобные системы должны выдавать сообщения об ошибке пользователя без дополнительного запроса с его стороны.

Промышленным автоматическим переводным словарям надо иметь специальные блоки обработки основных типов «некорректных» запросов для того, чтобы повысить эффективность своей работы и одновременно снизить уровень требований к языковой подготовке пользователя.

Однако решение проблемы действенной реакции на многокомпонентные «некорректные» запросы не является конечной точкой на пути повышения гибкости автоматического переводного словаря. Аналогичные проблемы возникают на уровне не только словосочетаний, но и отдельных слов. Автоматический словарь обычно выдает отказ на любое отдельно запрошенное слово, если оно отсутствует в его составе. Значительную часть подобных отказов составляют производные слова.

В традиционной переводной лексикографии производные слова рассматриваются составителями словарей как существенный резерв компрессии словаря, которая может осуществляться двумя способами: а) производные слова полностью исключаются; б) производные слова даются с отсылочными определениями. Первый способ в основном характерен для переводных словарей, второй — для толковых. Однако в обоих случаях необходимо последовательно включать в словарь в качестве самостоятельных вокабул наиболее продуктивные словообразовательные элементы с толкованием их значений и возможными переводными эквивалентами.

В идеале словообразование должно быть освещено и в грамматическом очерке, и в корпусе словаря. На деле же такие сведения, как правило, вообще отсутствуют. Лексикографы не имеют пока

единого мнения о способах представления производных слов в переводных словарях. При рассмотрении с этой точки зрения Большого англо-русского словаря и Русско-английского словаря, становятся очевидными существенные различия между ними. Наиболее последовательно и подробно словообразовательные элементы представлены в Большом англо-русском словаре. В нем отдельными вокабулами даны словообразовательные элементы суффиксального и префиксального типа. В таких блоках отражена семантизация значений словообразовательных элементов и нередко типовые русские переводные эквиваленты. Этот подход значительно расширяет возможности словаря, так как читатель с помощью подобных словарных статей может сам анализировать отсутствующие в словаре производные слова и синтезировать их переводные эквиваленты. Вместе с тем в таком словаре довольно часто встречаются и производные слова, хотя в некоторых случаях в этом даже нет необходимости, ибо они весьма просто и однозначно синтезируются из своих компонентов, которые также включены в словарь.

Менее последовательно проблема производных слов решается в Русско-английском словаре, в котором одни словообразовательные элементы (*не-*, *пере-*) присутствуют в качестве самостоятельных вокабул, а другие, не менее продуктивные (*без-* / *бес-*), отсутствуют.

В традиционной лексикографии существуют специальные словари словообразовательных элементов. В их словарных статьях дается и описание значений этих элементов. Словари такого типа могут быть полезны в учебных целях, а также при составлении традиционных и автоматических переводных словарей.

Серьезную проблему производные слова представляют для автоматических переводных словарей, ориентированных на лексику различных отраслей науки и техники. Деривационный потенциал научного языка в процессе развития самой науки непрерывно возрастает, так как в терминологических системах вследствие интенсивности протекающего отбора словообразовательных средств возникают собственные, характерные для данной области науки терминологические элементы, с участием которых формируются определенные словообразовательные структуры. Более того, в силу интеграции и дифференциации наук и их терминологий число производных терминов, образованных путем перекрестного заимствования стандартных терминологических элементов, постоянно увеличивается. В качестве примеров можно привести такие широко используемые в различных терминологических системах форманты, как *micro-*, *milli-*, *multi-*, *фото-*, *электро-* и др.

Включение в словарь всех производных терминов трудно выполнимо и нецелесообразно. Гораздо удобнее для каждого входящего в состав автоматического словаря языка иметь набор стан-

дартных терминологических элементов с указанием их значений и/или способов передачи на других языках, сопряженных с механизмом их вычленения в составе запрашиваемых лексических единиц.

Наличие в словарях блоков обработки «некорректных» запросов и словообразовательного анализа увеличивает их возможности, и поэтому естественно то большое внимание, которое разработчики автоматических переводных словарей и терминологических банков данных уделяют решению этих проблем. В пользу введения в словари блоков словообразовательного анализа говорит также то, что для полного описания плана содержания языка в равной степени важны лексико-семантический и морфо-семантический уровни. Данные о морфемной семантике приобретают особый вес при создании информационных языков и специальных программ для человеко-машинных систем в рамках ограниченных подязыков.

При анализе сложных слов в автоматическом переводном словаре пользователь получает переводные их эквиваленты, основы которых входят в состав анализируемого сложного слова, и на основании этой информации сам должен синтезировать конкретный переводной эквивалент данного слова, ориентируясь при этом на тот контекст, из которого он его запросил. Очевидно, что отказ от включения в состав переводных словарей производных слов целесообразен только в том случае, если после автоматического их разложения на производящую основу и словообразовательные элементы или сложных слов на составляющие их основы пользователь получит ответ в виде словарных статей этих основ и/или словарных статей словообразовательных элементов с описанием их значений и возможными переводными эквивалентами на выходном языке. Это означает, что в автоматическом словаре помимо словарных статей цельноформированных лексических единиц должны быть словарные статьи словообразовательных аффиксов, терминологических элементов и основ сложных слов.

Выбор правильного значения основы и словообразовательного элемента и соответствующих переводных эквивалентов в автоматическом словаре ложится на самого пользователя. Такое решение соответствует основной идеологии построения и использования автоматических переводных словарей, но не систем МП, в которых анализ входного и синтез выходного текстов должны осуществляться автоматически. В силу сложности построения алгоритмов анализа и синтеза производных слов большинство создателей промышленных систем МП предпочитают вводить такие слова целиком и обрабатывать теми же алгоритмами, что и производные слова.

Алгоритмы обработки «некорректных» запросов и словообразовательного анализа в автоматическом переводном словаре уменьшают количество отказов и повторных обращений пользователя к

словарю, сокращают его общий объем без уменьшения разрешающей способности. Вместе с тем такие алгоритмы в целом замедляют работу автоматического переводного словаря, поэтому, рассматривая вопрос о создании подобных алгоритмов, следует предварительно попытаться определить их конечную целесообразность, т.е. насколько каждый конкретный алгоритм может в целом повысить эффективность данного словаря. При этом следует учитывать степень распространенности и важности тех проблем, для решения которых предназначен такой алгоритм, трудоемкость лингвистических и программных работ, степень его надежности, а также его влияние на функционирование других частей словаря.

Алгоритмы обработки «некорректных» запросов и словообразовательного анализа в автоматическом переводном словаре должны играть вспомогательную роль. Все идиоматичные производные и многокомпонентные лексические единицы, а также лексические единицы с нерегулярным характером обработки следует включать в словарь в качестве самостоятельных вокабул. Вместе с тем, несмотря на вспомогательный характер алгоритмов обработки «некорректных» запросов и словообразовательного анализа, представляется бесспорной необходимость создания подобных алгоритмов, так как они играют определенную роль в превращении автоматических переводных словарей из «складов» лексической информации в «фабрики» по ее обработке.

5.9. Основные лингвистические характеристики функционирования терминосистем

Лингвистические характеристики любого научного или технического текста прежде всего зависят от его функциональной направленности. В пределах подязыка одной предметной области могут быть найдены примеры лингвистических особенностей, свойственных самым различным видам научной и технической литературы. Но нельзя не отметить, что эти особенности существуют, заметно проявляются на всех уровнях языка и значительно различаются в разных предметных областях в силу воздействующих экстралингвистических обстоятельств, о чем наглядно свидетельствует пример работотехники по сравнению с другими предметными областями.

В основном наиболее типичные особенности подязыка можно разделить на две группы:

– те, которые проявляются в изменении частоты употребления языковых норм и их вариантов, традиционно сложившихся в различных жанрах научной и технической литературы;

– и те, которые определяются качественными изменениями языковых норм.

Первая группа в большей степени касается грамматических и синтаксических характеристик подязыка. Вторая относится к лексическому его уровню. Она особенно важна для коммуникации, поскольку наиболее ярко проявляется в новой терминологии, и в частности при переводе новых терминов с одного языка на другой, когда в языке перевода еще не оформился устойчивый эквивалент.

Анализируя тексты по робототехнике на английском, русском и японском языках, можно прийти к выводу, что независимо от языка в научной и технической литературе в основном используются грамматические нормы письменной речи с небольшими изменениями частоты их употребления. Например, в английских текстах отмечается незначительное повышение частоты употребления пассивных конструкций. В японских текстах активизируется повелительное наклонение, довольно редко встречающееся в общелитературной письменной речи. Если сравнивать все три исследуемых языка между собой, то на примерах общих для них текстов при сохранении адекватности переводов интересен факт повышенного содержания в японском языке предположительных и вероятностных форм.

Суммируя результаты экспериментальных наблюдений, можно прийти к выводу, что синтаксис научных и технических текстов в общих чертах не имеет существенных отличий и каких-либо регламентированных особенностей, характеризующих подязыки конкретных предметных областей. Замеченные отклонения не только не обладают смыслоразличительным значением, но и не требуют какого-либо дополнительного анализа. В текстах на русском языке весьма часты случаи смешения падежей при склонении существительных с семантикой действия: *останов на смену инструмента* — «останов для смены инструмента»; *проверка на смещение* — «проверка величины смещения» (при роботизированной сборке); *коррекция позиционирования от люфта* — «коррекция (позиционирования) по величине люфта». В английском языке частым отступлением от нормы можно назвать отклонения в согласовании времен и крайне редкое употребление времен группы Perfect Continuous.

Если перечисленные грамматические и синтаксические неудачи и обращают на себя внимание в силу повышенной частотности употребления, то выражается это больше в плане оценки стилистики текста, чем его грамматической и синтаксической нормы.

Исходя из понимания того, что основной задачей научной и технической литературы является максимально точная и полная передача информационной ценности сообщения, доминирующую роль при подготовке текста играют прагматическая и антропоцентрическая составляющие стиля изложения. Как и в любом другом жанре, в научной и технической литературе структура и форма

изложения должны быть подчинены целям передачи авторской информации. Поэтому основным стилем изложения и в подъязыке робототехники служит формально-логический стиль, основанный на сильно формализованной синтаксической организации текста и подчиненности логической последовательности причин и следствий в иерархии событий предметной области.

Стилистические особенности научного и технического текста обычно оцениваются с двух позиций — с точки зрения полноты представления запланированных для информации фактов и с точки зрения раскрытия соотношений между ними. Стилистические параметры текста при макроанализе могут также определяться планом содержания и планом выражения. Первый в основном поддерживается лексическим уровнем подъязыка предметной области, а план выражения отражается в терминах стиля.

Прагматическая направленность научного и технического текста представляет собой определяющий фактор при установлении необходимого соотношения между планом содержания и планом выражения. Поскольку основная задача текста в данном случае — донесение до читающего фактической информации, то он строится таким образом, чтобы план выражения содействовал наиболее полному раскрытию плана содержания. На практике это реализуется в некоторых ограничениях развертывания плана выражения, проявляющихся в соблюдении следующих основных требований к стилю: композиционной четкости и логической последовательности изложения, экономичности и информативности используемых лексических средств, отсутствии эмоциональной окраски.

Научная и техническая литература как жанр имеет несколько видов, различительной характеристикой которых является степень специализации: от специфической узконаучной до научно-популярной литературы. Количественным показателем степени специализации текста служит его терминологическая насыщенность, т. е. отношение лексики, образующей терминологию конкретной предметной области, использованную в данном тексте, ко всей значимой лексике текста. Терминологическая насыщенность научной статьи может быть в несколько раз выше, чем научно-популярной.

Стиль различных видов научной или технической литературы в значительной степени определяется лексическим и прежде всего терминологическим составом текста. Многокомпонентные термины часто влияют на синтаксическую конструкцию предложения. Для многих видов научной и технической литературы характерно активное использование лексико-грамматических штампов — синтаксических моделей универсального типа.

Привнесение терминологией в научный текст нейтрального книжно-письменного стиля особенно четко прослеживается на

примере научных и технических текстов на японском языке. Известно, что японский язык — один из сложнейших в стилистическом отношении. В нем используется многоступенчатая иерархия степеней вежливости, касающаяся как грамматического, так и лексического уровней языка. Но высокая концентрация в тексте терминов со специфической аффиксацией требует нейтрализации глагольных конструкций в стилистическом отношении.

При общей оценке влияния сложных терминов на стилистические качества текста замечено, что благодаря увеличению частотности вводных слов, союзов и послелогов, которые, казалось бы, затрудняют восприятие предложения, на самом деле более четко прослеживаются устойчивые грамматические конструкции.

Влияние терминологической насыщенности на стиль английского текста менее заметно, чем в японском языке, но и здесь существуют определенные закономерности. В текстах, отобранных для анализа, наблюдается большее, чем в текстах на русском языке, количество образных конструкций. Реже используются сокращенные формы. За счет увеличения числа именных конструкций с семантикой действия упрощаются формы сказуемого.

Высокая терминологическая насыщенность научных и технических текстов, как правило, в большинстве случаев обеспечивает лаконичность и экономность изложения. Термины сами по себе несут точное и экономное определение научных или технических понятий. Системность терминологии упрощает ее функционирование в тексте. Влияние терминосистемы на новую терминологию приводит к сокращению количества терминообразовательных моделей и их унификации. Включение в терминосистему сложных терминов закрепляет логическую связь между отдельными простыми терминами, упрощая восприятие их значений в наиболее полном объеме и обуславливая последующее сокращение их форм, что отвечает принципу рациональной номинации. Такой процесс активно идет во всех трех сравниваемых подъязыках.

Наряду с высокой терминологической насыщенностью, которая в большей или меньшей степени свойственна различным областям науки и техники, лексический состав подъязыка робототехники отличается большим удельным весом наречий, союзов и предлогов, участвующих в формировании вводных конструкций. Экстралингвистический подход к пониманию такой особенности оправдывает ее как стремление приблизить логику изложения к строгой логике робототехнических процессов с программным управлением.

Требование подробного отражения в тексте действий-операций робототехнического устройства, предназначенного заменить аналогичные действия человека, приводит к расширению использования описательных средств и увеличению словарного состава подъязыка. При описании параллельного выполнения нескольких

процессов в текстах по робототехнике чаще, чем в обычном тексте, возникает необходимость включения в высказывание вводных конструкций; в результате увеличивается длина предложения, что особенно заметно в английском языке. Следствием этого является повышение частоты употребления указательных слов *this (these)*, *that (those)*, *it*, *such*, *same*, а также местоимения *one (ones)* в функции определителей при терминах. Углубленное смысловое различительное значение в научных и технических текстах приобретают артикли.

Поскольку робототехника — интернациональная предметная область, во всех трех рассматриваемых текстах используется много заимствованных терминов из других языков.

5.10. Инструментальность термина в составе коммуникативной деятельности

Одна из важных составляющих речевой коммуникации — организация взаимодействия участвующих в ней субъектов, однако сама по себе их речевая деятельность понимается в смысле более широком, чем просто передача информации. Речевая деятельность обеспечивает координацию совместного труда, в интересах которого предпринята коммуникация, и лишь в качестве одного из конституирующих факторов вплетается в совместную трудовую деятельность.

Границы каждого отдельного акта коммуникации определяются прежде всего задачей конкретного взаимодействия людей, поскольку единственной отличительной характеристикой процесса коммуникации в конкретной предметной области от коммуникации в утилитарно-бытовых условиях, в основном обслуживаемой общей лексикой языка, является именно задача взаимодействия в конкретных установках предметной области с целью достижения ожидаемого (или предполагаемого) результата. Что же касается лексического уровня языка, то здесь основные положения приобретают определенную интерпретацию при условии соответствия лексического состава задачам конкретного взаимодействия людей.

Рассматривая состояние и развитие какой-либо определенной сферы деятельности людей, следует признать объективность оснований для широкого включения в речь терминологизированной лексики. Участие терминов в коммуникативном процессе в качестве выразителей главной идеи и номенов предметов или явлений делает их и объектом, и средством достижения цели коммуникации. Высокочастотная включенность терминологии характерна не для всякой совместной деятельности, а прежде всего для такой, которая осуществляется в той или иной специальной сфе-

ре. Но здесь важно понимание того, что использование отдельных терминов в коммуникации возможно в условиях и других подсистем. Объективная необходимость в широком употреблении терминологии главным образом бывает обусловлена специфичностью предметной сферы, в которой разворачивается и координируется совместная деятельность.

Понимание языка как глобальной системы систем создает возможность для представления упорядоченных совокупностей терминов в виде терминосистем, не выходящих за границы лексической системы языка, являющихся ее подсистемами, функционирующими вместе со всеми другими подсистемами в качестве ее элементов и занимающих в структуре лексической системы языка вполне определенное место. Эта точка зрения противостоит взглядам некоторых исследователей, выводящих терминологию из системы языка, наделяющих ее некоторыми искусственными функциями и подменяющих во многих случаях лингвистическую методологию исследования и оценки терминологии исключительно экстралингвистическими методиками.

Наличие таких крайних точек зрения создает поле исследования в области представления терминологии как самостоятельного раздела лексического состава языка и предоставляет для такого анализа возможности рационального использования лингвистических и экстралингвистических методов.

Изучение терминосистемы может быть полноценным только с позиций ее функционирования и развития, а основу развития терминосистемы молодой предметной области составляет адаптация новой терминологии.

5.11. Терминологические проблемы перевода

Исходя из принятого в лингвистике деятельностного представления о языке, большинство процессов получения информации и ее обмена можно описать в терминах аналитико-синтетической обработки, под которой принято понимать классификацию и систематизацию, связанные с кодированием, библиографическим описанием и аннотированием информации, обобщением, реферированием и переводом с одного языка на другой. Во всех этих случаях результатом аналитико-синтетической обработки информации является вторичная информация.

С позиций понимания аналитико-синтетической сущности коммуникативной деятельности, порождением и результатом которой становится первичная информация, текст представляет собой «непосредственное коммуникативное средство первого порядка»¹.

¹ Сидоров Е. В. Основы системной концепции текста: автореф. дис. ... д-ра филол. наук. — М., 1998. — С. 27.

Первичная коммуникативная деятельность состоит из целостной совокупности этапов, главные из которых: ориентирование в условиях достижения цели, сличение полученного с намеченной целью и отображение его в коммуникативной деятельности, например, в виде текста. В самом деле, текст — это продукт первичной коммуникативной деятельности, ее предметно-знаковая реализация.

Одним из наиболее типичных образцов такой переработки информации является перевод с одного языка на другой. Перевод научных и технических текстов представляет собой единый аналитико-синтетический процесс, при котором производится прежде всего анализ исходного научно-технического текста, т. е. выявление главной темы и способов ее выражения на языке оригинала, и синтез полученной информации в виде текста, т. е. передача (трансляция) темы высказывания средствами языка перевода. Промежуточный этап между анализом и синтезом текстов представляет собой собственно перевод, т. е. этап преобразования одних языковых форм в другие.

5.11.1. Лингвистические аспекты

В переводе с одного языка на другой, как мы уже отмечали, можно выделить две основные группы проблем — лингвистические и экстралингвистические. Лингвистические проблемы в какой-то мере будут общими для всех видов перевода, существующих в пределах одной пары языков. Экстралингвистические проблемы, как правило, представляют собой макропроблемы, которые чаще всего поддаются делению на ряд более простых задач, многие из которых решаются теми же методами, что и лингвистические. В состав последних входит и занимает среди них особое место большая группа лексических проблем.

Здесь необходимо уточнить, что в специальных сферах человеческой деятельности характеристикой лексического состава текстов является его высокая насыщенность терминологией. Во многих случаях лексические проблемы невозможно решать только лингвистическими методами, для этого требуются еще и экстралингвистические знания.

Практика научно-технического перевода доказывает, что лексические проблемы имеют комплексный характер, включающий лингвистический и экстралингвистический компоненты. А потому они довольно часто требуют решения экстралингвистическими методами. Точнее, занимаемое лексическими проблемами место можно в зависимости от рассматриваемой ситуации признавать и центральным, и пограничным между лингвистическими и экстралингвистическими проблемами:

1) разграничивая оба эти круга проблем, они создают центробежные явления в формировании методологических подходов к научному и техническому переводу;

2) объединяя лингвистические и экстралингвистические методы решения своих задач, лексические проблемы порождают центристренительные явления.

Такое положение и такое значение лексических свойств служат важнейшими доказательствами системности и структурности всего комплекса проблем перевода — лексические проблемы восходят в ранг регламентирующих основ научного и технического перевода.

Как мы уже упоминали, Р. Г. Пиотровский считает, что в формализованных деловых и научно-технических документах до 80 % синтаксической и семантико-прагматической информации содержится в лексике и фразеологии текста, и правильный перевод 70 % словоупотреблений иностранного текста приводит к удовлетворительному пониманию содержания текста¹.

С практической точки зрения наибольший интерес представляет поиск наиболее универсальных путей решения лексических проблем перевода. Насколько значимы эти проблемы, говорит статистическое исследование, выполненное на материале довольно большого массива статей по нескольким разделам маркетинга. Все лексические единицы текстов были разделены на две группы: терминологическая и нетерминологическая лексика. Анализ, проведенный методом экспертной оценки, показал, что доля терминологической лексики в среднем составила 54 %, причем это соотношение не зависело от языка оригинала (рассматривались статьи на английском, японском и русском языках). В каждом случае эксперту предлагались тема статьи и список всех использованных в ней словоупотреблений. В результате эксперимента мы получили список терминологической лексики. (По такой технологии работают терминологи Департамента переводов Европейского экономического сообщества.)

Можно утверждать, что основные трудности переводчика научной и технической литературы связаны с терминологической лексикой. Они еще усугубляются тем, что в информационных потоках происходят не только количественные, но и качественные изменения, среди которых для научно-технического перевода наиболее важна тематическая интерференция.

Трудно найти такую область, где в настоящее время не использовалась бы вычислительная техника. Автоматизация от простейших манипуляторов до гибких производственных систем, роботизированных комплексов и заводов-автоматов широко прони-

¹ См.: Пиотровский Р. Г. Предисловие // Л. Л. Нелюбин. Перевод и прикладная лингвистика. — М., 1983. — С. 4.

кает в науку, технику, сферу обслуживания. Все предметные области озадачены проблемами охраны окружающей среды, энергосбережения и снижения материалоемкости. И это требует значительного расширения кругозора переводчика. Значение терминологической работы повышается также из-за сильной специализации общенациональных языков, основная причина которой в бурном развитии научно-технического прогресса.

Терминологические системы отдельных предметных областей функционируют настолько активно, что оказывают значительное влияние на язык повседневного общения. Нередки случаи проникновения специфического синтаксиса длинных терминологических словосочетаний в общенациональный язык. Все это обуславливает большую зависимость переводчика от терминологического обеспечения. Поэтому среди резервов повышения качества и увеличения объемов выполняемых переводов наиболее существенным остается терминологическая работа. Переводческий корпус все требовательнее относится к своим основным инструментам — словарям.

Наиболее динамичное развитие практической лексикографии началось сравнительно недавно и в нем имеется еще немало не до конца использованных путей. Новая терминология — это особый пласт лексического уровня подязыка предметной области. Уровень национального развития предметной области и его соотношение с мировым опытом определяют интенсивность притока в нее новых терминов. Знание появляющейся терминологии в значительной степени повышает когнитивную значимость перевода в аналитико-синтетическом процессе распространения информации.

5.11.2. Прагматические аспекты

Прагматическая составляющая языкового общения, насыщенного новой терминологией, требует лингвистического обеспечения, предусматривающего создание специалисту комфортности в сфере общения.

Неравномерность национально-территориального развития предметных областей углубляется различиями социальных условий. На примере предметной области маркетинга можно убедиться в значительной разности социальной мотивации терминологии. Современная система обмена информацией, обладающая высокой технической оснащенностью, имеет явные тенденции к антропоцентризму и прагматике.

Какому бы анализу не подвергался аналитико-синтетический процесс, и в частности процесс перевода, и с каких бы позиций он не рассматривался, все его уровни и этапы в той или иной мере связаны с обработкой лексических средств языка оригинала

или языка перевода. И основную часть такой обработки составляет осмысление терминов. Причем процесс этот постоянно углубляется и функционально зависит от научно-технического прогресса.

Терминообразование значительно опережает словообразование и идет как путем создания терминов-словосочетаний, так и путем присвоения новых и расширяющих значений уже известным словам. Углубляется многозначность слов, что в значительной степени определяется специализацией отдельных узких направлений областей знаний, т.е. дифференциацией предметных областей, обусловленной развитием новых технологий и новых материалов.

Если всего лишь 50—60 лет назад, например, керамическое производство могло обходиться двумя-тремя сотнями терминов, то сейчас изданы солидные словари по таким узкоспециализированным направлениям, как бытовая керамика, медицинская керамика, промышленная керамика. Что же касается последнего из них, то уже трудно ограничиваться таким понятием, как «терминология предметной области промышленной керамики». В этой области уже выделились такие направления, как электрокерамика, металлокерамика, оптокерамика и др. В электрокерамике, в свою очередь, сформировались терминологические подсистемы сикнетокерамики, пьезокерамики и т.д. Разработка пьезоэлектрических двигателей готовит революцию в робототехнике. С позиций языковых средств все процессы такого бурного развития керамики практически происходили в среде стабильной лексики, которая адаптировалась с появлением новых специализаций. Преимущественным способом адаптации оставалось углубление многозначности слов, выражающих определенный круг понятий, характеризующих материалы и технологию керамики.

Терминологическое обеспечение должно быть системным и основанным на изучении внутренних процессов, происходящих в терминосистеме подязыка предметной области, и динамических процессов, определяющих закономерности формирования терминологии. В этом смысле большое значение приобретают исследования образования и функционирования новых терминов.

Таким образом, проблемы терминологического обеспечения составляют две основные группы:

- внутренние проблемы лингвистического происхождения, решаемые в основном методами лингвистических исследований;
- внешние проблемы преимущественно экстралингвистического происхождения, решаемые в основном методами нелингвистических исследований.

Характер исследований находится в прямой зависимости от особенностей предметной области. Применительно к научному и техническому переводу эти проблемы чаще всего могут быть решены различными методами инженерных исследований.

5.12. Автоматические переводные словари

5.12.1. Лексический состав и объем

Отбор лексики для любого словаря обусловлен прежде всего тем, для решения каких задач данный словарь предназначен и кто будет его основным пользователем. Составитель словаря сам определяет, какие пласты лексики следует в него включать. Основной объем в нетерминологических переводных словарях отведен нормативной лексике общего языка, но немалую часть занимает и специальная лексика. В них включаются также, но в значительно меньших масштабах, устаревшая и диалектная лексика, регионализмы, экзотизмы и собственные имена. Весьма остро стоит вопрос о введении ненормативной лексики.

Промышленные системы МП и автоматические переводные словари предназначены в первую очередь для повышения оперативности перевода научно-технической литературы и документации. Это объясняется особой ролью научно-технического перевода в информационной деятельности развитых стран. Причем его объемы и значение постоянно возрастают. Такая ориентация промышленных систем МП и автоматических словарей предопределяет тот факт, что основную часть их состава занимает терминологическая лексика.

Теоретические различия между терминологией и номенклатурой не оказывают никакого влияния на практику работы с иноязычными текстами, поскольку главные пользователи переводных терминологических словарей — переводчики, редакторы и технические специалисты — не делают различий между термином и номеном. Именно поэтому большинство терминологических словарей включают и термины, и номены. Для автоматических переводных словарей и систем МП не существует дилеммы термин — номен, так как отсутствие в словарном запасе системы МП или переводного словаря той или иной лексической единицы, относящейся к терминологии или к номенклатуре, одинаково отрицательно сказывается на результатах обработки входного текста. Термины и номены с одинаковым правом должны быть включены в состав словарей этих двух типов.

Как правило, крупные промышленные автоматические словари содержат лексику нескольких тематических областей, поэтому проблема тематического пересечения для них стоит столь же остро, как и для традиционных терминологических словарей. Однако если для традиционной лексикографии эта проблема фактически равнозначна дилемме — включать или не включать слово, то в автоматическом переводном словаре, в силу его политематичности и большого объема, задача чаще сводится к решению, к какой тематической области отнести данную лексическую единицу.

Создатели промышленных автоматических переводных словарей уделяют решению этой задачи самое серьезное внимание и разрабатывают специальные тематические классификации лексики. Тематическое кодирование в таком словаре предназначено в первую очередь для решения проблемы омонимии и многозначности терминов и установления границ запроса. Поэтому чем глубже используемая в словаре тематическая классификация, тем успешнее могут быть решены эти задачи. Однако углубление классификации сопряжено с увеличением возможностей ошибок лексикографа при составлении словарных статей и затруднений пользователя при формировании запроса. Ошибки тематической атрибуции лексической единицы, в свою очередь, либо приводят к сбоям в работе автоматического словаря, либо вынуждают пользователя расширять поиск необходимой ему лексики на несколько тематических областей, что в конечном счете снижает качество и скорость работы словаря и повышает эксплуатационные расходы.

Трудности создателей автоматических переводных словарей усугубляются еще и тем, что пока не разработаны общепризнанные теоретические принципы терминологической классификации человеческой деятельности, и поэтому каждый словарь и каждый терминологический банк руководствуются своей собственной, часто несовместимой с другими, классификацией, что осложняет координацию усилий в этом направлении и обмен лексическими массивами.

Комплектование лексических массивов автоматических переводных словарей, как и любой другой автоматизированной системы обработки текста, обычно начинается с выявления лексического состава подязыков, на обработку которых нацелена создаваемая система. При этом объектом лексико-статистического исследования выступает не сам подязык или комплекс подязыков, а некое множество текстов, определяемое теми или иными прикладными задачами. Такое множество в принципе может быть бесконечным или, во всяком случае, открытым. На основании исследования определяется некоторый лексический массив, который и вводится в первую очередь в состав словаря. Такой подход позволяет добиться высокой степени покрытия текстов, для обработки которых предназначена данная система.

Вместе с тем формирование лексических массивов автоматических переводных словарей только на основании предварительно обработанных текстов сужает круг задач, которые он может решать, и понижает его рентабельность. Разработчики автоматических переводных словарей, как правило, стремятся расширить лексический состав главным образом за счет близкой по тематике лексики, имеющейся на магнитных носителях в составе либо других автоматических переводных словарей, либо систем МП, либо автоматизированных лексикографических систем, что позволяет

избежать трудоемкой и дорогостоящей операции записи лексической информации на магнитный носитель. Для комплектования автоматического переводного словаря широко используются также традиционные терминологические словари и лексика, поступающая в материалах организаций или от отдельных пользователей словаря. В результате терминологическая лексика составляет от 60 до 80 % объема такого словаря.

Крупный автоматический переводной словарь промышленного типа должен быть «тексто-независимым», т.е. ориентированным на обработку текстов самых различных видов, поэтому в него следует включать всю терминологическую лексику по основной тематике словаря, наиболее употребительную лексику из соседних тематических областей, общенаучную и общетехническую лексику, а также наиболее общеупотребительную. Необходимость включения общей лексики объясняется тем, что при работе с техническими текстами трудности могут скрываться в незнании не только специальной лексики, но и некоторых слов общего языка.

Создатели автоматических переводных словарей не могут пойти по пути составителей традиционных терминологических словарей, которые обычно не включают в их состав общую лексику, справедливо полагая, что читатель найдет ее в обычных переводных словарях. Отсутствие в переводном словаре значительного числа слов из обрабатываемого текста заставляет пользователя слишком часто прерывать свою работу и обращаться за помощью к традиционному словарю. Не отрицая самой возможности обращения по конкретным вопросам к традиционным лексикографическим источникам, следует подчеркнуть, что чем чаще пользователь вынужден это делать, тем меньшими будут для него преимущества автоматического словаря перед традиционными и тем менее выгодным покажется ему использование автоматического переводного словаря в процессе обработки иноязычных текстов.

Дополнительным доводом в пользу включения общей лексики в автоматический переводной словарь может служить тот факт, что такие словари иногда используются для выдачи переводных эквивалентов всех слов текста в форме традиционных алфавитных или текстоориентированных переводных словарей.

Большое внимание проблеме нормирования лексики уделяется и в терминологической лексикографии. Стремительное развитие науки и техники ведет к значительным изменениям в существующих терминологических системах, к возникновению новых их вариантов, создаваемых на основе собственной возникающей терминологии, а также за счет заимствования терминов из смежных наук. Бурный неконтролируемый процесс терминоворчества, безусловно, затрудняет общение между работниками разных специальностей и даже между специалистами одного профиля, но принадлежащими к разным школам. Все это настоя-

тельно требует практического решения проблемы нормирования в терминологии.

Следует отметить, что в этом направлении терминологическая лексикография достигла большего, чем общая лексикография. Одним из разработанных ею средств решения этой проблемы являются терминологические стандарты. В настоящее время создано значительное количество стандартов и рекомендаций по терминологии на международном, региональном и национальном уровнях. В некоторых странах стандартизованные термины и выражения становятся обязательными для употребления во всех документах, исходящих от правительственных организаций, в контрактах, в учебной литературе и в научных публикациях. Однако многие термины, рекомендованные к употреблению, в практике научно-технического общения используются весьма редко или же вовсе не используются. Иногда специалисты не только не принимают стандартизованные термины, но и активно выступают с их критикой. Это особенно остро ощущается в новых и развивающихся отраслях науки и техники. Включение в терминологические словари только стандартизованной лексики приводит к тому, что специалисты и переводчики не получают необходимой информации о множестве разговорных и сленговых синонимов рекомендованных терминов.

В некоторых языках, например в английском, разговорная лексика вообще и нереконмендованные термины в частности довольно часто встречаются в научно-технических текстах, так что их отсутствие в терминологическом словаре становится реальной помехой для понимания текстов.

Автоматические переводные словари предназначены для обработки всех видов научно-технических текстов, поэтому они должны включать все термины, как стандартизованные и рекомендуемые, так и нереконмендуемые, как устоявшиеся, так и новые, которые еще не успели получить единого понимания и не вошли в общепринятую лексикографическую обработку.

Одно из основных преимуществ автоматического словаря перед традиционным состоит не в том, что он автоматический и потенциально превосходит традиционный словарь по скорости поиска нужной лексики, а именно в том, что он включает всю используемую терминологию, постоянно и часто пополняет свой состав новыми терминами и, как следствие, качественно превосходит по составу традиционный словарь. Вместе с тем при таком свободном подходе к отбору лексики, для того чтобы словарь пользовался у читателя доверием, целесообразно сопровождать каждую лексическую единицу информацией, уточняющей сферу ее употребления, указанием на то, является ли данный термин общепринятым, зафиксирован ли он в стандартах, в авторитетных лексикографических изданиях. В общей лексикографии эту

задачу в определенной степени решают стилистические пометы. В терминологической лексикографии стилистические пометы — пока явление весьма редкое.

В идеологии автоматических переводных словарей подобная информация получила название индекса надежности, кода качества, весового коэффициента. Учитывая, что разработчикам таких словарей приходится обрабатывать большие лексические массивы в весьма жесткие сроки, индекс надежности может быть достаточно удобным практическим решением проблемы нормы для терминологической лексики.

В некоторых автоматических переводных словарях и терминологических банках в основе определения индекса надежности лежит тип источника поступления всей словарной статьи или ее фрагмента. Высший индекс надежности поручает информация, взятая из национальных или международных стандартов либо из рекомендаций авторитетных терминологических служб. Далее следует индекс, указывающий, что термин уже зафиксирован в авторитетных переводных словарях. Низший индекс надежности получает термин и его переводной эквивалент, которые взяты из научно-технических текстов и их переводов и еще не прошли проверку практикой. Один и тот же термин может иметь несколько переводных эквивалентов с разными индексами надежности, которые могут изменяться в ходе эксплуатации переводного словаря.

Существенное значение приобретает выбор основной единицы словника. Для крупных переводных словарей общей лексики такой единицей является слово; словосочетания обычно даются в составе словарных статей отдельных слов в качестве иллюстрации значений этих слов либо как пример сочетаемости заглавной единицы. В лексикографической практике основным критерием включения слова в словарь служит его употребительность. Одновременно с этим составитель переводного словаря должен учитывать частотность не только слов входного языка, но и их переводных эквивалентов в выходном языке.

В словарях систем МП в качестве основной единицы также чаще всего выступает слово. Словари фразеологических оборотов и различных словосочетаний в них организованы с привязкой к отдельным словам. По-иному обстоит дело в традиционных переводных терминологических словарях, которые в большей мере содержат словосочетания. Это подтверждают специальные исследования на материале переводных терминологических словарей и одноязычных тезаурусов научно-технических терминов.

В научно-технических словарях доля словосочетаний составляет 75—80 % всего словарного состава. Такое положение в терминологической лексикографии вызывает критику многих исследователей.

Составители терминологических словарей по-разному толкуют понятие «термин», по-разному определяют и его границы, а также «нижний» и «верхний» пределы этого понятия. В результате словники терминологических словарей часто содержат «балластный материал», не имеющий практической ценности.

Не оправдывая в целом такой способ комплектования словника терминологического словаря, следует отметить, что в основе этого довольно распространенного явления лежат субъективные и объективные причины. К субъективным относится простота лексикографических решений данного подхода, при котором составитель словаря не должен строго решать вопрос о границах многокомпонентных терминов и о переводимости их составляющих. К объективным причинам можно отнести, во-первых, то, что в словосочетаниях заложено заметно больше структурно-стилистических связей текста, чем в отдельных словах, что весьма существенно при работе со связным текстом. Во-вторых, включая в состав словаря большое количество словосочетаний, разработчик стремится точнее отразить запросы своего читателя: приблизительно 80 % обращений к терминологическому словарю составляют словосочетания.

С точки зрения основной единицы словника автоматические переводные словари близки к традиционным терминологическим: 80—90 % их лексического массива представлено словосочетаниями. Происходит это потому, что крупные автоматические словари в большинстве своем фактически являются терминологическими банками и их основная функция направлена на решение тех же задач, что и у традиционных терминологических словарей. Интересная особенность автоматических словарей состоит в том, что верхняя граница словосочетаний в них еще более размыта, чем даже у традиционных. В крупные автоматические словари нередко включаются в качестве самостоятельных статей отдельные предложения и даже фрагменты текстов. Подобный подход к комплектованию словника в традиционной лексикографии не используется. Однако здесь такое решение в определенных условиях оказывается весьма удобным. Дело в том, что переводные словари показали себя эффективным инструментом полуавтоматического перевода стандартизованных текстов, например товаросопроводительной документации или юридических документов, в которых отдельные части регулярно повторяются без существенных изменений.

Количество свободных словосочетаний и единиц текста в автоматическом переводном словаре в первую очередь определяется его типом и задачами. Чем более ограничены цели словаря, чем он уже тематически, тем в большей степени такой словарь оказывается текстоориентированным, тем больше в нем может и должно быть свободных словосочетаний и единиц текста. Чем шире

диапазон задач автоматического переводного словаря, чем шире спектр тематик и различных слов, которые он охватывает, тем более он должен быть системоориентированным, тем меньше в нем должно содержаться свободных словосочетаний и единиц текста.

Одним из самых продуктивных способов словообразования в современную эпоху является аббревиация. Она характерна для всех языков. Компрессии подвергаются словосочетания и даже отдельные слова. Образование усечений, т.е. сокращений слов, давно практикуется в письменных формах общеупотребительной лексики. В настоящее время одно- и двусложные усечения получили распространение и в общенаучной и терминологической лексике. Наиболее распространенный способ образования таких сокращений — отсечение конца слова.

В терминологической лексике четко прослеживается тенденция к свертыванию многокомпонентных терминологических единств и замена их в сфере функционирования различными сокращениями: инициальными аббревиатурами, акронимами, сращениями и другими вариантами. Образование всех видов сокращений в терминологической лексике — процесс в основном стихийный, никем не управляемый. Появление среди них все новых приводит к возникновению синонимии и омонимии как между терминосистемами разных наук, так и в пределах одной и той же терминосистемы, что безусловно осложняет межъязыковое и внутриязыковое общение ученых и специалистов разных направлений. Рекомендации авторитетных международных и национальных органов, к сожалению, пока не оказывают решающего влияния на процесс порождения терминологических сокращений.

В таких условиях удобным практическим решением является включение всех сокращений, обнаруженных в сфере определенного функционирования, в специальные словари с указанием степени употребительности. Такие словари регулярно издаются для основных языков мира и для главных направлений развития науки и техники.

Помимо специальных словарей сокращения вводятся практически во все переводные терминологические словари либо как самостоятельные единицы, либо, чаще всего, в составе специальных приложений, в которых обычно приводятся несокращенная форма терминологической единицы, ее полный переводной эквивалент и соответствующее сокращение на выходном языке. Таким образом, можно сказать, что проблеме сокращений в практике традиционной лексикографии уделяется серьезное внимание.

Автоматические переводные словари также не устраняются от решения этой проблемы. Как правило, промышленные автомати-

ческие словари включают аббревиатуру с отсылкой к основной несокращенной форме данной лексической единицы. Индекс надежности, который служит в таких словарях критерием степени рекомендованности термина или его переводного эквивалента, может присваиваться также и сокращениям. Динамичность автоматических переводных словарей позволяет изменять индекс надежности любой единицы, в том числе и сокращения, либо совсем исключить ее из своего состава в зависимости от частотности или рекомендаций специальных терминологических органов. Включение всех действующих в определенной тематической сфере сокращений и последующее прослеживание их употребительности и статуса безусловно осложняет задачи составителей автоматических переводных словарей и увеличивает их объем. Однако этим значительно облегчается задача пользователя и возрастает ценность такого словаря как источника лексикографической информации.

Сопоставляя перечень характеристик, на основании которых строят свои типологии разные исследователи, можно сделать вывод, что наибольшее значение все отводят принципам отбора лексики, организации и интерпретации лексического материала. Объем является как бы вторичной характеристикой, производной от времени разработки словаря и возможностей составителя (или их коллектива), и существенно не влияет на тип словаря. Это косвенно подтверждается большим разнообразием объемов словарей одного и того же типа. Например, регулярно издаются различные переводные словари общей лексики для одной и той же пары языков объемом от 2 тыс. до 100 тыс. слов — карманные, краткие, средние, большие. Примерно такое же положение и в переводной терминологической лексикографии. Однако в этом разнообразии объемов существует одна четкая закономерность: с увеличением объема количество словарей резко уменьшается. Так, если количество переводных словарей объемом до 20 тыс. слов для развитых лексикографий исчисляется сотнями, то словарей с объемом от 20 до 50 тыс. слов десятки, а от 50 до 100 тыс. — уже единицы. Переводные словари объемом 150 тыс. слов и более явление исключительное практически для всех развитых лексикографий. Это объясняется не только огромными трудностями, с которыми сталкиваются создатели таких словарей, но и объективными недостатками традиционных «бумажных» словарей. Словари подобных объемов в книжной форме содержат тысячи страниц, в большинстве своем являются многотомными и, как следствие, неудобными в работе и дорогостоящими.

Вместе с тем развитие общества в целом, науки и техники в первую очередь, требует формирования лексикографических фондов для языка во всех его проявлениях, создания новых типов словарей, органично включающих в себя толкования, перевод-

ные эквиваленты, контексты, лексическую сочетаемость и многие другие виды лингвистической и экстралингвистической информации.

Основываясь на сопоставлении объемов традиционных переводных словарей, с одной стороны, и многоязычных автоматических словарей и терминологических банков — с другой, а также учитывая современные тенденции развития вычислительной лексикографии, можно сделать вывод, что объем является существенной характеристикой, принципиально различающей традиционные и автоматические переводные словари. Автоматические словари объемом в несколько миллионов словарных статей стали в настоящий момент реальностью, в то время как максимальные объемы традиционных переводных словарей достигают 150—200 тыс. словарных статей. Практика показывает, что создание словарей объемом 300—500 тыс. слов в книжной форме нерационально. Первичной для словарей должна стать компьютерная форма, а книжная может быть производной, проекцией автоматизированных лексикографических баз данных по составляющим, структуре и номенклатуре словарных статей.

5.12.2. Динамичность автоматических переводных словарей

Одним из существенных свойств словарей в традиционной книжной форме является их статичность, что обнаруживается на уровне состава словника и на уровне содержания словарных статей. Сложность лексикографических работ, редакционно-издательские, типографские, а также экономические факторы не позволяют издавать большие словари с необходимой периодичностью.

Сопоставление данных о периодичности переиздания переводных и одноязычных словарей убедительно свидетельствует о том, что традиционная лексикография в среднем на 7—10 лет отстает от процесса фактического обогащения или изменения терминологических систем современных активно развивающихся направлений науки и техники. В качестве возможного пути частичного решения проблемы пополнения словарей используется выпуск различных дополнений к ним, которые выходят либо самостоятельно, либо в составе того же словаря при его стереотипном переиздании. Запаздывание с включением новой лексики отмечается переводчиками как один из основных недостатков традиционных словарей.

Словарная статья «бумажного» издания статична и по своей структуре, и по содержанию. Это означает, что любые неточности или ошибки могут быть исправлены, а дополнительная информация, даже самая актуальная, может быть внесена только при очередном переиздании.

Разные задачи и различные уровни языковой подготовки читателя словаря предъявляют заметно различающиеся требования к составу и объему информации в словарных статьях, что, в свою очередь, вызывает необходимость создания словарей разных типов. Так, в переводных словарях интенсивного типа словарная статья содержит помимо переводных эквивалентов разнообразные сведения семантического и стилистического характера о заглавном слове и его эквивалентах в выходном языке, примеры употребления, типовые и идиоматические словосочетания, синонимы, антонимы и т. п. Как правило, значительная часть этой информации для читателя с хорошим знанием языка оригинала, каковым, например, является профессиональный переводчик, избыточна. В то же время в переводных технических словарях, большинство которых представляет собой словари экстенсивного типа, словарные статьи значительно беднее и чаще всего не содержат практически никакой информации, кроме переводных эквивалентов, что для читателя со слабым или средним уровнем знания языка оригинала явно недостаточно. В условиях традиционной лексикографии противоречие между уровнем языковой подготовки читателя, ожидаемым разнообразием задач и целей словаря, с одной стороны, и объемом и характером информации в его словарных статьях — с другой, в рамках одного «бумажного» словаря не разрешимо из-за его статичности.

В отличие от традиционных словарей словарные компоненты систем МП и автоматические переводные словари обладают высокой динамичностью. Отбор лексики продолжается в течение всего времени функционирования словаря; постоянно происходит исключение оказавшихся ненужными слов и включение новой лексики, коррекция и пополнение информации в любой зоне словарной статьи. Переменность состава и способность воспринимать изменения в ходе функционирования являются принципиальными характеристиками, отличающими компьютерные словари от обычных. Динамичность состава осуществляется в основном за счет гибкости программного обеспечения. На лингвистическом уровне достижение динамичности зависит от четкой структуры словарной статьи и возможности записи однотипной информации строго в одной и той же зоне всех словарных статей, от единообразия интерпретации больших массивов лингвистических объектов на основе всей суммы накопленных знаний. На организационном уровне резко возрастает роль слежения за динамикой изменения состава словаря, а также состава и объема информации внутри словарных статей.

Следовательно, автоматические переводные словари обладают принципиальной возможностью преодоления противоречий между словарями интенсивного и экстенсивного типов, предоставляя пользователю право самому определять состав и объем необходимой ему информации.

5.12.3. Состав и структура словарных статей

Словарная статья — основная коммуникативная и композиционная единица словаря как особого вида справочной литературы. Ее состав и структура определяются многими факторами, главные из которых целевая направленность словаря, доминирующие лексикографические традиции и реальные технические возможности.

По мнению Ю. Н. Караулова, существуют две разнонаправленные линии развития словарного дела. Центростремительная линия предусматривает разработку универсального словаря, охватывающего в идеале все сведения о данном языке — грамматические, синтаксические и экстралингвистические. На практике создание такого идеального словаря оказывается недостижимым, но почти для каждого языка с развитой лексикографией имеются словари, обладающие универсальными свойствами. Центробежная линия, напротив, нацелена на создание однопараметрических словарей, описывающих отдельные явления данного языка. Это словари орфографические, словообразовательные, словари синонимов, сочетаемости и т. д. Их словарные статьи содержат информацию только одного типа, поэтому при практической работе обычно требуется одновременно несколько разных однопараметрических словарей¹.

Таким образом, первоочередной проблемой при разработке словаря, обладающего достаточно богатой информацией, но в то же время не столь сложного, как универсальный, является определение минимально достаточного состава словарной статьи. Она должна иметь, по меньшей мере, следующие компоненты:

- 1) заглавное слово;
- 2) его формальные характеристики (грамматические, орфоэпические, орфографические и т. д.);
- 3) его семантизацию;
- 4) извлечения из текстов, иллюстрирующие ту или иную формальную или семантическую его особенность;
- 5) указания на «соседей» заглавного слова в лексической системе языка по разным осям ее семантического пространства;
- 6) отсылки и справки разного характера и назначения.

Такая модель предлагается для словарной статьи одноязычного филологического словаря. Для переводного словаря к этому перечню следует добавить переводные эквиваленты, т. е. слова или словосочетания выходного языка, которые составитель приписывает заглавному слову и которые в идеале должны передавать максимум информации, содержащейся в заглавном слове.

¹ Подробнее см.: Караулов Ю. Н. Лингвистическое конструирование и тезаурус литературного языка. — М., 1981.

В терминологической лексикографии идеальную словарную статью можно представить как объединение требований собственно энциклопедического и филологического описаний. Согласно требованиям энциклопедического описания, она должна включать:

- 1) заглавное слово (термин);
- 2) научный раздел;
- 3) этимологию;
- 4) первое определение (с цитатой на языке оригинала);
- 5) современное определение;
- 6) дополнительные разъяснения и замечания относительно различных определений или трактовок значения термина;
- 7) синонимы;
- 8) аббревиатуры;
- 9) литературу;
- 10) автора словарной статьи и год ее написания.

Подобный состав, как и идеальная словарная статья филологического словаря, ориентирован на одноязычные словари. Требования переводной лексикографии требуют расширения этого формата. В этом случае статья, также идеальная, должна включать переводной эквивалент, фонетическую и грамматическую характеристики, пометы сфер употребления как в смысле стилистическом (*разг.*, *офиц.*, *проф.* и т. п.), так и в смысле концептуально-системном.

В практике создания общефилологических и терминологических словарей эти идеальные варианты существенно упрощаются. Как уже было сказано, в настоящее время нет универсальных общефилологических словарей, хотя статьи некоторых толковых словарей обладают весьма богатой информацией. Наиболее разработанные переводные словари, например «Большой англо-русский словарь», включает бóльшую часть компонентов описанного идеального варианта. Однако в количественном отношении все-таки пока преобладают словари с ограниченным составом словарных статей. Такое упрощение особенно наглядно прослеживается в терминологической лексикографии. Большинство переводных терминологических словарей не включают и половины указанных в идеальном случае характеристик и чаще всего ограничиваются краткой грамматической информацией (индекс грамматического класса), переводным эквивалентом (или эквивалентами) и определением.

Ю. Н. Караулов выделяет 66 параметров для лексикографического описания языка¹, однако даже разрабатываемые в последнее время проекты универсальных словарей охватывают менее половины из них. Упрощения словарных статей для практики

¹ См.: Караулов Ю. Н. Лингвистическое конструирование и тезаурус литературного языка. — М., 1981. — С. 75—76.

создания словарей в книжной форме стало естественными. Причины этого кроются и в трудностях совмещения в пределах одной словарной статьи всех сведений о слове, и в огромной трудоемкости подбора такого материала. Словарная статья, охватывающая все языковые и неязыковые сведения о слове, может сделать словарь столь сложным, что для массового читателя он будет практически бесполезным.

Подобные же явления наблюдаются и в словарных компонентах различных систем автоматической обработки текстов. Создатели некоторых автоматизированных систем, предусматривающих общение с компьютером на естественном языке, систем МП и автоматических словарей разработали для них довольно сложные и глубокие по содержанию словарные статьи. К числу постоянных параметров, характерных для всех частей речи, отнесены следующие лингвистические компоненты:

- 1) орфографический, т. е. сама словарная единица;
- 2) часть речи;
- 3) семантическое поле;
- 4) словарная дефиниция;
- 5) нерегулируемая словоизменительная форма;
- 6) словообразовательная морфология;
- 7) синонимы, включая синонимические отсылки;
- 8) антонимы;
- 9) примеры употребления для каждого случая, имеющего дефиницию.

Некоторые разработчики систем, рассчитанных на общение человека с компьютером на естественном языке, считают необходимым давать в словаре в декларативной форме индивидуальные сведения об особенностях конкретных лексем. Объем лингвистической и прагматической информации в словарных статьях таких систем весьма велик, тем не менее его недостаточно для обработки любых текстов русского языка. Однако дальнейшее их расширение разработчики считают нецелесообразным по следующим причинам:

– трудности воссоздания полной лингвистической модели русского языка;

– проблемы реализации такой модели в виде компьютерных программ, правильно и в разумное время обрабатывающих любые тексты на естественном языке;

– чрезмерный объем информации о слове, который делает процесс занесения этой информации в систему и ее отладки практически невыполнимым.

Структура, состав и объем словарной статьи переводных автоматических словарей и терминологических банков довольно четко отражают то промежуточное положение, которое они занимают между системами МП и традиционными словарями в книжной

форме. В разных автоматических переводных словарях эти параметры весьма различны. Словарные статьи могут содержать в вариативных сочетаниях следующие сведения: заглавное слово, один или несколько его иноязычных эквивалентов, тематическую принадлежность заглавного слова и его иноязычных эквивалентов, грамматическую информацию, неформализованные толкования или стандартизованные дефиниции, контексты, лексическую сочетаемость, семантические иерархические связи заглавного слова, сокращения, стилистическую окрашенность, лингвогеографические ограничения, ключевые слова (если заглавная лексическая единица является словосочетанием), источник и дату поступления, фамилию автора словарной статьи, различную служебную информацию и неформальные комментарии.

Колебания в составе словарных статей в разных автоматических переводных словарях весьма велики. Общими компонентами фактически являются только само заглавное слово, его тематическая принадлежность и иноязычный переводной эквивалент. Что же касается способов и глубины описания всех лексикографических параметров, то в этом мнения разработчиков разных автоматических переводных словарей расходятся еще больше.

5.13. Место переводного словаря в процессе перевода

Для того чтобы определить место переводного словаря в процессе перевода, необходимо первоначально решить вопрос о разделении всего процесса перевода на отдельные последовательные этапы. Следует отметить, что в подходах к его решению в теории машинного перевода и в теории традиционного «человеческого» перевода существуют заметные различия.

В практике машинного перевода, начиная с создания самых первых систем МП, естественным путем выделялись этапы вычленения лексических единиц в тексте, морфологического и синтаксического анализа, поиска по словарю входного языка для определения переводных эквивалентов, синтаксического и морфологического синтеза и другие более мелкие и специфические этапы этого процесса. Все названные стадии в различных теоретических моделях машинного перевода сводятся к трем этапам: анализу, промежуточному этапу, основной задачей которого является выбор переводных эквивалентов для лексических единиц входного языка, и синтезу. В промежуточный этап включаются поиск по словарю, разрешение неоднозначности лексическими средствами, обработка устойчивых и фразеологических словосочетаний, установление лексических и синтаксических соответствий между входным и выходным языками. Этот этап в ряде работ по машинному переводу получил название «трансфер».

В теоретических работах по традиционному человеческому переводу в этом вопросе нет единодушия, характерного для исследований по машинному переводу. Некоторые теоретики перевода в своих построениях не делят в явном виде процесс перевода на этапы. Так, А. Д. Швейцер пишет, что анализ переводов свидетельствует о том, что, как и любая речевая деятельность, он, по всей вероятности, не осуществляется по единой модели. В процессе перевода находят применение и грамматические трансформации, и лексико-синтаксические перифразы, и семантические преобразования, соответствующие модели «ситуация — текст». При этом выбор оптимального способа анализа исходного текста и построение соответствующего текста на другом языке диктуется конкретными условиями межъязыкового коммуникативного акта. Думается, что наряду с предварительной обработкой исходного текста путем грамматических трансформаций или лексико-синтаксического перефразирования вполне возможно установление прямых соответствий между высказыванием на языке оригинала и соответствующим высказыванием и последующим его редактированием на языке перевода¹. В то же время другие исследователи вполне однозначно выделяют в процессе перевода отдельные последовательные этапы.

Основная задача традиционного переводного словаря — дать соответствия типа лексическая единица языка оригинала — лексическая единица (или единицы) языка перевода. Переводной словарь косвенным образом помогает также решать некоторые задачи анализа текста оригинала за счет включения в предисловие грамматического описания входного языка, а также за счет включения в словник в качестве отсылочных статей некоторых словоформ входного языка, образуемых нерегулярными способами. Составители традиционных переводных словарей иногда вводят в них грамматический очерк выходного языка. В таких случаях можно говорить, что эти словари до некоторой степени помогают осуществлять синтез текста.

Тем не менее анализ входного и синтез выходного текста не являются основными задачами традиционных терминологических словарей. Их составители, как правило, задач такого типа перед собой даже не ставят.

5.14. Практическое использование автоматического переводного словаря

Многоязычный электронный словарь АBBYY Lingvo. В настоящее время в продаже имеется двенадцатая версия многоязычного словаря Lingvo российской фирмы АBBYY. В последние годы эта

¹ См.: Швейцер А. Д. Перевод и лингвистика. — М., 1973. — С. 61.

словарная система развивалась небывалыми темпами. Каждый год появлялась ее новая версия. По лексическому наполнению словарей, входящие в Lingvo, легко выигрывают в конкурентной борьбе с традиционными «бумажными» словарями. Оперативность публикации в них значительно превышает традиционные технологии. Тематика словарей отвечает самым современным запросам.

Словарь Lingvo 12.0 — многоязычная компьютерная система словарей. До одиннадцатой версии (2006 г.) фирма шла в основном по пути наращивания количества словарей и суммарного словарного объема, и лишь незначительно совершенствовала программные средства. Двенадцатая версия — это абсолютно новое решение лексического обеспечения автоматизированного рабочего места переводчика. Достаточно обратить внимание на то, что только англо-русская и русско-английская части словаря представляют собой огромную библиотеку из 47 профессиональных словарей общим объемом 2 307 740 словарных статей. Коллекция словарей, входящих в «английскую часть» Lingvo 12.0, делает его абсолютно универсальным и достаточным для современного автоматизированного рабочего места переводчика.

В связи с практически сплошной компьютеризацией текстовой деятельности в современном переводе комфортность использования электронных словарей резко повышает спрос на них.

Библиотеку словарей Lingvo 12.0 отличает новейший дружелюбный интерфейс с очень приятной и легкой эстетикой оформления всех карточек и сообщений. После обработки запроса поис-

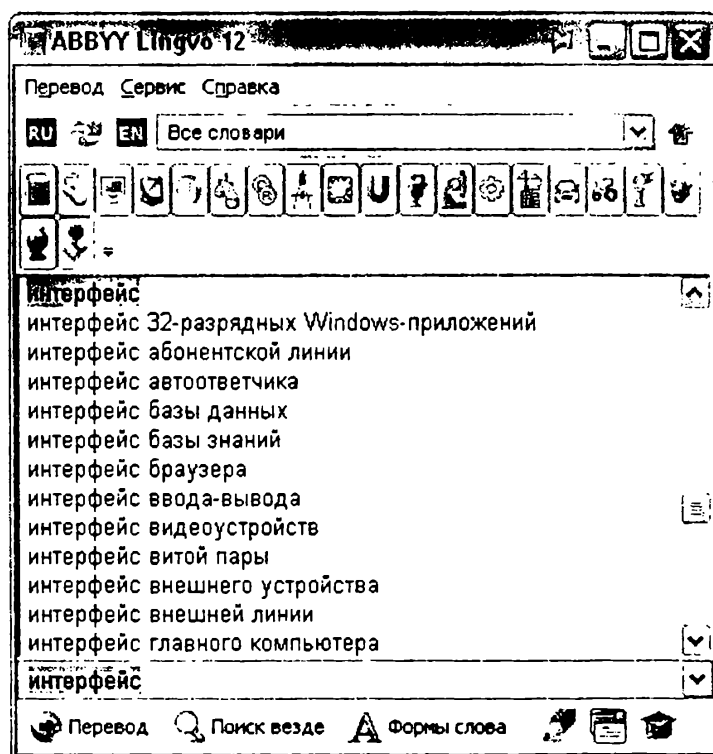


Рис. 25

ковая система выдает пользователю одновременно два окна. На рис. 25 показано первое основное окно, в котором выведен фрагмент всей возможной для данной тематики словарной базы. Вывод достаточно большого списка удобен для переводчика тем, что по нему он еще может уточнить свой запрос, увидев близкие словосочетания с запрашиваемым термином. Длина списка — это параметр, легко настраиваемый пользователем.

Второе окно показано на рис. 26. Левое его поле содержит всю имеющуюся в словарной системе информацию о запрашиваемом термине. В правом дан список словарей, в которых можно найти сведения о нем.

Систему характеризует высокая степень «интеллектуальности» поиска: если введено задание с неточностью или опечаткой, поиск все же осуществляется (рис. 27), в итоге выводится меню с разными вариантами, в той или иной степени соответствующими запросу.

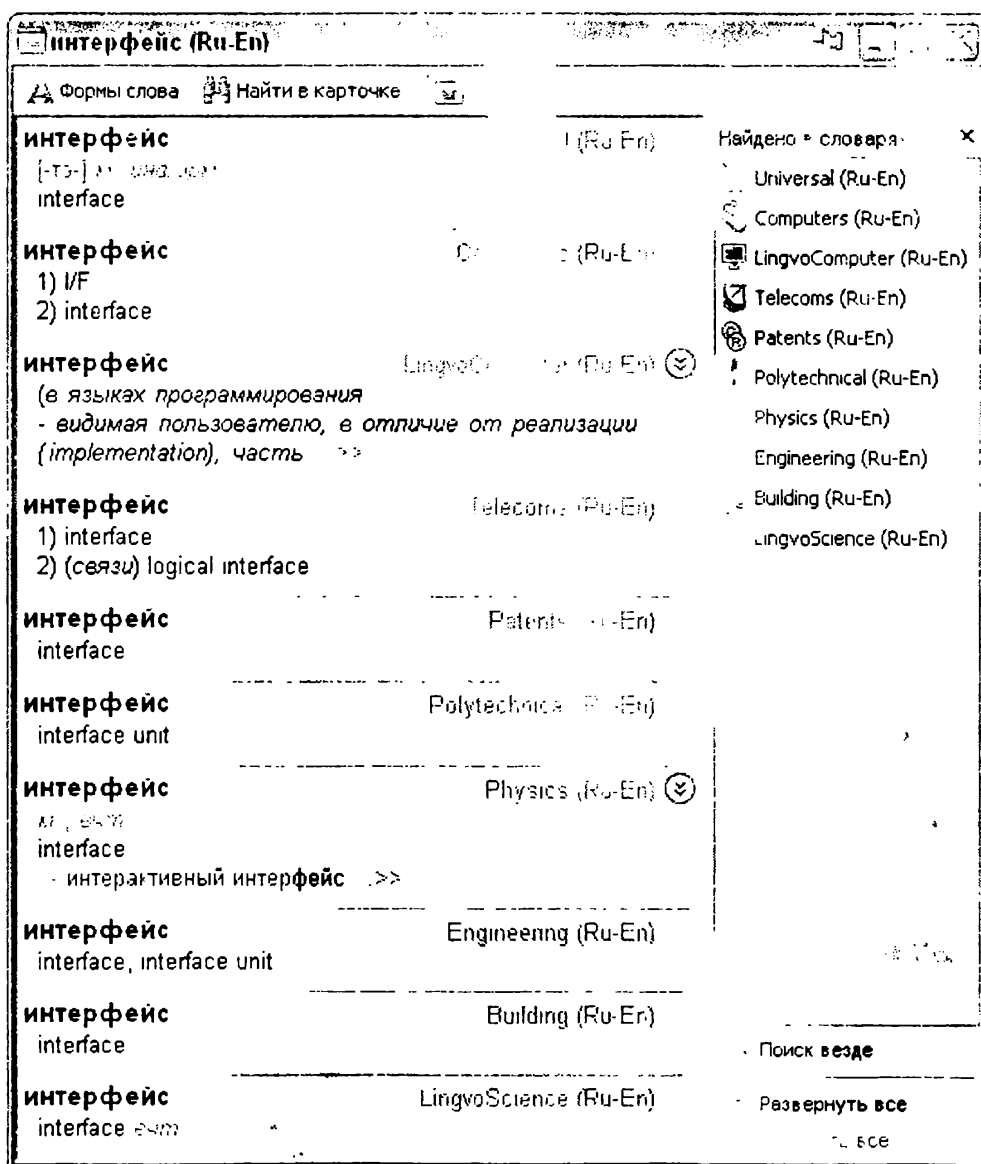


Рис. 26

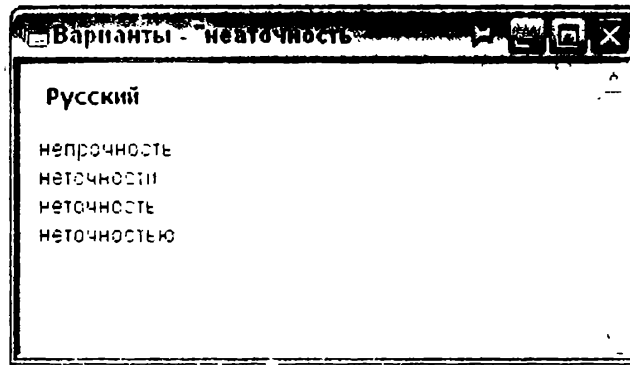


Рис. 27

Поскольку Lingvo полностью совместим с текстовым редактором MSWord и открывается в его среде, нужную информацию из словаря можно просто отметить с помощью «мыши» и «перетащить» в текст.

Удобным новшеством последней версии Lingvo является всплывающая подсказка, что продемонстрировано на рис. 28 (для иллюстрации мы воспользовались текстом предыдущего абзаца). При загруженном Lingvo чтение или просмотр незнакомого материала оказывается очень комфортным: нужно только подвести «мышь» к любому слову и подсказка всплывает автоматически. При необходимости по выбранной подсказке можно вызвать полную информацию о слове из всех словарей.

Lingvo — система многоязычная. Современный переводчик чаще всего владеет более чем одним иностранным языком. Поэтому многоязычная библиотека — это значительный шаг вперед по пути повышения комфорта труда переводчика. Никакой перенастройки для перехода с одного языка на другой не требуется — следует

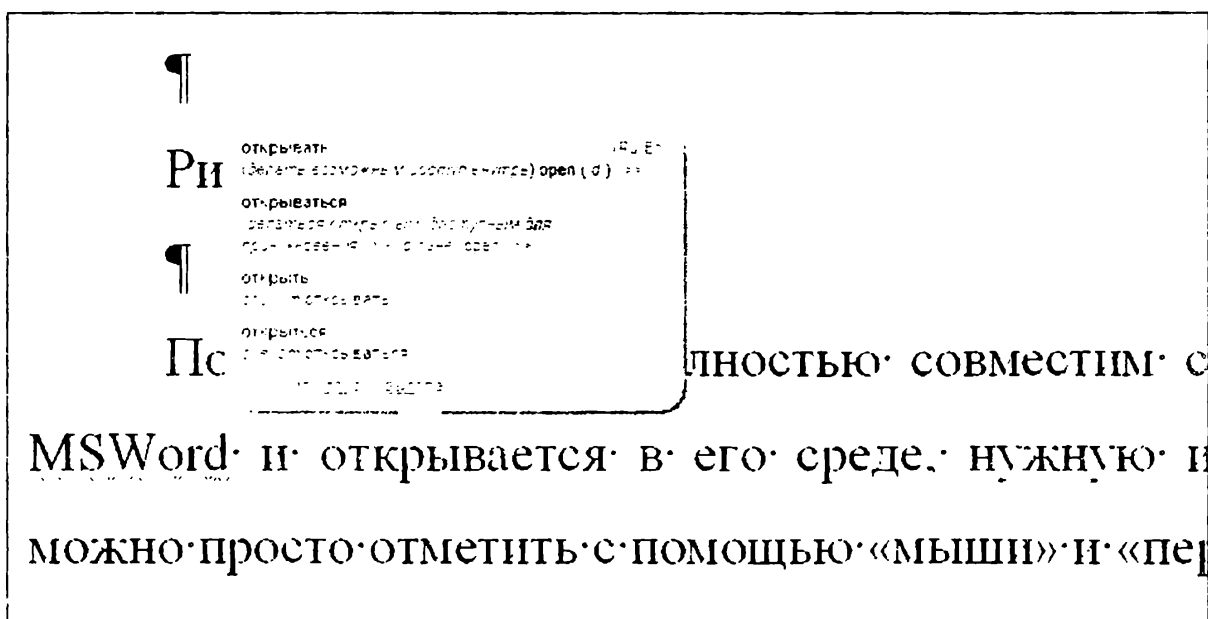


Рис. 28

лишь указать пару языков в простом меню. Для этого не нужно перезагружать программу, что важно в работе со смешанными текстами и с текстами, имеющими значительные включения на другом языке (цитаты, ссылки, примечания и др.).

В составе Lingvo есть несколько сервисных блоков. Грамматический блок отображает парадигмы как русских слов, так и иностранных; фонетический — дает транскрипции, а произношение наиболее частотных слов можно услышать в дикторском исполнении.

Кроме того, новая версия стала многоплатформенной: Lingvo можно установить на персональный компьютер, на карманный компьютер Pocket PC и на смартфон.

Одна из многих сервисных функций — возможность фиксирования своего собственного словаря внутри общей системы. В работе с Lingvo предусмотрено быстрое добавление личных словарных карточек и редактирование уже имеющихся.

Контрольные вопросы и задания

1. Что понимается под лексическим составом языка?
2. Каковы причины лексических ошибок в переводе?
3. Назовите основные проблемы формирования современного словаря.
4. На примере своего электронного словаря выпишите в конспект известные вам преимущества электронного словаря по сравнению с традиционным.
5. Назовите определяющие признаки системности терминологии.
6. Назовите и кратко охарактеризуйте три ведущих компонента автоматического переводного словаря.
7. Каковы основные аспекты асимметрии языковых вариантов терминосистемы?
8. Назовите наиболее распространенные типы параллельных текстов.
9. Что понимается под обратимостью переводного словаря?
10. Проведите поиск какого-либо слова в электронном словаре на вашем компьютере. Найдите обратный перевод полученного эквивалента в этом же электронном словаре. Сравните обе словарные статьи и дайте характеристику обратимости вашего электронного словаря.
11. Назовите наиболее значимые средства повышения гибкости автоматического словаря.
12. Попробуйте найти одно и то же слово в небольшом учебном словаре и в самом большом доступном вам «бумажном» словаре. В последнем случае вы, возможно, получите больше информации (не всегда необходимой), но ощутите неудобство сложности поиска. Найдите это же слово в электронном словаре, объем которого, безусловно, превышает объем самого большого «бумажного» словаря. Оцените таким образом зависимость удобства и надежности работы от объема электронного словаря.

ПАМЯТЬ ПЕРЕВОДОВ — TRANSLATION MEMORY**6.1. Истоки ТМ-инструментов**

Рассматривая проблемы систем МП, мы уже упоминали японского ученого Макото Нагао из университета Киото. В 1982 г. в одной из своих статей он предложил новую концепцию машинного перевода, которая была основана на утверждении, что тексты должны переводиться по аналогии с текстами, ранее переведенными вручную. Для этого целенаправленно формируется большой массив билингв (исходных текстов и их переводов), который вводится в мощный компьютер с большим объемом памяти. В процессе перевода новых текстов специальная поисковая система выбирает аналоги фрагментов этих текстов из массива билингв. М. Нагао назвал свой подход к машинному переводу «Example based translation» (перевод, основанный на примерах). Идея М. Нагао была использована некоторыми разработчиками систем МП, но наибольшее распространение она получила в связи с развитием новейшей технологии перевода — систем класса Translation Memory (память переводов), которые в последнее время чаще называют ТМ-инструментами.

6.2. Идеология ТМ-инструментов

ТМ-инструменты предназначены для сохранения пар предложений в переводческой базе данных. Каждая такая пара состоит из предложения из оригинала и его перевода на другом языке. Возможно размещение фрагментов текста и других форматов — превышающих длину предложения или являющихся лишь его частью. Но в автоматическом режиме сохраняются именно предложения, поэтому довольно часто такие программы называют «память предложений» (от англ. Sentence Memory).

Принцип действия ТМ-инструментов прост. Для овладения этими программами не требуется никаких дополнительных знаний, а навык пользования ими приходит очень быстро. Как правило, они интегрируются с такими обычными офисными программами, как, например, Word. Некоторые из них имеют собственные средства редактирования текста, интерфейсы которых мало отличаются от привычных для современного переводчика интерфейсов текстовых редакторов.

Современный ТМ-инструмент представляет собой сложную компьютерную систему, в которую входит множество программ, обладающих различными функциями. Эти функции можно подразделить на две группы — основные и сервисные. Таким образом построена, например, наиболее распространенная система TRADOS, основными программами которой являются Translator's Workbench и MultiTerm.

Программа Translator's Workbench автоматически вводит предложение, подлежащее переводу, в базу данных (память переводчика). Обычно длина записи определяется точками — от точки до точки. Затем также автоматически к исходному предложению в базе данных прикрепляется выполненный перевод. Закончив работу, переводчик получает выполненный им перевод, а в базу данных помещаются два фрагментированных параллельных текста. Следующий переводимый текст будет обработан таким же образом и так же размещен в базе данных.

В первое время использование ТМ-инструмента никак не повлияет на производительность труда переводчика. Но база данных расширяется довольно быстро — со скоростью работы переводчика. В момент начала работы над переводом любого нового предложения ТМ-инструмент автоматически проверяет в базе данных наличие такого же или похожего среди переведенных ранее.

Общей тенденцией в современном переводческом бизнесе становится жанрово-тематическая специализация. Это означает, что и бюро переводов, и индивидуальные переводчики, как правило, выполняют однотипную работу. В настоящее время на потребительском рынке одновременно предлагается несколько десятков, например, электрических чайников; каждый сопровождается инструкцией для пользователя. Тексты этих инструкций однотипны. После перевода первой из них во второй переводчику встретится не более половины нового текста. Если первая инструкция была переведена с использованием ТМ-инструмента, во второй предстоит перевести лишь отличающиеся фрагменты, а совпадающие будут предоставлены переводчику ТМ-инструментом. Ему остается либо согласиться с полученным предложением, либо отредактировать его в необходимой мере.

Количество однотипных текстов так же огромно, как огромен потребительский рынок. Но дополнительным мощным источником частично совпадающих текстов является постоянное совершенствование существующих товаров. Примером могут служить «Руководства пользователей» многочисленных программных продуктов. Вскоре после выпуска на рынок версии 5.0 некоего программного продукта, для которого было полностью переведено на русский язык «Руководство пользователя», появляется его новая версия — 5.01. С большой вероятностью можно быть уверенным, что «Руководство пользователя» для нее будет отличаться от

предыдущего не более, чем на 10—15%. Но оно также должно быть издано на русском языке, следовательно, его тоже нужно переводить. Если оно представляет собой объемный текст, да еще со сложной структурой, то его просмотр, поиск несовпадений и механическая коррекция могут оказаться делом весьма трудоемким и не имеющим гарантий того, что какие-то мелкие изменения останутся незамеченными. ТМ-инструмент с недоступной для человека педантичностью просмотрит весь новый текст по фразам, предложит варианты полностью совпадающих предложений, обратит внимание на частично совпадающие, выделив в них отличающиеся фрагменты, и покажет возможности редактирования несовпадающих предложений. Большим достоинством ТМ-инструментов является еще и то, что они полностью сохраняют при переводе формат текста-оригинала. Таким образом, переводчик освобождается от необходимости следить за используемыми шрифтовыми выделениями, абзацными отступлениями, табуляцией и т.д.

Самое важное преимущество использования ТМ-инструментов — освобождение переводчика от повторяющихся монотонных операций. Возможно, эти программы не повышают производительность труда отдельного человека, но они резко увеличивают производительность переводческого бизнеса.

Программа *Translator's Workbench* располагает возможностью различных видов настройки на выполняемый перевод. Рисунок 29 иллюстрирует варианты настройки, которые могут быть использованы при переводе. Программа предлагает выбрать рабочую величину совпадений (*Minimum Match Value %*). В приведенном на рисунке варианте выбрано 70%, это соответствует характеристике совпадений текста оригинала с имеющимися в базе данных билингвами. Какова логика выбора этой величины? Предположим, что предложение оригинала состоит из 10 слов. Если назначить минимальный объем совпадений 10%, то можно ожидать, что практически все переводимые предложения попадут в выборку программы и будут выведены как вариант перевода, ведь служебные слова также участвуют в определении объема совпадений. Если объем совпадений установить в диапазоне 90—100%, подсказки программы будут чрезвычайно редки, и ее работа окажется неэффективной. Объем совпадений 70% — наиболее рентабельный режим работы системы.

Второй параметр настройки — количество совпавших вариантов, имеющихся в базе билингв (*Maximum Number of Hits*). При выполнении переводов различных текстов могут встречаться почти одинаковые предложения, но в зависимости от контекста их переводы будут отличаться. Программа *Translator's Workbench* позволяет накапливать несколько возможных вариантов билингв. Если в базе уже содержится от одного до пяти билингв, то они все

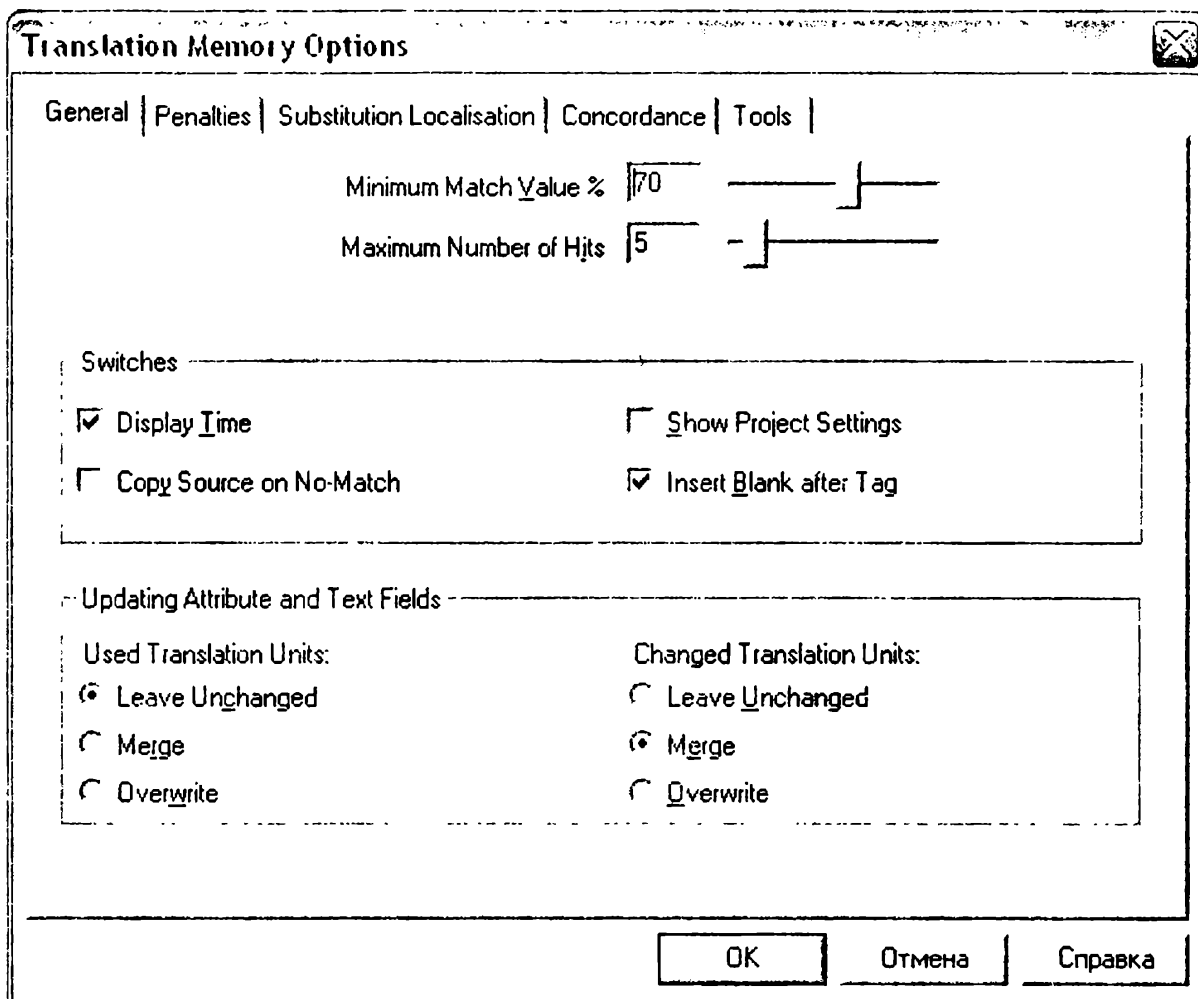


Рис. 29

будут предложены для выбора, так как из рисунка видно, что заказано именно «5». Если их имеется больше, то все же будут предложены только пять.

Остальные опции настроек касаются некоторых сервисных программ, которые не входят в наше описание.

Большинство ТМ-инструментов существуют в двух версиях: для инсталляции на персональном компьютере и для сетевого использования. Инсталлированный в локальной сети какого-либо бюро переводов ТМ-инструмент превращает всю переводческую деятельность бюро в коллективную. В создании базы данных участвуют все его сотрудники, а наличие и использование общей базы приводит к выработке единого переводческого стиля всего бюро и значительно повышает качество выпускаемых переводов, поскольку решение многих переводческих проблем также становится коллективным. При срочном переводе больших текстов часто приходится включаться в работу сразу несколькими переводчикам. Без использования ТМ-инструмента совместить фрагменты, выполненные несколькими людьми, — задача не из легких. Основной принцип перевода — единство терминологии — неизбежно будет нарушен. ТМ-инструмент в таком случае оказывается един-

ственным средством выхода из сложной ситуации, поскольку он возьмет на себя функции координатора действий и результатов отдельных переводчиков. Повышение качества работы бюро автоматически происходит еще и потому, что база данных ТМ-инструмента находится в постоянной работе, т.е. пополнение базы осуществляется и в режиме редактирования. А совершенствование этого режима, в свою очередь, открывает возможности для исправления ошибок и недопущения их в последующих переводах.

Своеобразное разделение труда (скорее — использование лучших достижений) может быть достигнуто и в условиях индивидуальной переводческой деятельности на персональном компьютере. В большинстве действующих ТМ-инструментов предусмотрено подключение к уже накопленным на других компьютерах базам данных.

Не меньшее значение для повышения качества и производительности переводческого бизнеса имеет оперативное терминологическое обеспечение. Существует много электронных переводных словарей, которые интегрируются с текстовым редактором MSWord. Они хорошо зарекомендовали себя, получили всеобщее признание и широкое распространение не только среди переводчиков. Эти словари постоянно пополняются. Во многих случаях можно регулярно через Интернет пополнять лицензионную копию словаря на своем персональном компьютере. Периодически и довольно часто в продажу поступают дополненные версии имеющихся словарей. Но в современной ситуации этого не всегда бывает достаточно. Многие бюро переводов и индивидуальные переводчики создают свои узкоспециализированные словари терминов и профессионализмов. У большинства ТМ-инструментов имеется функция терминологического обеспечения. В ТМ-инструменте TRADOS эту функцию выполняет вторая основная программа MultiTerm. При обнаружении в переводимом тексте слова или словосочетания, которое в прежних переводах было помечено как термин, MultiTerm обращает внимание переводчика на этот фрагмент текста. Таким образом, реализуется еще один очень важный параметр повышения качества перевода — принцип единства терминологии, который должен неукоснительно соблюдаться. Естественно, в условиях локальной компьютерной сети этот принцип будет соблюдаться всем коллективом переводческого бюро.

При подключении программы MultiTerm интерфейс ТМ-инструмента TRADOS имеет вид, показанный на рис. 30. Правое его поле предназначено для вывода информации из терминологической базы. Эта терминологическая база формируется так же, как и база билингв, т.е. в процессе работы переводчика или бюро переводов, и не требует больших трудозатрат. Данная функция легко вызывается после запуска программы MultiTerm (рис. 31).

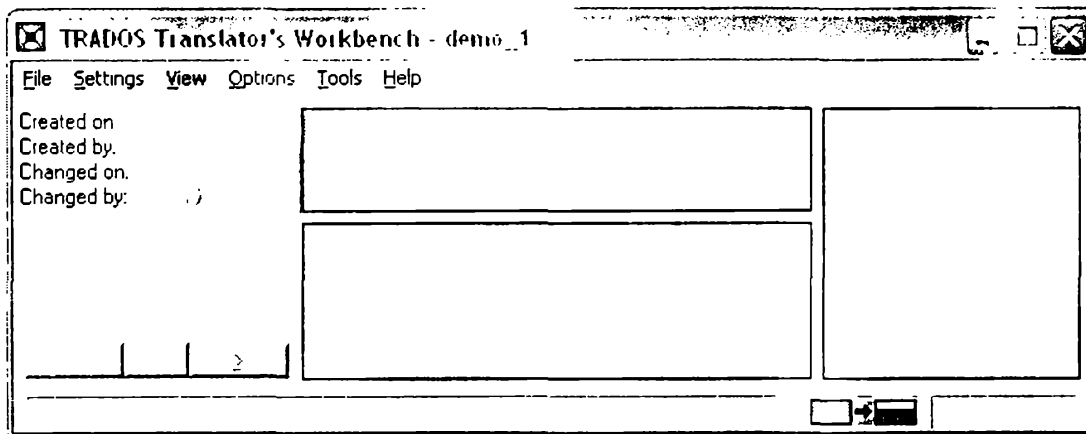


Рис. 30

В имеющейся в настоящее время версии ТМ-инструмента TRADOS 5 Freelance отсутствует блок русского языка. Но принципиально система не зависит от языка и очень мало зависит от направления перевода, что не мешает использовать ее при переводе на русский язык и создавать на этом языке базы билингв и терминов.

Важной особенностью терминологической базы программы MultiTerm является возможность многоаспектной структуризации. На рис. 32 приведен довольно большой список рубрик, по которым может быть распределена вводимая в базу терминология.

На рис. 33 показано поле для ввода информации. Им можно воспользоваться дважды: для ввода исходного термина и для ввода эквивалентов.

В настоящее время уже существуют готовые терминологические базы данных, которые могут быть подключены к ТМ-инструмен-

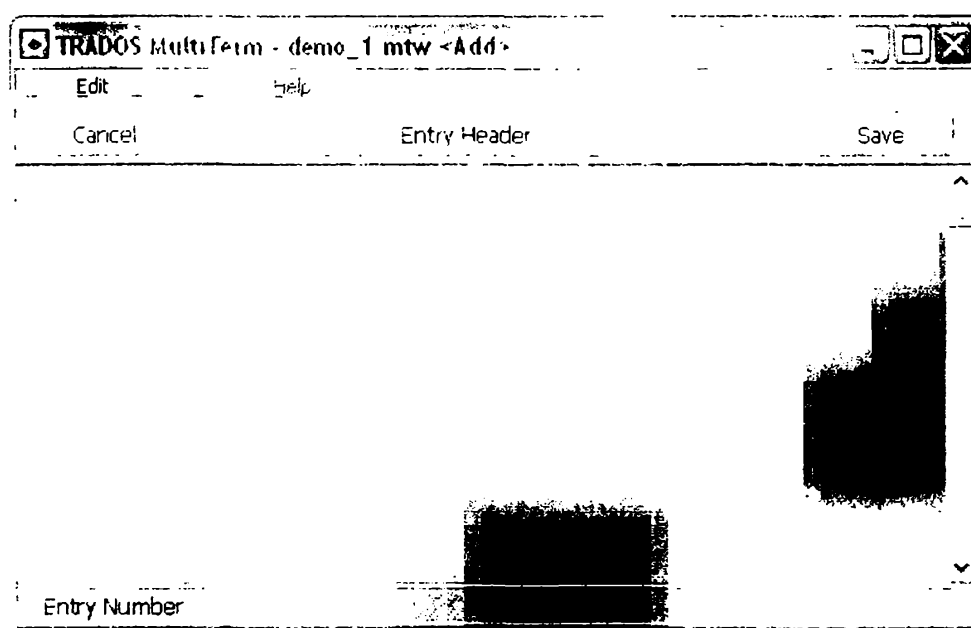


Рис. 31

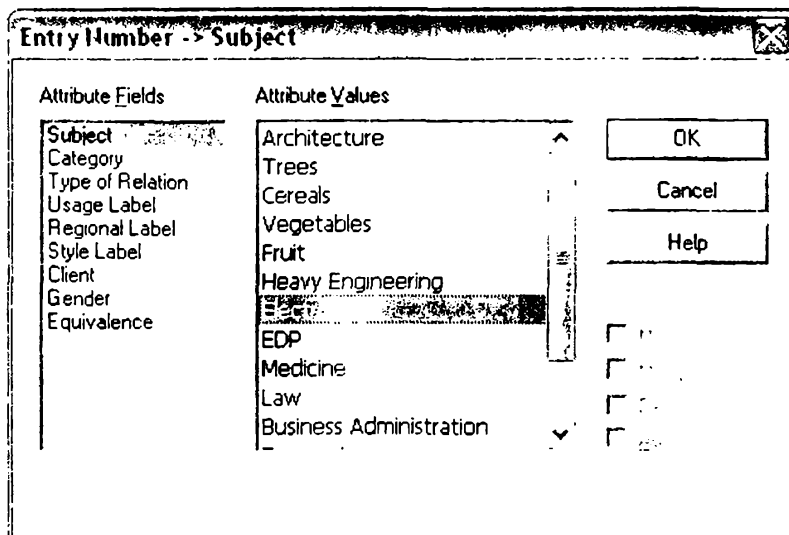


Рис. 32

там в виде дополнительных, сменных или временных блоков. Такие терминологические списки, как правило, создаются производителями экспортируемой продукции. При выдаче заказа на перевод сопроводительной документации к своей продукции фирмы обычно передают переводчику терминологическую базу в виде временного блока для ТМ-инструмента, что не только ускоряет процесс перевода и повышает его качество, но еще и унифицирует документацию фирмы.

Программа MultiTerm может работать независимо от других программ системы TRADOS. Это означает, что она может быть использована как оболочка для формирования любой персональной справочной системы. Допустимый объем записей достаточно велик для создания справочной системы, например, по рабочим документам, их фрагментам или цитатам.

Информационное поле позволяет размещать в нем не только текстовую информацию со сложной структурой (таблицы, шриф-

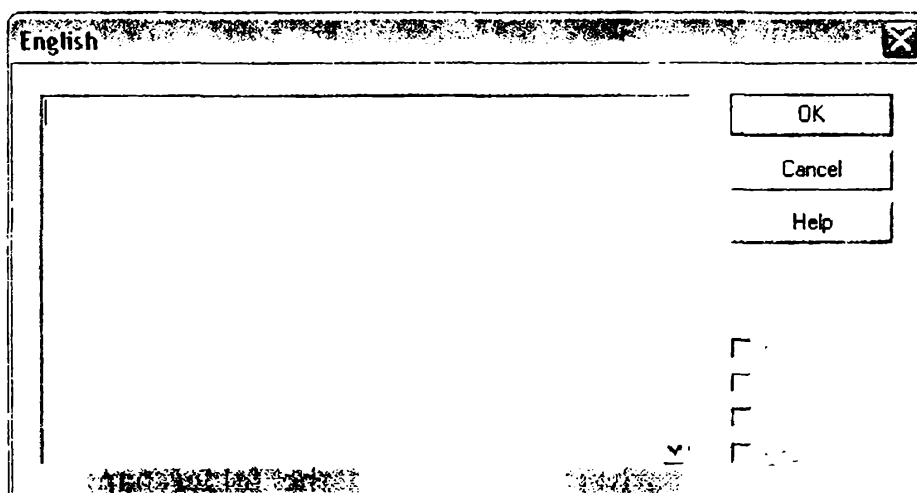


Рис. 33

товые выделения, графические символы или пиктограммы), но и изображения. Примером могут служить иллюстрированные номенклатурные списки продукции фирмы и т.п. Наиболее типичные объемы параметров терминологической базы таковы (цифровые данные отражают современные предельные возможности — максимум):

- терминологическая база данных может содержать 20 языков;
- длина записи — 32 тыс. символов;
- каждая запись может иметь 500 полей;
- длина поля может составлять 4 096 символов;
- допускается включение графики;
- интерфейс, существующий между MultiTerm и Word, обеспечивает (для пользователя) возможность вставки терминов из MultiTerm непосредственно в документ;
- интерфейс активного распознавания терминологии, существующий между MultiTerm и Translator's Workbench, предоставляет (для Workbench) возможность автоматического поиска терминов в MultiTerm с последующей их вставкой в документ в процессе перевода.

6.3. Практическое использование системы TRADOS 5 Freelance

Рассмотрим конкретный пример перевода с использованием ТМ-инструмента.

Имеются две инструкции для пользователей по быстрой установке запоминающих устройств, произведенных различными фирмами.

Инструкция 1:

EasyDisk Installation Guide

EasyDisk is a removable electronic hard disk with the storage media of flash memory. Data transferring between personal computer and EasyDisk can be accomplished via the USB port.

System Required.

Hardware: IBM PC, Compatible PC or Macintosh with USB interface.

Operating System. Windows 98/SE, Windows ME, Windows 2000/XP, Macintosh and Linux, which support USB mass storage protocol.

Installation Method.

Windows 98/SE System

Installation Method:

Insert your EasyDisk driver diskette, find the folder named Driver for Win98, then double click «Setup, exe». An information bar will prompt, click OK button. Select install or update driver and click Start Button, then system

will install the driver automatically. The progress bar is filled full and done prompt displayed, click Exit button, plug EasyDisk into USB port, system will install the hardware automatically.

Инструкция 2:

MW Quick Installation Guide

MW is an electronic mini hard disk. Data transferring between personal computer and your new MW can be managed via the USB port.

System Requirements.

Hardware: IBM PC, Compatible PC with USB interface.

Operating System. Windows 98/SE, Windows ME, Windows 2000/XP, which support USB mass storage protocol.

Installation Method.

Windows 98/SE System

Installation Method 1:

Insert your MW driver CD in your PC. Find the folder named «Driver for Win98», then select «Setup, exe». An information bar will prompt. Select OK button. Select install driver and click Start Button. The system will install the driver automatically. Click Exit button.

Plug MW into USB port. Your system installs the hardware automatically. Now you can operate MW as your local hard disk.

Один из переводчиков, окончивших Национальные курсы повышения квалификации при Всероссийском центре переводов научно-технической литературы и документации, предложил следующий перевод первой инструкции, который вполне можно признать удачным.

Руководство по установке ED

ED — это съемный электронный жесткий диск с флеш памятью большого объема. Обмен данными между ПК и ED может осуществляться через порт USB.

Требования к системе:

Аппаратные средства: IBM PC, совместимые ПК или Macintosh с USB-интерфейсом.

Операционная система: Windows 98/SE, Windows ME, Windows 2000/XP, Macintosh и Linux, которые поддерживают USB протокол хранения данных.

Способ установки:

Для систем Windows 98/SE

Вставьте дискету с драйвером вашего ED, найдите папку с именем Driver for Win98, затем дважды щелкните по кнопке «Setup, exe».

Появится информационное окно с подсказкой; щелкните по кнопке ОК.

Выберите опцию установки или обновления драйвера и щелкните по кнопке «Start Button», после этого система установит драйвер автоматич-

чески. Когда указатель хода процесса заполнится и появится информация об исполнении, щелкните по кнопке «Exit».

Подсоедините ED к порту USB, система автоматически установит это устройство.

Сравнивая оригиналы обеих инструкций, нельзя не заметить, что их тексты различаются не более, чем на 15—20%. Было бы нерентабельно выполнять перевод второй инструкции традиционным способом, выдав задание другому переводчику. Поскольку первый предусмотрительно выполнил свой перевод на компьютере с установленной системой Trados 5 Freelance, то второй текст был переведен с минимальными затратами.

Наш ТМ-инструмент был настроен на 70%-ное совпадение. Это означало, что при совпадении менее, чем на 70%, текст рассматривается как новый и система подсказок не дает. Так, без подсказок пришлось перевести заголовки и первое краткое предложение. Сразу же после ввода второго предложения мы оказались в ситуации, отраженной на рис. 34. В верхней части экрана интерфейс программы Translator's Workbench, в нижней — обычный интерфейс текстового редактора MSWord. На верхнее поле интерфейса Translator's Workbench выводится сегмент (предложение) текста для перевода (это поле помечено значком текстового редактора

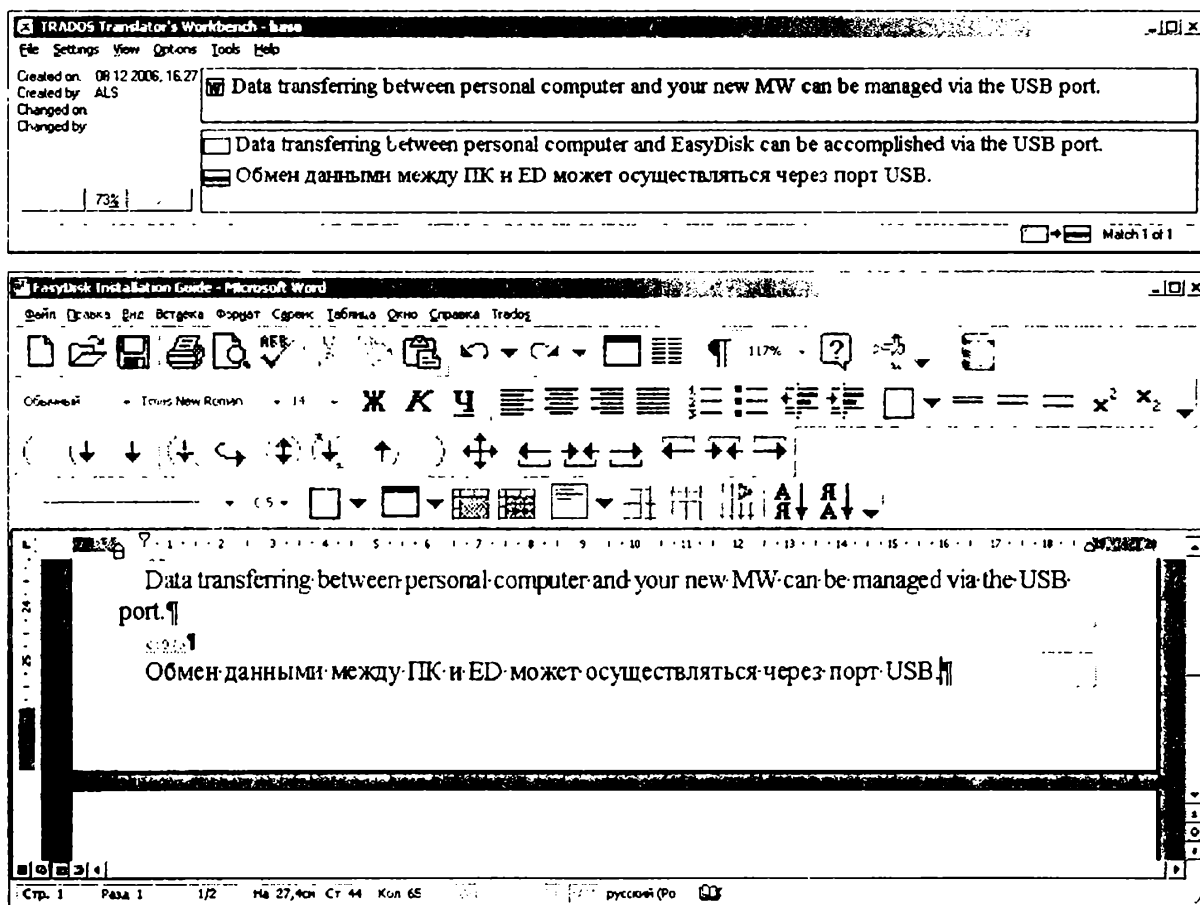


Рис. 34

MSWord). Затем в его нижнее поле выводится билингв, найденный в базе данных. В текстовом редакторе MSWord автоматически появляются два временных поля, выделенные разными подцветками. В верхнем из них предложение для перевода, в нижнем — билингв из базы.

ТМ-инструмент обнаруживает, что предложения отличаются наличием в первом двух дополнительных слов *your* и *MW*, а также несовпадением еще двух пар: *new* — *EasyDisk* и *managed* — *accomplished*. В такой ситуации, конечно, рациональнее не переводить предложение заново, а просто отредактировать то, что дала система. После редактирования, как это показано на рис. 35, выполняется несложная операция перехода к следующему предложению. Она заключается в нажатии «мышью» кнопки (⇅) интерфейса программы Translator's Workbench. Эта операция представляет собой единственное неудобство по сравнению с работой без ТМ-инструмента. Но высокая рентабельность новой технологии полностью компенсирует это маленькое неудобство.

При переводе следующего предложения (рис. 36) обнаружено полное соответствие, которое следует просто подтвердить. Но обратите внимание на то, что система полностью сохраняет оформление исходного текста: новая строка, курсив, двоеточие.

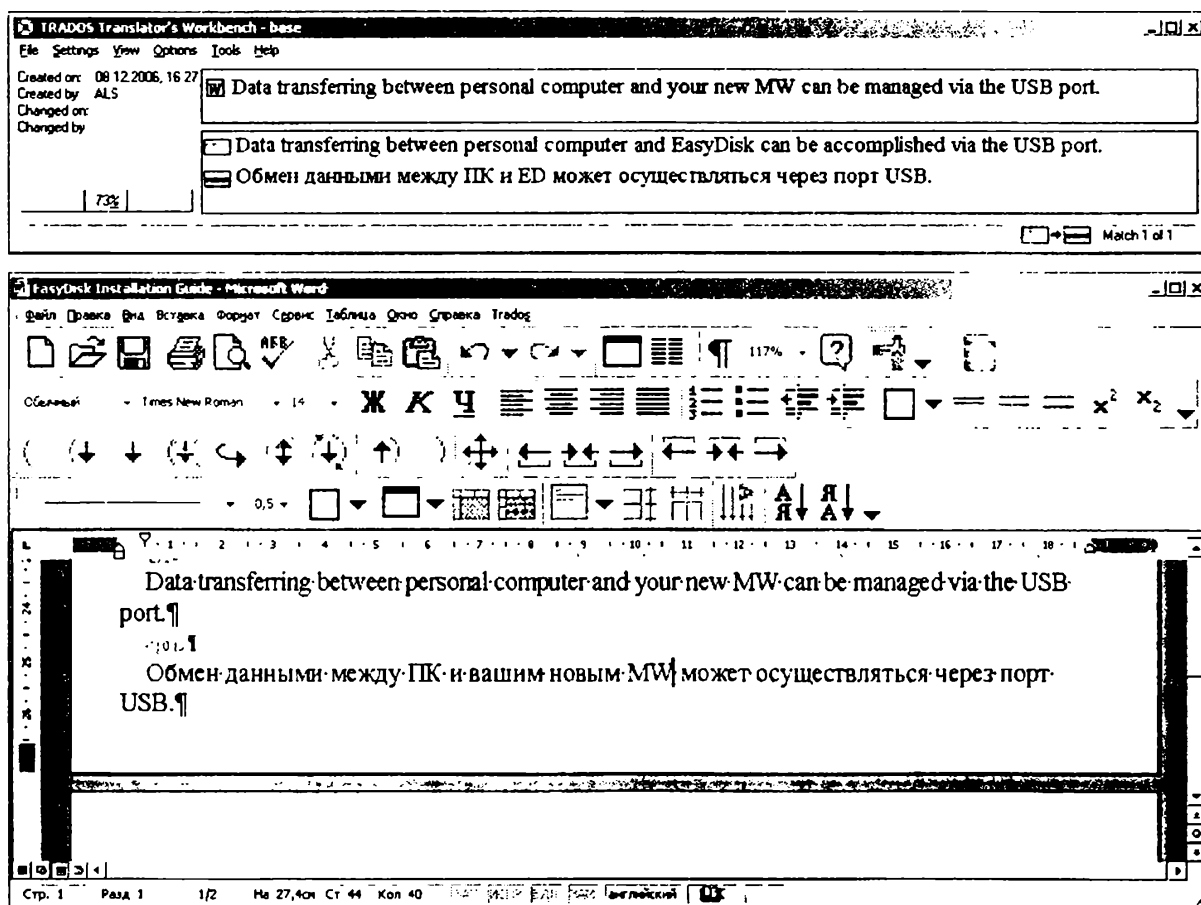


Рис. 35

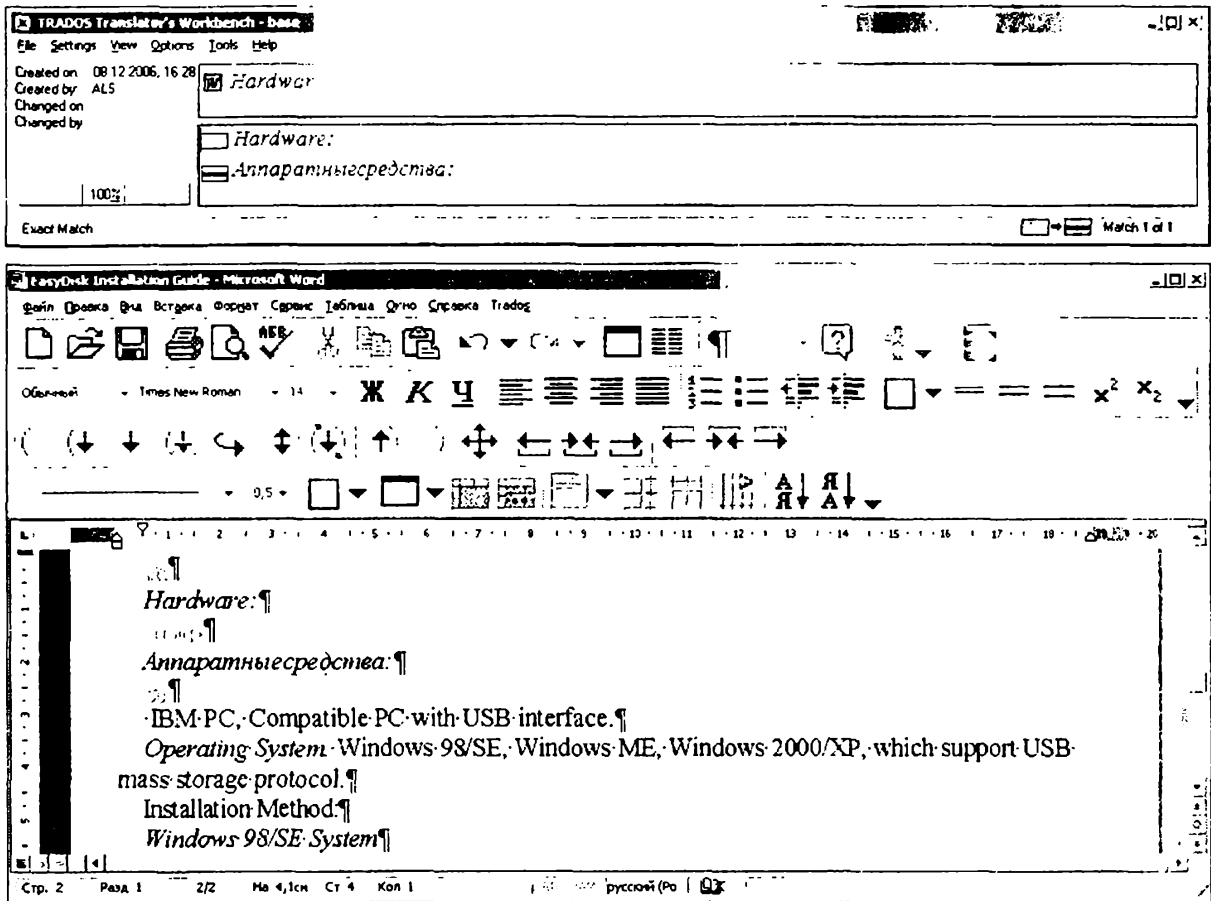


Рис. 36

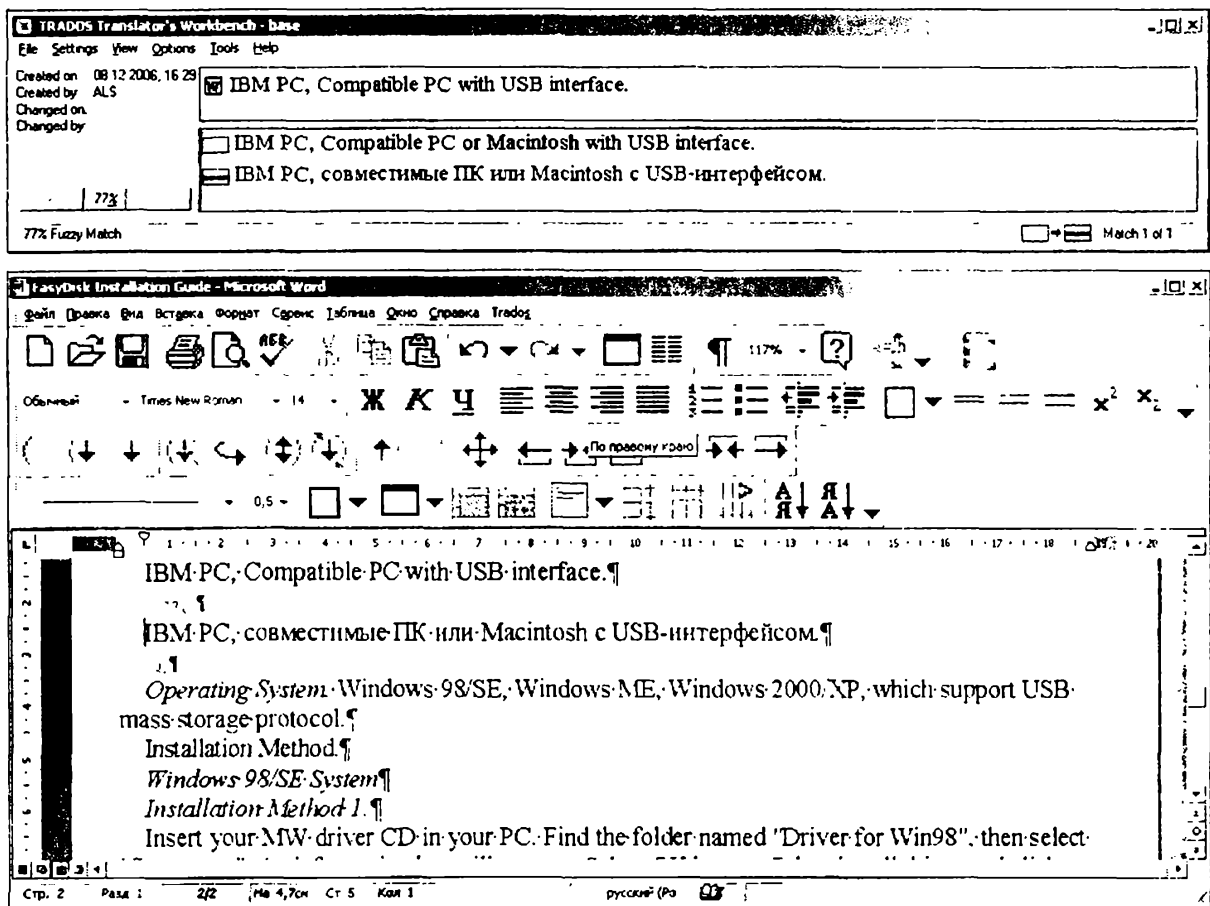


Рис. 37

Следующий пример, показанный на рис. 37, интересен тем, что при достаточно высоком совпадении нового предложения с имеющимся в базе данных, в нем все же отсутствуют два элемента (*or Macintosh*), на что система обращает внимание переводчика, выделяя их другим цветом.

Как видно из представленных иллюстраций, вся работа происходит в текстовом редакторе MSWord. Это означает, что переводчик может использовать все обычные возможности: обратиться к словарю, выполнить любые операции правки текста и т.д. Естественно, работа программы *Translator's Workbench* — процесс не непрерывный. Начать ее можно в любом месте текста. Если программа уже загружена в оперативную память, к работе она вызывается нажатием «мышью» кнопки (↓), а останавливается нажатием кнопки (↑). После того как программа остановлена (это необходимо делать обязательно!), можно обычным способом перейти к другому фрагменту текста в текстовом редакторе. Это очень удобно, когда переводу подлежит не весь текст, а лишь отдельные его фрагменты.

6.4. Общий принцип Translation Memory

База данных ТМ-инструмента состоит из пар сопоставленных друг с другом сегментов (TU — Translation Units) на языке источника и языке перевода. Обычно в качестве сегмента выделяется предложение, однако, по желанию пользователя, в качестве единиц перевода в базу данных можно вводить и фрагменты предложений, и отдельные абзацы.

Основной идеей концепции Translation Memory является исключение необходимости снова переводить предложение (или его фрагмент), которое когда-то ранее уже было переведено пользователем и занесено в базу данных (в «память переводчика»).

Предусмотрены два режима поиска соответствий: точные (*Exact matching*) и неточные соответствия (*Fuzzy matching*). В первом случае программа сопоставляет предложенный сегмент на языке источника с теми, которые имеются в базе данных, и предлагает его перевод только в том случае, если обнаруживает их полное соответствие. В результате же неточного поиска в базе биллингов в качестве варианта перевода программа выдает предложения, схожие с запрашиваемым сегментом. Степень сходства в режиме неточного поиска определяется самим пользователем в зависимости от того, хочет ли он просмотреть все предложения, имеющие хотя бы какую-то общую часть с текущим, или же он предпочтет обратиться только к тем, которые незначительно отличаются от сегмента, который ему нужно перевести. Таким образом, его задача сводится к переводу только тех фрагментов, которые им никогда ранее не переводились.

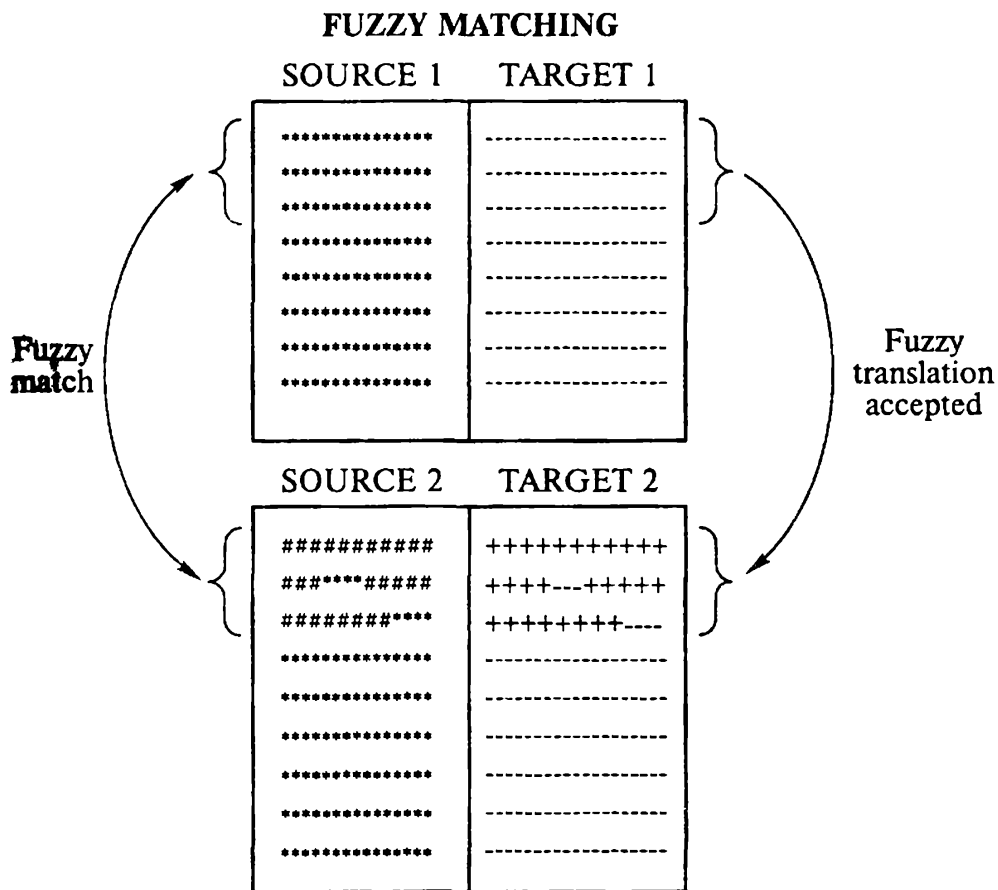


Рис. 39

Первая ситуация весьма характерна для начала работы сразу же после установки ТМ-инструмента на рабочее место переводчика, когда еще не накоплена достаточная база билингв, а терминологическая база еще мала. Но может быть и так, что переводчик уже располагает достаточным объемом баз по определенной тематике, но вынужденный переход на другую тематику требует создания новой тематической базы. Переводчик практически работает в традиционном режиме с использованием также и других средств автоматизации труда, но при этом программа Translator's Workbench уже запущена в его компьютере и все действия совершаются в ее среде. Схематически это выглядит следующим образом:

Новый перевод сегмента (текста) в среде Translator's Workbench.



Редактирование и подтверждение перевода.

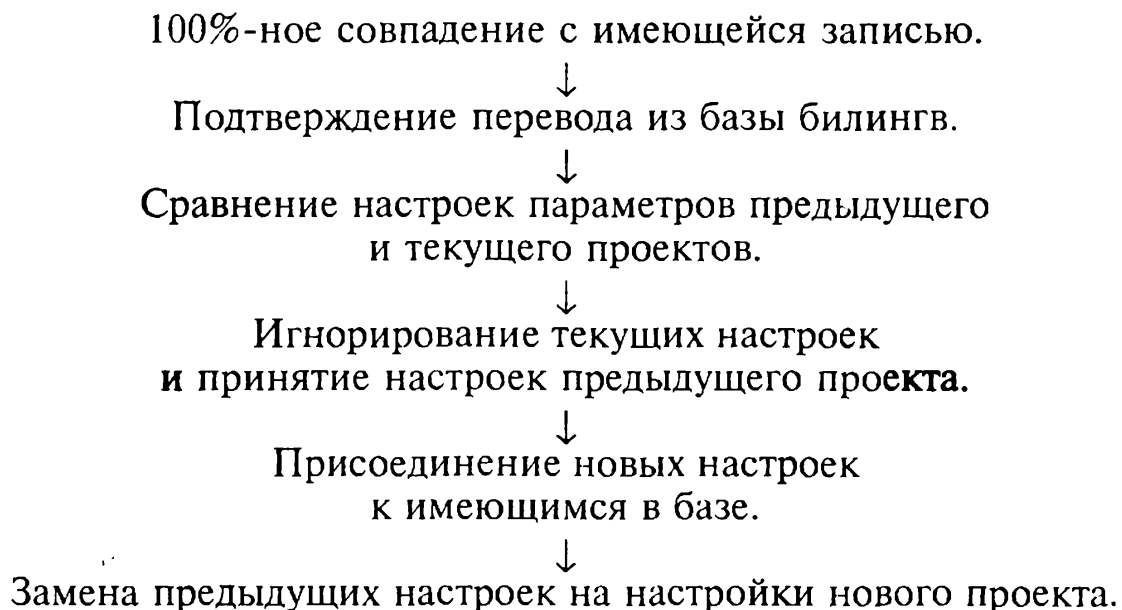


Создание билингвы, прикрепление служебной информации
и запись в базу.

Вторая ситуация возникает, когда программа Translator's Workbench находит полное (100 %) совпадение введенного для перевода сегмента (предложения) с имеющейся в базе данных

билингвой. Если переводчик принимает предложенный программой перевод без каких-либо изменений, он все же должен сделать выбор в отношении служебной информации, которая, как упомянуто выше, также имеет значение для работы системы. Поскольку служебная информация предыдущего проекта прикреплена к билингве, то в части перевода можно принять ее и проигнорировать текущие настройки программы. В части коррекции базы билингв можно прикрепить новые настройки к предыдущей записи или полностью заменить их на настройки текущего проекта-перевода.

Схематически операция выглядит так:



Третья ситуация также связана с нахождением программой Translator's Workbench полного совпадения с имеющейся в базе записью, но переводчик все же вынужден по определенным причинам вносить некоторые изменения в перевод. Это может быть вызвано требованиями контекста или какими-то дополнительными требованиями к переводу.

После получения переводного варианта из базы программы вступает в силу вторая ситуация — изменяется перевод и, следовательно, необходимо принять решение о редактировании записи в базе билингв. Запись может быть оставлена без изменений, если переводчик точно знает, что текущий проект имеет частный характер и вероятность его повторения очень мала. Она может быть дополнена новой служебной информацией или старая информация заменена данными нового проекта. Наконец, может быть сделана целиком запись новой билингвы вместе с новой служебной информацией.

Четвертая ситуация, в которой программа Translator's Workbench находит одно или несколько неполных совпадений. Переводчик выбирает одно из них, вносит изменения и подтвер-

ждает отредактированный перевод сегмента по методике третьей ситуации или, отвергнув все предложения программы, переводит сегмент традиционным способом и создает новую запись в базе билингв (по первой ситуации).

6.5. Сервисные программы

Программа WinAlign. Необходимой составляющей любого ТМ-инструмента является программа «Слияния параллельных текстов» (Alignment tool), которая позволяет использовать материалы накопленных переводов для построения базы билингв ТМ-инструмента. Работая с такой программой, переводчик указывает адреса файлов, содержащих оригинальные тексты и соответствующие им переводы, задает ряд параметров и правил для корректного сопоставления сегментов и затем, после автоматического разбиения текстов на сегменты и установления связей между ними, проверяет правильность связей и корректирует их вручную. На рис. 40 представлен весь технологический процесс создания базы билингв по оригиналам и переводам, выполненным не в среде ТМ-инструмента.

В такой ситуации необходимо вручную для каждого английского предложения в левой колонке найти соответствующее немецкое предложение в правой колонке и установить между ними связь для записи в базу билингв со всей сопроводительной служебной информацией.

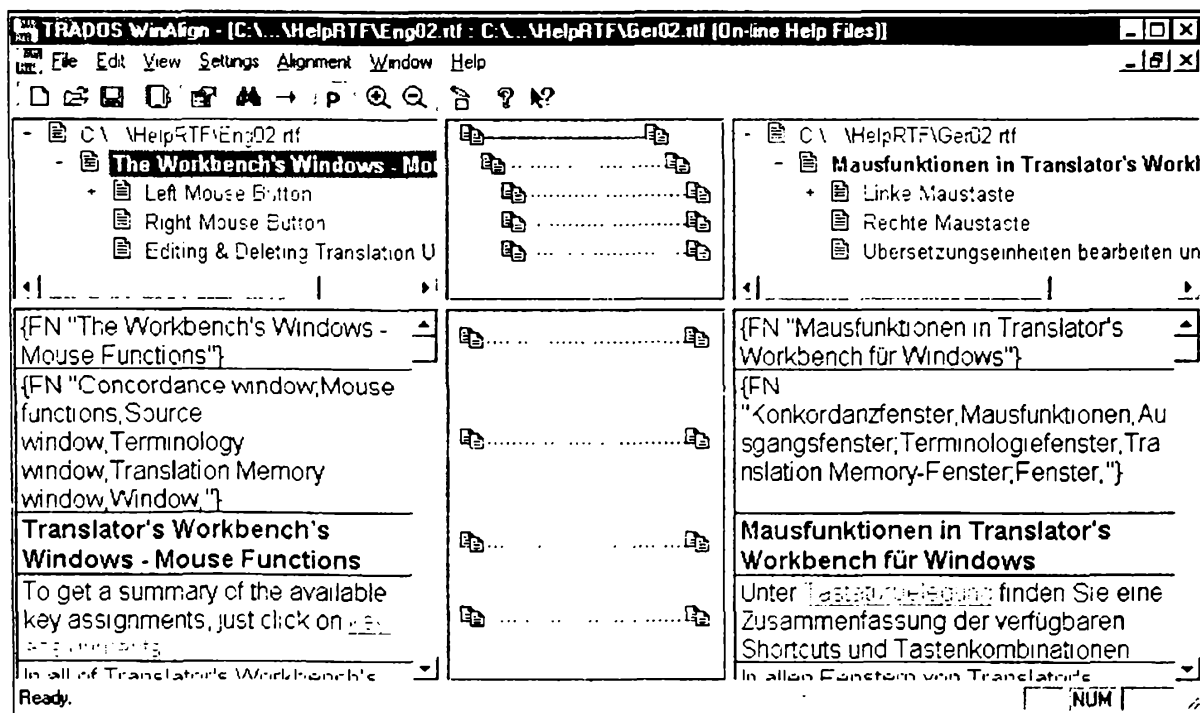


Рис. 40

Такой способ работы оправдан только в том случае, если пользователю предстоит сталкиваться с обновлением материалов исходного текста в последующей работе с ТМ-инструментом (например, при переводе какого-либо каталога, только более позднего выпуска). Часто случается, что корпус параллельных текстов хотя бы частично хранится на бумаге. В таком случае для адекватной работы программы слияния текстов неизбежно сканирование текстов и их тщательное форматирование. На это уходит время, и прежде чем приступить к этому, бюро переводов или переводчик должны быть уверены в том, что сформированная база данных непременно пригодится для перевода текстов похожего содержания. Практика показывает, что рентабельнее сделать новый перевод, начав с пустой базы билингв, чем создавать базу искусственным способом, сопоставляя старые оригиналы и переводы.

Программа TagEditor в составе системы TRADOS позволяет выполнять переводы текстов во многих форматах, отличных от форматов .doc (HTML, SGML, XML, ASP, JSP, TaggedRTF и TRADOSTag). Она осуществляет защиту тегов, т.е. они пропускаются в процессе работы ТМ-инструмента и автоматически переносятся в текст перевода. А это означает, что текст перевода может быть использован также, как и оригинал.

6.6. Некоторые вспомогательные функции

Concordance. Очень часто программа Translator's Workbench не находит удовлетворительного аналога достаточно длинного сегмента, в составе которого все же могут быть частично совпадающие фрагменты. Тогда целесообразно выделить этот фрагмент в поле перевода и воспользоваться функцией Concordance. Эта функция может работать самостоятельно, имеет свой интерфейс и открывается в отдельном окне.

Ввод термина в терминологическую базу в процессе перевода. Этот вспомогательный инструмент позволяет пополнять терми-

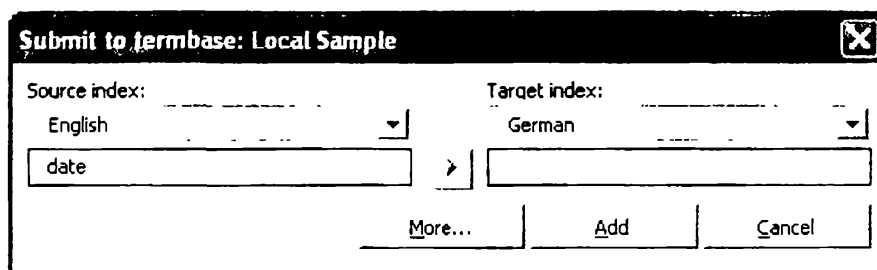


Рис. 41

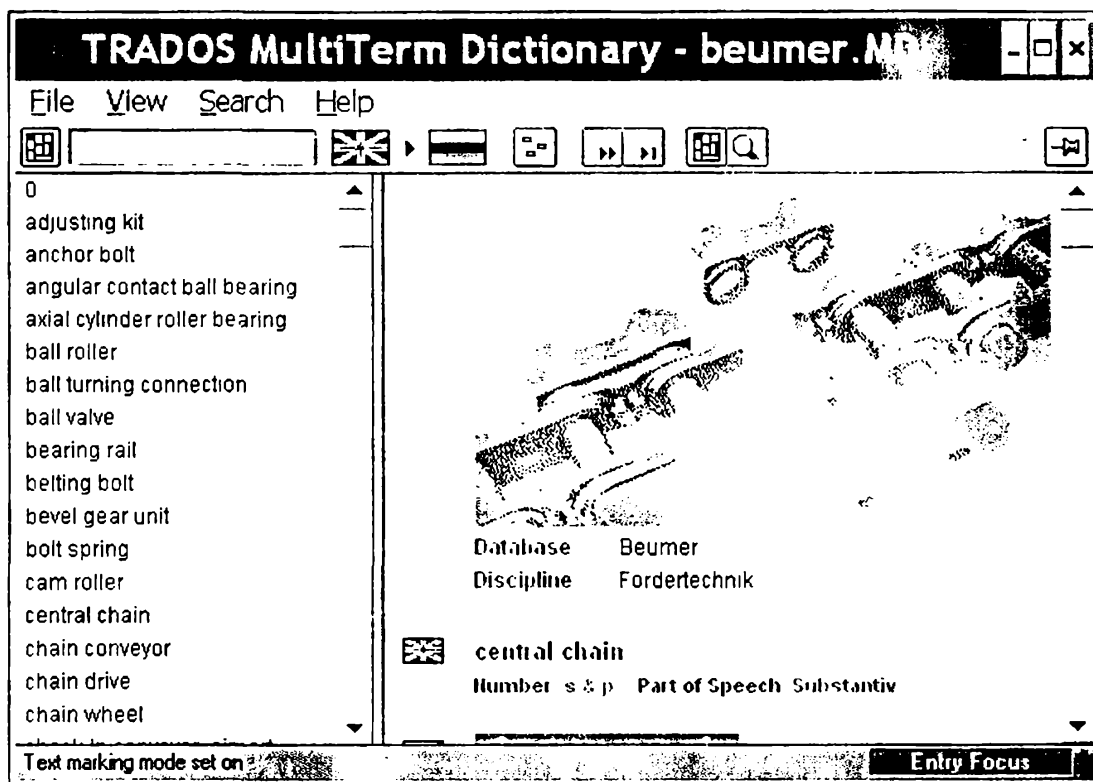


Рис. 42

нологическую базу программы MultiTerm в процессе перевода. Как это происходит, легко понять с помощью рис. 41.

Словари. Вместе с пакетом программ системы TRADOS пользователь получает несколько полезных словарей. В некоторых из них имеется и русская часть. Пример списка возможных словарей представлен на рис. 42.

6.7. Другие системы класса Translation Memory

В последние годы системы класса Translation Memory приобретают все большую известность среди людей, связанных с письменным переводом. Мировой рынок предлагает множество продуктов, которые уже приобрели популярность и широко используются в процессе автоматизации переводческой деятельности. В табл. 6 указаны наиболее известные из них.

Все эти системы обладают схожим набором инструментов и функций, однако по некоторым параметрам и ценам довольно сильно отличаются одна от другой. Это позволяет пользователю выбирать то, что необходимо ему, исходя из его профессиональных предпочтений. На сайтах фирм-производителей имеются очень подробные демонстрационные ролики и клипы.

При определении рентабельности внедрения в свою переводческую технологию такой системы следует уточнить, не требуется ли замена текстового редактора на более новый.

Компания	Название систем	Рабочие языки
IBM (США) www.4.ibm.com/ software/ad/ translat/lmt.	Personal Translator Personal Translator Plus PeTra (for OS/2) Personal Translator Translation Manager (Translation Memory)	Английский, французский, испанский, итальянский
TRADOS (США) www.trados.com	TRADOS Translation Solutions (Translation Memory)	Практически не зависит от языка
LANT (Франция) www.lant.com	EUROLANG OPTIMIZER (Translation Memory)	С шести языков (англ., нем., франц., исп., итал., голланд.) на более чем десять других языков (англ., нем., франц., исп., итал., голланд., дат., норв., фин., швед. и португ.)
STAR (Швейцария) www.star-group. neth	Transit (Translation Memory) TermStar WebTerm	Поддержка более сорока языков
ATRIL (США) www.atril.com	Deja Vu (Translation Memory)	Практически не зависит от языка

Контрольные вопросы и задания

1. Опишите принцип действия ТМ-инструментов.
2. В чем состоит дополнительное преимущество использования ТМ-инструментов в бюро переводов?
3. Каковы принципы терминологической функции ТМ-инструмента?
4. Подберите несколько однотипных текстов и попытайтесь определить, в какую из четырех ситуаций в технологии перевода с применением ТМ-инструмента могли бы попасть эти тексты.
5. Определите для себя рентабельность использования сервисной программы WinAlign.
6. Посетите web-сайты основных производителей ТМ-инструментов и попытайтесь сопоставить и оценить их возможности.

ЭРГОНОМИКА ТРУДА ПЕРЕВОДЧИКА**7.1. Общие положения**

Об охране и гигиене труда чаще говорят в связи со специальностями, где используются вредные для здоровья вещества, возможна радиация или опасна технология их обработки. Гигиене интеллектуального труда, разновидность которого представляет собой текстовая деятельность, и в частности перевод, крайне редко уделяют хотя бы какое-то внимание. Наша задача — подчеркнуть необходимость создания комфортных условий труда переводчика, что является достаточно плодотворным источником повышения его производительности и качества переводческой продукции, а потому обратить на это самое серьезное внимание своих читателей, будущих переводчиков.

Мы не беремся давать какие-либо рекомендации медицинского характера и указывать на противопоказания, которые, к сожалению, все же имеются. Но полагаем, что не лишним будет порекомендовать переводчику периодически встречаться со своим лечащим врачом специально для того, чтобы обсудить условия своего труда, возможности их изменения и свои наблюдения за реакцией организма на эти изменения.

В данном разделе нашего учебника мы хотим дать советы, продиктованные личным опытом организации своей собственной работы, а также опытом, накопленным многими высококвалифицированными профессиональными переводчиками. Наши советы в основном касаются двух параметров условий труда переводчика, которые органично взаимосвязаны, — это комфортность рабочего места и самочувствие человека.

В подавляющем большинстве случаев переводчик сам обустроивает свое рабочее место у себя дома или в офисе. Даже если он трудится в учреждении, службы охраны труда (к сожалению, имеющие тенденцию к сокращению во многих организациях) обходят переводчиков своим вниманием. Официальные нормативные документы в кодексе законов о труде уже давно не отражают реальное положение в сфере переводческой деятельности (хотя, пожалуй, и в любой другой). Вот почему главное, о чем мы хотели бы попросить задуматься переводчиков, заботящихся об условиях своего труда, — это самостоятельно и с самого начала своей трудовой деятельности правильно оценить ситуацию на своем рабочем месте и постараться сделать все возможное для того, чтобы

она была оптимальной. При этом следует исходить из того, что в конечном счете все мероприятия по организации рабочего места оказывают непосредственное влияние и на здоровье, и на благосостояние (через производительность).

До сих пор пока еще действуют Санитарные правила и нормы СанПин 2.2.2.542-96 «Гигиенические требования к видеодисплейным терминалам, персональным электронно-вычислительным машинам и организации работы», утвержденные постановлением № 14 Госкомсанэпиднадзора РФ от 14 июля 1996 года.

Ниже мы приводим отдельные из них.

Конструкция монитора должна обеспечивать возможность фронтального наблюдения экрана путем поворота корпуса в пределах $\pm 30^\circ$ в горизонтальной плоскости вокруг вертикальной оси и в вертикальной плоскости вокруг горизонтальной оси с фиксацией в заданном положении. Ручки регулировки яркости и контраста обеспечивают возможность изменения этих параметров от минимальных до максимальных значений. Для защиты от электромагнитных и электростатических полей допускается применение приэкранных фильтров, специальных экранов и других средств.

Клавиатура располагается на поверхности стола на расстоянии 100—300 мм от края, обращенного к пользователю или на специальной, регулируемой по высоте рабочей поверхности, отделенной от основной столешницы.

Помещения с персональными компьютерами должны иметь естественное и искусственное освещение. Естественное освещение осуществляется через светопроемы, ориентированные преимущественно на север и северо-восток, и обеспечивает коэффициент освещенности не ниже 1,2 % в регионах с устойчивым снежным покровом и не ниже 1,5 % на остальной территории. Указанные значения коэффициента нормируются для зданий, находящихся в III световом климатическом поясе.

Расположение рабочих мест в подвальных помещениях не допускается. В случаях производственной необходимости эксплуатация персональных компьютеров в помещениях без естественного освещения может проводиться только по согласованию с органами и учреждениями Госсанэпиднадзора.

Площадь на одно рабочее место с персональным компьютером составляет не менее 6,0 м², а объем не менее 20,0 м³.

Звукоизоляция ограждающих конструкций помещений с персональными компьютерами должна отвечать гигиеническим требованиям и обеспечивать нормируемые параметры шума.

Помещения с персональными компьютерами должны оборудоваться системами отопления, кондиционирования воздуха или эффективной приточно-вытяжной вентиляцией.

Для внутренней отделки интерьера таких помещений используются диффузно-отражающие материалы с коэффициентом от-

ражения для потолка — 0,7—0,8; для стен — 0,5—0,6; для пола — 0,3—0,5. Полимерные материалы для отделки интерьера должны быть разрешены для применения органами и учреждениями Госсанэпиднадзора. Поверхность пола должна быть ровной, без выбоин, нескользкой, удобной для очистки и влажной уборки, обладающей антистатическими свойствами.

В производственных помещениях, где работа на персональном компьютере является вспомогательной, температура, относительная влажность и скорость движения воздуха на рабочих местах должны соответствовать действующим санитарным нормам.

В производственных помещениях, в которых работа на персональном компьютере является основной, должны обеспечиваться оптимальные параметры микроклимата. Для повышения влажности воздуха в таких помещениях следует применять увлажнители воздуха, заправляемые ежедневно дистиллированной или прокипяченной питьевой водой.

Содержание вредных химических веществ в воздухе производственных помещений, в которых работа на персональном компьютере является вспомогательной, не должно превышать предельно допустимых концентраций вредных веществ в воздухе рабочей зоны. А там, где эта работа является основной, их содержание не должно превышать предельно допустимых концентраций загрязняющих веществ в атмосферном воздухе населенных мест.

Запрещается проводить ремонт вычислительной техники непосредственно в рабочих помещениях.

Переводчики относятся к категории специалистов, для которых уровень шума в рабочей зоне не должен превышать 60 дБА. Снизить его уровень в помещениях с персональными компьютерами можно использованием звукопоглощающих материалов с максимальными коэффициентами звукопоглощения в области частот 63—8000 Гц. Дополнительным звукопоглощением служат однотонные занавеси из плотной ткани, гармонирующие с окраской стен и подвешенные в складку на расстоянии 15—20 см от ограждения. Ширина занавеси должна быть в 2 раза больше ширины окна.

Искусственное освещение в помещениях эксплуатации персональных компьютеров осуществляется системой общего равномерного освещения. В производственных и административно-общественных помещениях при преимущественной работе с документами допускается применение комбинированного освещения (к общему освещению дополнительно устанавливаются светильники местного назначения, предназначенные для освещения зоны расположения документов).

Освещенность на поверхности стола в зоне размещения рабочего документа должна достигать 300—500 лк. Допускается установка дополнительных светильников для подсветки документов. Однако местное освещение не должно создавать бликов на поверх-

ности экрана и увеличивать его освещенность. Яркость светящихся поверхностей (окна, светильники и др.), находящихся в поле зрения работающего, не должна превышать 200 кд/м².

Необходимо ограничивать отражение поверхностей (экрана, стола, клавиатуры и др.) за счет правильного выбора типов светильников и расположения рабочих мест по отношению к источникам естественного и искусственного освещения, при этом яркость бликов на экране не должна превышать 40 кд/м², а яркость потолка при применении системы отраженного освещения быть не более 200 кд/м².

В качестве источников света при искусственном освещении применяются преимущественно люминесцентные лампы типа ЛБ. При устройстве отраженного освещения допускается использование металлогалогенных ламп мощностью до 250 Вт и ламп накаливания в светильниках местного освещения. Такие светильники должны быть снабжены непросвечивающим отражателем с защитным углом не менее 40°.

Для обеспечения нормируемых значений освещенности в помещениях, где используются персональные компьютеры, следует проводить чистку стекол оконных рам и светильников не реже двух раз в год и своевременно заменять перегоревшие лампы.

Рабочие места с персональными компьютерами располагаются по отношению к световым проемам таким образом, чтобы естественный свет падал сбоку, преимущественно слева. При этом необходимо учитывать расстояния между рабочими столами с мониторами (в направлении тыла поверхности одного монитора и экрана другого монитора), которое должно быть не менее 2,0 м, а расстояние между боковыми поверхностями мониторов — не менее 1,2 м.

Рабочие места с персональными компьютерами при выполнении творческой работы, требующей значительного умственного напряжения или высокой концентрации внимания, следует изолировать друг от друга перегородками высотой 1,5—2,0 м.

Конструкция рабочего стола должна учитывать возможность оптимального размещения на рабочей поверхности необходимого оборудования с учетом его количества и конструктивных особенностей. При этом допускается использование рабочих столов различных конструкций, отвечающих современным требованиям эргономики. Высота рабочей поверхности стола для взрослых пользователей должна регулироваться в пределах 680—800 мм; при отсутствии такой возможности ее высота должна составлять 725 мм. Должно быть предусмотрено пространство для ног высотой не менее 600 мм, шириной — не менее 500 мм, глубиной на уровне колен — не менее 450 мм и на уровне вытянутых ног — не менее 650 мм.

Конструкция рабочего стула (кресла) должна обеспечивать поддержание рациональной рабочей позы при работе на персональном компьютере, возможность изменения позы с целью сни-

жения статического напряжения мышц шейно-плечевой области и спины для предупреждения развития утомления. Рабочий стул (кресло) должен быть подъемно-поворотным, регулируемым по высоте и углам наклонов сиденья и спинки, а также по расстоянию спинки от переднего края сиденья; при этом регулировка каждого параметра должна быть независимой, легко осуществляемой и имеющей надежную фиксацию.

Поверхности сиденья, спинки и других элементов стула (кресла) должны быть полумягкими, с нескользящим, неэлектризующимся и воздухопроницаемым покрытием, обеспечивающим легкую очистку от загрязнений.

В его конструкции должны соблюдаться следующие параметры: ширина и глубина поверхности сиденья не менее 400 мм с закругленным передним краем; регулировка высоты сиденья в пределах 400—550 мм с углами наклона вперед до 15° и назад до 5°; высота опорной поверхности спинки — 300 ± 20 мм, ширина — не менее 380 мм с радиусом кривизны горизонтальной плоскости — 400 мм; угол наклона спинки в вертикальной плоскости в пределах $\pm 30^\circ$; регулировка расстояния спинки от переднего края сиденья в пределах 260—400 мм; стационарные или съемные подлокотники длиной не менее 250 мм и шириной 50—70 мм; регулировка подлокотников по высоте над сиденьем в пределах 230 ± 30 мм и внутреннее расстояние между ними 350—500 мм.

Рабочее место должно быть оборудовано подставкой для ног, имеющей ширину не менее 300 мм, глубину не менее 400 мм, с регулировкой по высоте до 150 мм и по углу наклона опорной поверхности подставки до 20°. Поверхность подставки должна быть рифленой и иметь по переднему краю бортик высотой 10 мм.

В помещениях с персональными компьютерами должна ежедневно проводиться влажная уборка. Необходимо, чтобы рабочие комнаты были оснащены аптечками первой помощи и углекислотными огнетушителями.

Таковы в основном требования официальных «Санитарных правил и норм».

7.2. Самоконтроль

Перевод фактически стал относиться к ненормируемым видам деятельности. Совершенно очевидно, что в любых видах устного перевода невозможно установить какие-либо нормы и в организационном плане речь может идти только о мере материального возмещения неестественных нагрузок на речемыслительный аппарат переводчика. В сфере письменных видов перевода тоже весьма затруднительно определить градацию срочности выполнения работы и универсальную ее тарификацию.

Опыт многих квалифицированных переводчиков предписывает соблюдение двух исходных количественных параметров труда письменного переводчика. Считается, что наиболее продуктивно переводчик может работать 6 часов в день со скоростью 500 слов в час. Это — средние параметры, и к ним нужно относиться лишь как к ориентирам для самоконтроля.

Очень часто возникает необходимость работать гораздо более 6 часов — по 10, а иногда и более часов на протяжении нескольких дней. Такое периодически возможно, и человек в состоянии позволить себе столь явную перегрузку, четко осознавая то, что она неизбежно вызовет переутомление и даже нервное перенапряжение. И, как показывает опыт, после напряженной работы обязательно наступает период, в течение которого переводчик может трудиться лишь с нагрузкой много ниже средней. Возникает естественная потребность в более или менее продолжительном отдыхе. Как правило, средняя сочетанная нагрузка этих двух периодов — повышенной интенсивности и спада — приводит все к тому же показателю — 6-часовому рабочему дню.

Можно ли увеличить производительность не обращаясь к новым технологиям? Без потери качества выполняемой работы сколь-нибудь значительное повышение скорости традиционного перевода без применения средств автоматизации и компьютеризации текстовой деятельности практически невозможно. Срочно выполненный перевод со скоростью более 500 слов в час может быстро и эффективно удовлетворить оперативные информационные потребности. Но после завершения этапа получения самой поверхностной информации неизбежно оказываются необходимыми различные доработки. В итоге суммарная скорость выполнения перевода, скорее всего, вернется к средней величине — 500 слов в час.

Не хотелось бы, чтобы приведенные выше доводы вселили хоть каплю пессимизма в душу начинающего переводчика и лишили его стремления работать как можно больше, лучше и быстрее. Еще раз хотим подчеркнуть, что в переводческой практике количество срочных и сверхсрочных работ значительно превышает количество плановых переводов. Наша главная цель в данном случае — убедить начинающего переводчика в том, что он имеет заслуженное моральное право позволить себе дополнительный отдых после срочной работы.

К сожалению, в нашем обществе довольно часто приходится сталкиваться с неправильным разделением труда на тяжелый и легкий. Зачастую работу переводчика относят к разряду легких. На самом деле переводческая деятельность требует высокой выносливости. Срочность письменных переводов, специфичность, а иногда и экстремальность устных работ, нарушают привычный режим труда и отдыха, питания и гигиены. В состав комплекса постоянных мер по совершенствованию профессионального мас-

терства, кроме языковой практики и всестороннего расширения кругозора, современный переводчик обязан включать эффективные меры по укреплению своего здоровья.

Выше мы постарались доказать необходимость и естественность периодов релаксации после длительных срочных работ. Но и в «штатном» режиме необходимо сочетать интенсивную работу с небольшими перерывами. Такой ритм полезен во всех профессиях, а для переводчика он особенно актуален в силу специфики его труда. Деятельность переводчика связана с постоянным переключением: с одного языка — на другой, с чтения — на запись, со слухового восприятия — на порождение речи и т. д. Естественно предположить, что организм переводчика более, чем организм представителей других профессий, требует циклических переключений напряжений и релаксаций.

При выполнении устного перевода во время международных мероприятий у переводчика почти нет возможностей уделить внимание самому себе, и здесь мало какие советы по повышению комфортности труда пойдут ему на пользу. Остается только научиться изыскивать и ловить любой момент и любую возможность для «передышки».

Другое дело в ситуации письменного перевода, когда переводчик имеет возможность установить свой режим производственных перерывов. Через каждые два часа интенсивной работы необходим 10—15-минутный перерыв. Основная цель такого перерыва — снять мышечное напряжение, связанное с сидячим положением на рабочем месте. Для этого полезны физические упражнения, уместные в окружающей среде и показанные по состоянию здоровья. Достаточное внимание следует уделять дыхательной гимнастике и проветрить помещение, если по другим условиям нельзя делать этого постоянно. Опыт показывает, что сухой воздух при температуре выше 22 °С резко снижает производительность. Если условия изменить нельзя, то во время перерыва полезно умыться холодной водой и прополоскать рот и нос. Такие процедуры полезны и для снятия статического электричества.

Каждый час рекомендуется делать 3—5-минутные перерывы. Во время таких перерывов лучше покинуть рабочее место, немного походить, распрямив грудную клетку, выпить воды, чая или кофе, сделать телефонный звонок или перебраться несколькими словами со свободным коллегой.

Между этими паузами, если позволяют обстоятельства, было бы неплохо сделать небольшой перерыв, не связанный с отвлечением от работы. Просто, сидя за рабочим столом, можно откинуться на спинку кресла, расправить плечи, выполнить несколько глубоких вдохов.

Для релаксации зрительного напряжения полезно регулярно делать небольшие «передышки» для глаз. Каждые 15—20 минут

нужно «оторвать» глаза от экрана и несколько раз перевести взгляд на отдаленные предметы — на расстоянии 5 — 6 м от вашего рабочего места. Полезно также раз десять учащенно поморгать — это особенно важно, если приходится работать в излишне пересушенном теплом воздухе.

Врач, безусловно, может дать дополнительные профессиональные рекомендации в соответствии с физическим состоянием обратившегося к нему человека. Наши советы, исходящие из опыта и здравого смысла, носят универсальный характер, доступны и полезны всем независимо от физического состояния и возраста.

7.3. Комфортность рабочего места

Как мы уже говорили, очень важно с самого начала трудовой деятельности критически и правильно оценить комфортность своего рабочего места. Конечно, очень большое значение имеют достаточная и обдуманно направленная освещенность, кондиционирование воздуха, качество отделочных материалов и многое другое, на что уже многократно указывалось в соответствующей литературе. Позволим себе остановиться лишь на специфических рекомендациях, обусловленных особенностями труда переводчика.

Сейчас не найдется переводчиков любого возраста и любой квалификации, которые не согласятся с тем, что центром современного рабочего места является компьютер с многочисленной и очень разнообразной периферией.

Собственно компьютер начинается с системного блока с процессором, оперативной памятью, диском для хранения данных и портами для ввода и вывода данных. Специалисты по компьютерной технике обязательно в качестве совета объявят вам, что частота процессора должна быть как можно выше, оперативная память как можно больше и, конечно, нужно обзавестись, по меньшей мере, двумя жесткими дисками, скажем, по 300 Гб. Да, это неплохие советы! К сожалению, они, наверняка, ничего не скажут о том, что все это будет изнурительно гудеть около вашего уха не менее 6 ч в сутки, причем именно в то время, когда вам необходимо сосредоточенно работать одновременно на двух языках, экстремально нагружая свою нервную систему.

Даже при условии использования самого продвинутого программного обеспечения для организации современного рабочего места переводчика вполне подойдет системный блок средней мощности из разряда «компьютер для дома». Бесшумность его работы и эргономичность корпуса гораздо важнее избыточной мощности. Основным источником шума является система охлаждения процессора. При небольшой, но достаточной для вас мощности можно организовать почти бесшумное охлаждение. Материальная экономия на

мощности позволит вам использовать корпус с дополнительными средствами шумоподавления. Обратите внимание на то, чтобы конструкция корпуса обеспечивала вывод разъемов различных портов и звуковой карты на лицевую или верхнюю панель.

Очень важно правильно выбрать в рабочей зоне место установки системного блока. Многие делают ошибку, располагая системный блок на столе. Большинство современных рабочих столов снабжено специальным отделением для системного блока. К сожалению, это серьезная ошибка конструкторов офисной мебели. Сам рабочий стол, а тем более такое специальное отделение оказываются очень эффективными (в самом отрицательном смысле) резонаторами, многократно усиливающими шум системного блока. Переместите системный блок из отсека стола на толстый резиновый коврик рядом, и вы сразу достигните комфортного эффекта.

Оптические дисководы, сканеры и принтеры тоже шумят, но они не требуются постоянно, и их шумовыми характеристиками можно пренебречь.

Оптические дисководы обычно размещаются в верхней части корпуса системного блока. Даже если системный блок стоит почти на полу рядом с рабочим столом, менять диски будет удобно. Расположение принтера не играет существенной роли в организации рабочего места. Если для получения бумажной копии перевода придется встать и сделать несколько шагов, это пойдет вам только на пользу.

Сканер необходимо разместить таким образом, чтобы, удобно сидя в рабочем кресле, можно было работать с ним обеими руками. У большинства переводчиков он расположен слева на рабочем столе.

Большую роль в создании комфорта на рабочем месте играет монитор. На него приходится смотреть весь рабочий день и наблюдать по нему за всем, что вы делаете, читать электронные письма, выбирать информацию из Интернета.

Прежде всего следует решить, какой величины экран будет наиболее удобен для вас в условиях вашего рабочего места. Величину экрана измеряют по диагонали в дюймах. Наиболее распространенные размеры: 15, 17, 19, 20 и 21 дюйм. Большинство мониторов выполнены с соотношением сторон экрана 4:3, но уже достаточно много с соотношением 16:9. Последние имеют панорамный вид и очень удобны для работы не только с документами, но и в Интернете и для просмотра фильмов.

Как определить размер экрана для вашего рабочего места? Если вы удобно сидите в рабочем положении в своем кресле, монитор должен располагаться прямо перед вами на расстоянии приблизительно 80—90 см от глаз. Весь экран должен находиться в поле вашего зрения, но его верхняя кромка — выше линии вашего взора на 6—7 см. Таким требованиям, скорее всего, удовлетворит

монитор с экраном 19 дюймов. При большем экране вы получите слишком увеличенное изображение, способное ввести вас в заблуждение при макетировании текста. Учтите, что на экране 19 дюймов текст выглядит примерно вдвое крупнее, чем на самом деле, экран размером 21 дюйм дает еще большее увеличение. К этому, конечно, можно быстро привыкнуть, но целесообразность использования такого большого экрана для текстовой деятельности весьма сомнительна. Если же монитор недостаточно велик и вам хотелось бы поставить его ближе, мы все же не советуем делать этого. Гораздо лучше, например, увеличить на экране размер шрифта и иконок интерфейсов используемых программ.

При подборе монитора надо учитывать два параметра: качество изображения и разрешающую способность. Но главным все же является комфортность вашего восприятия. Перед приобретением монитора постарайтесь достаточно продолжительно посмотреть, как работают различные популярные модели. Независимо от характеристик смело берите тот, восприятие которого будет вам приятнее. Кроме этого простого приема проведите еще одно специальное испытание — попробуйте посмотреть динамичную сцену из какого-либо игрового фильма. Если, например, за резкими перемещениями руки на экране не тянется «шлейф», если краски остаются яркими и контрастными, если вы нигде не видите статичных темных точек, вы подошли к правильному выбору.

Многие современные мониторы имеют очень удобное дополнительное оборудование — стереофонические аудиокolonки, разъемы USB и разъемы звуковой карты. Хотя все мониторы снабжены удобными подставками, которые позволяют регулировать не только высоту, но и наклон экрана, все же мы рекомендуем размещать экран строго вертикально. Современный экран обязательно имеет антибликовую поверхность, тем не менее мы советуем расположить рабочее место так, чтобы экран все же занимал перпендикулярное положение относительно окна помещения.

Вокруг монитора должно быть достаточно света. Вообще требования к освещению рабочего места весьма индивидуальны. Вы сами должны выбрать себе комфортный уровень освещенности рабочего места. Заметим только, что даже самый качественный и изначально контрастно приятный экран на слишком затемненном рабочем месте создает повышенную нагрузку на зрение и способствует накоплению общей усталости.

Перевод — это прежде всего умственный труд. Но в письменном переводе имеется еще достаточно большая доля физического труда. При средней нагрузке переводчик за рабочий день не менее 20 тыс. раз ударяет по клавишам клавиатуры. Этот факт заставит вас уделить внимание выбору клавиатуры. Современный рынок предлагает их около 2 тыс. моделей нескольких типов. Обычно классические клавиатуры имеют 102 клавиши, но, вообще, их число

может быть от 17 (у специальных цифровых клавиатур, которые не представляют интереса для переводчиков) до 153 с расширенными мультимедийными возможностями и со специальными клавишами для удобной навигации в Интернет-браузерах.

Тип клавиатуры определяется механизмом работы клавиш. Они бывают мембранными, механическими или полумеханическими. Мы не будем здесь рассматривать принцип действия клавиш. Обратим лишь внимание на то, что наиболее распространены клавиатуры мембранного типа — в основном из-за их низкой цены. Их долговечность составляет 20 млн срабатываний. Вообще, это цифра внушительная. Но если учесть, что переводчик делает не менее 20 тыс. ударов в день, то подобная долговечность оказывается не такой уж большой. Следует учесть еще и то, что обычно пальцы привыкают к клавиатуре, и частая ее смена будет снижать рентабельность вашего труда. Но у мембранных клавиатур имеется одно существенное преимущество — они не издают характерных шелкающих звуков. Если даже такие шелчки вам не мешают, они могут сбивать с рабочего ритма ваших коллег, которые трудятся рядом.

Более надежны механические клавиатуры, рассчитанные на 100 млн нажатий, но они «шумят» и стоят значительно дороже мембранных. Но вот полумеханические клавиатуры совмещают плавность работы мембранных и довольно высокую надежность (но ниже, чем у механических).

При активном использовании клавиатуры, что обычно для работы переводчика, большое значение приобретает длина рабочего хода клавиши. К сожалению, у каждой клавиатуры своя длина хода и колеблется она приблизительно от 2 до 4,5 мм. Это — весьма важная характеристика клавиатуры. Со временем у каждого человека вырабатывается «память физического действия». Пальцы привыкают совершать постоянно повторяющиеся действия. Постарайтесь вспомнить, как часто вы пальцами чувствуете, что, по-видимому, сделали опечатку. Почему это происходит? Да потому, что ваши пальцы совершили какую-то непривычную комбинацию движений, и вы сразу заметили это. Вполне понятно, что за продолжительный период пальцы привыкают к длине хода клавиш на освоенной клавиатуре; ее вынужденная замена и переход с одной клавиатуры на другую почти всегда будут для вас не очень приятны.

Переводчики старшего поколения, имеющие огромный опыт работы на пишущих машинках, чаще всего предпочитают клавиатуры с большим рабочим ходом клавиш. Если вы хорошо владеете клавиатурой и печатаете текст достаточно бегло, имеет смысл приучить свои руки к мягким клавиатурам с небольшим (до 3 мм) ходом клавиш. Но в любом случае целесообразно выбирать надежные клавиатуры, чтобы менять их как можно реже.

Клавиатура — устройство асимметричное. Она совмещает в себе два блока — алфавитный и цифровой. Для переводчика основным является алфавитный блок, середина которого находится приблизительно в зоне клавиши «b/и». Эргономичным считается размещение центра клавиатуры на оси, проходящей от центра вашего рабочего кресла к центру монитора.

Необходимо расположить клавиатуру на такой высоте, чтобы при работе ваши предплечья были в строго горизонтальном положении, кисти рук продолжали положение предплечий, а пальцы свободно касались клавиш, не выгибая кисти вверх. Такое положение способствует минимальному напряжению и снижает усталость локтевого и лучезапястного суставов. Большинство современных рабочих столов оборудованы специальной выдвижной полкой для клавиатуры. Чаще всего ее высота соответствует удобным для пользователя требованиям, но иногда необходима некоторая подгонка. Дорогие хорошо оборудованные столы имеют несложные устройства для регулирования высоты.

Использование выдвижной полки иногда создает проблему освещения клавиатуры. Даже если вы владеете «слепым» методом набора текста, клавиатура должна быть хорошо освещена. Сейчас рынок предлагает несколько клавиатур с подсветкой клавиш. На первый взгляд это очень удобно, и многим такая клавиатура может понравиться. Но такая подсветка осуществляется с помощью светодиодов — не совсем привычных источников света. Кроме того, у вас наверняка уже имеется какое-либо дополнительное освещение рабочего стола. Следует учитывать, что постоянное переключение с одного вида освещения на другое не очень полезно, и отнюдь не относится к числу эффективных изменений.

Неотъемлемым элементом рабочего места переводчика является компьютерная «мышь». По принципу действия «мыши» делятся на оптомеханические и оптические. В оптомеханическом варианте установлен перекатывающийся резиновый шарик, приводящий в движение внутренний механизм, конвертирующий движения в дискретные сигналы, управляющие курсором на экране. Шарик быстро пачкается, в результате весь механизм начинает «залипать». Необходимы частые чистки, которые резко снижают долговечность и без того не очень надежного механизма. И хотя такие «мыши» чрезвычайно дешевы, спрос на них падает и объемы их производства сокращаются. Вполне возможно, что в ближайшее время они полностью будут заменены двумя типами оптических «мышей» — лазерными или светодиодными. Принципиальные различия в их устройстве вряд ли заинтересуют переводчиков. Важно, что в них нет движущихся частей, а это гарантирует их сверхвысокую надежность. Некоторые опасаются лазерных «мышей» из-за возможного опасного воздействия на организм лазерного излучения. В инструкциях по эксплуатации гарантируется

полная безопасность этого маломощного излучения. И при использовании такой «мыши» исключительно по назначению ее маломощный лазерный луч никак не может даже кратковременно попасть ни на ваши глаза, ни на какие-либо другие участки тела.

На современной «мыши» имеется от 1 (фирма «Apple») до 12 клавиш, часть которых может быть перепрограммирована на специальные операции, которые переводчику вряд ли потребуются. Самые распространенные «мыши» имеют две клавиши и колесо прокрутки между ними. У некоторых моделей вместо колеса прокрутки встроен манипулятор типа «джойстик». Это очень удобно для переводчика. При необходимости часто перемещаться по большому тексту приходится постоянно вращать колесо прокрутки, что создает повышенную нагрузку на указательный палец. Когда «мышь» снабжена манипулятором «джойстик», прокрутка происходит при нажатии в нужном направлении и удержании его. Но следует заметить, что найти такую «мышь» довольно трудно.

Удобнее всего размещать «мышь» на поверхности рабочего стола. При обычной высоте стола (70—72 см) вы можете пользоваться ей почти вытянутой рукой с легким сгибом в локтевом суставе. Лучезапястный сустав при этом будет опираться о поверхность стола, а кисть свободно управлять ее перемещениями.

* * *

Современный рынок компьютерной техники и офисного оборудования предлагает очень много продукции с пометкой «эргономичный». Нередко это действительно весьма эффективные и удобные вещи, но, к сожалению, еще чаще предметы из разряда эргономичных не выдерживают никакой критики при длительном использовании. Все оценки меры эргономичности не могут не быть субъективными. Достаточно сказать, что нам не приходилось видеть ни одной клавиатуры из разряда эргономичных в мощных офисах по подготовке текстовых данных, где работают профессионалы и где организация производства почти идеальна.

Оценка подлинной эргономичности — это исключительно ваша задача и ее мерой является ваше вдумчивое и внимательное отношение к собственной индивидуальности.

Контрольные вопросы и задания

1. Попробуйте спроектировать рабочее место переводчика в соответствии с санитарными нормами и правилами.
2. Попробуйте критически оценить свое собственное рабочее место. Соответствует ли оно результатам, которые вы получили, выполнив предыдущее задание?
3. Каковы требования к освещению рабочего места переводчика?

4. Укажите требования к физическим параметрам стола и кресла.
5. Назовите пороги снижения производительности труда переводчика в зависимости от продолжительности непрерывной работы, кондиционирования воздуха и т. п.
6. Каковы средние требования к основным аппаратным средствам рабочего места переводчика: к системному блоку, монитору, клавиатуре и др.?
7. Попробуйте сформулировать свое отношение к понятию эргономичности оборудования для переводчика.

ЛИТЕРАТУРА

- Акопян А. В. Программы профессионального перевода. — М., 2003.
- Бондаренко С., Бондаренко М. 100 лучших программ для офиса. — СПб., 2005.
- Карр Н. Дж. Блеск и нищета информационных технологий. — М., 2005.
- Марчук Ю. Н. Проблемы машинного перевода. — М., 1983.
- Нелюбин Л. Л. Компьютерная лингвистика и машинный перевод. — М., 1991.
- Убин И. И. ЭВМ и словарь (методическое пособие). — М., 1992.
- Хроменков П. Н. Анализ и оценка эффективности современных систем машинного перевода: автореф. дис. ... канд. филол. наук. — М., 2000.
- Dunne K. J. (ed.). Perspectives on Localization. — Am-dam: Benjamins, 2006.
- Quah C. K. Translation and Technology. — Hampshire: Palgrave Macmillan, 2006.
- Somers H. (ed.) Computers and Translation: A translator's guide. — Am-dam: Benjamins, 2003.

ПРЕДМЕТНЫЙ УКАЗАТЕЛЬ

Автоматический переводной словарь
11, 98, 106
— переводчик 18
— словарь 53, 105
адекватность 17, 32, 34, 62, 63, 81
алгоритм 6, 8, 56, 57, 65, 66, 69, 71,
73, 83, 123, 148, 152, 153, 155, 157,
158
— перевода 6
анализ 19, 56
аналитико-синтетическая (ий) обра-
ботка 16, 17, 26, 34
аналитико-синтетический процесс 19
аннотирование 16
аппроксимация 23, 25, 34
База данных 199
— переводческая 191
библиографическое описание 16
билингва 12, 187, 189, 200
билингвальная (ый) компетенция
22—25
— контакт 23, 24
— посредник 24
Вариантное соответствие 56
вторичная информация 16, 17, 26
— коммуникативная деятельность 134
входной язык 56
выходной язык 56
Глубинная структура 21
грамматика предложения 20
Декодирование текста 19
денотатная структура 62
Джорджаунский эксперимент 5
дискурсный анализ 25
дружелюбный интерфейс 183
Жанр 24
— письменного перевода 17
Иммунитетная жесткость 110, 111
интерлингва 58
интерпретация 34

интерредактирование 63
информатика 79
информационно-справочная база 15
искусственный интеллект 52, 54, 59
исследовательская система 54
Когерентный дискурс 34
кодирование 16
коммерческая система 54, 55
коммуникативная (ый) деятельность
16
— ситуация 20
— текстовая деятельность 29
— эквивалентность 69
— дискурс 18
компрессия текстов 53
компьютеризация 14
— переводческой деятельности 14
— текстовой деятельности 36, 38, 42
конвертирование 15
конечный пользователь 66
контекстологические связи 34
контрастивная текстология 22, 32—34
критерий адекватности 64
— понятности 64
Лексический синтез 72
лемматизация 152
лингвистический ингредиент 36, 37,
38
лингвистический (ое) процессор 54
— обеспечение 55
локализация 15, 38, 43, 44, 45, 46
Многозначность 40, 42
многоуровневая иерархическая струк-
тура 21
многоязычие 141
— словаря 140
модель перевода 19
морфологический анализ 55, 71
— синтез 55, 72
Направление перевода 53
Обратимость 146

обратимый переводной словарь 142, 144, 145

Память переводов 12, 15

— переводчика 188

— предложений 187

парадигматические связи 31

первичная информация 16, 17

— коммуникативная деятельность 16, 134

перевод-подстрочник 55

переводимость 42, 94

переводное (ой) соответствие 56—58, 68, 69

— эквивалент 9, 11, 55, 124

переводческие инструменты 38

перекодирование 19

перефразирование 42

перцептивный потенциал оригинала 35

поисковая система 47, 49

пользовательский словарь 39

постредактирование 42, 55

прагматический компонент 24, 31, 32

— трансфер 17

предредактирование 42, 55

программное обеспечение 14, 67

промышленная система 55

психолингвистический трансфер 23

Рабочая (ий) среда 14

— язык 53

режим реального времени 52

референционная семантика 22

Сегмент 199, 204

семантическая (ие, ий) эквивалентность 34, 142

— комплексы 21

— анализ 55

семиотическое перекодирование 23

— членение 29, 119

синтагматические связи 30, 31

синтаксема 21

синтаксическая (ые, ий) конструкция 12

— комплексы 21

— анализ 55, 71

— синтез 72

синтез 19, 21, 56

система машинного перевода 40

— прямого типа 54

— семантического типа 54

Ситуационная модель перевода 20

ситуационно-аналогичные тексты 20

словарь сочетаемости 143

словник 122, 142

содержательный перевод 25

специализированный словарь 81

среда функционирования текста 27, 32

стилистическая эквивалентность 62

Текстовая (ый) деятельность 14

— процессор 54

текстоориентированный словарь 104

терминологическая (ий, ос) база 205

— номинация 96, 97

— банк 135, 139

— обеспечение 39, 166, 167

терминосистема 102, 103

технологический подход 13, 36

ТМ-инструмент 38, 43, 44, 189

трансфер 19, 54, 56

трансферная (ые) система 54

— преобразования 72

трансформационное (ый) соответствие 74, 75

— анализ 21

Устойчивые словосочетания 100

Формальное соответствие 21

форматирование 15

фрагментированный перевод 15

функциональная равнозначность 43

Эквивалентное соответствие

56, 69, 75

эквивалентность 17, 20, 21

экспериментальная система 54, 55

экстралингвистические (ий) знания

58, 59, 64, 127, 164

— критерий 66

Язык-посредник 58

ОГЛАВЛЕНИЕ

Предисловие	3
-------------------	---

ЧАСТЬ I

ТЕОРЕТИЧЕСКИЕ ОСНОВЫ ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ В ПЕРЕВОДЕ

Глава 1. Истоки новых информационных технологий в переводе	5
1.1. Первые опыты машинного перевода.....	5
1.2. Автоматические переводные словари	9
1.3. ТМ-инструменты (Translation Memory)	12
1.4. Информационные технологии и бизнес	12
Глава 2. Лингвистические основы информационных технологий в переводе	16
2.1. Перевод как аналитико-синтетический процесс	16
2.2. Модели процесса перевода	18
2.3. Основные теоретические модели	19
2.4. Типы переводов	24
2.5. Коммуникативные особенности аналитико-синтетического процесса	26
2.6. Лингвистическое понимание среды и системы в коммуникативном процессе	27
2.7. Контрастивная суть аналитико-синтетического процесса	32
Глава 3. Область практической реализации информационных технологий	36
3.1. Лингвистический ингредиент	36
3.2. Терминологичность информации	39
3.3. Системы машинного перевода	40
3.4. Основные принципы локализации при переводе	43
3.5. Рентабельность технологии перевода	45

ЧАСТЬ II

ПРАКТИЧЕСКИЕ ОСНОВЫ ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ В ПЕРЕВОДЕ

Глава 4. Системы машинного перевода	48
4.1. Системы машинного перевода и Интернет	48
4.2. Система машинного перевода как лингвистический процессор ...	53
4.3. Современные требования к системам машинного перевода	61

4.4. Типология ошибок при работе систем машинного перевода	67
4.5. Оценка эффективности основных действующих систем машинного перевода	76
Глава 5. Автоматические переводные словари	95
5.1. Принципы номинативной деятельности	95
5.2. Формирование словаря	100
5.3. Становление компьютерной лексикографии	105
5.4. Основные признаки системности терминологии	109
5.5. Внешние и внутренние признаки типологизации терминологии	121
5.6. Условия соотнесенности терминов и понятий	128
5.7. Вариативность терминосистемы	129
5.8. Коммуникативные условия функционирования терминосистемы	138
5.9. Основные лингвистические характеристики функционирования терминосистем	158
5.10. Инструментальность термина в составе коммуникативной деятельности	162
5.11. Терминологические проблемы перевода	163
5.12. Автоматические переводные словари	168
5.13. Место переводного словаря в процессе перевода	181
5.14. Практическое использование автоматического переводного словаря	182
Глава 6. Память переводов — Translation Memory	187
6.1. Истоки ТМ-инструментов	187
6.2. Идеология ТМ-инструментов	187
6.3. Практическое использование системы TRADOS 5 Freelance	194
6.4. Общий принцип Translation Memory	199
6.5. Сервисные программы	203
6.6. Некоторые вспомогательные функции	204
6.7. Другие системы класса Translation Memory	205
Глава 7. Эргономика труда переводчика	207
7.1. Общие положения	207
7.2. Самоконтроль	211
7.3. Комфортность рабочего места	214
Литература	220
Предметный указатель	221