

ФЕДЕРАЛЬНОЕ АГЕНТСТВО ПО ОБРАЗОВАНИЮ
АЛТАЙСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
Факультет журналистики

МАШИННЫЙ ПЕРЕВОД
МЕТОДИЧЕСКИЕ УКАЗАНИЯ



Барнаул

Издательство Алтайского
государственного университета
2009

Составители:

канд. филол. наук, доцент *О.С. Саланина*

ст. преподаватель *М.С. Наумова*

Рецензент:

канд. филол. наук, доцент *И.Н. Шадрина*

Методические указания содержат краткий обзор базовых понятий общей теории перевода, а также истории, современного состояния и новейших достижений в области машинного перевода.

Предназначены для магистров, студентов дневного и вечернего отделений факультета журналистики.

План УМД 2009 г., п. 111

Подписано в печать 15.09.2009. Формат 60x84/16.

Усл. печ. л. 1,0. Тираж 100 экз. Заказ 310.

Типография Алтайского государственного университета:

656049, Барнаул, ул. Димитрова, 66

Пояснительная записка

В последнее время знание иностранных языков требуется повсеместно – в учебе и быту, при общении с иностранцами, чтении Web-страниц иноязычных сайтов и книг на иностранном языке. В этом случае требуются услуги переводчика. Однако процесс перевода можно провести самостоятельно и оптимизировать с помощью средств вычислительной техники. Основные знания в области машинного перевода, необходимые умения и опыт использования соответствующих программных средств и технологий студенты приобретут в процессе освоения данного курса.

Программа курса «Машинный перевод» создана для расширения спектра образовательных услуг для студентов факультета журналистики и рассчитана на студентов, имеющих базовую подготовку по информатике и иностранному языку. Количество часов, предусмотренное в базовом курсе на овладение информационными технологиями, предполагает первичное ознакомление с программами и принципами обработки данных, таким образом данный курс направлен на изучение основ машинного перевода, его средств, принципов их работы, возможностей и назначения, а также на выработку устойчивых навыков использования компьютерных технологий при переводе иноязычных текстов. Курс содержит обзор базовых понятий общей теории перевода, которые помогут студентам осуществить качественное редактирование текстов, переведённых с помощью программ машинного перевода.

Преподавание данного курса базируется на неразрывной связи между теорией и практикой, поэтому рекомендуется проводить уроки, сочетающие изучение нового материала с практической реализацией полученных знаний за компьютером.

Рекомендуемое аппаратное и программное обеспечение – IBM совместимая компьютерная техника, сканер, принтер, мультимедийный проектор, электронные словари (Lingvo, Контекст 3.0, «МультиЛекс 1.0, 2.0, 3.0», Polyglossum и др.), системы машинного перевода (Сократ, Prompt и др.), подключение к Internet, броузеры, телеконференции, информационные службы, поисковые системы.

Программа курса

Введение (4 ч.)

Перевод, его виды, коммуникативная эквивалентность текстов. Критерий точности перевода, правильный и адекватный переводы. Возможности средств и методов информатики для перевода иноязычных текстов. Достоинства, недостатки, применимость и ограничения машинного перевода. Роль и место машинного перевода в деятельности специалистов гуманитарного профиля. Особенности английского газетно-информационного стиля.

Тема 1. История машинного перевода (2 ч.)

Возникновение концепции машинного перевода (40-е гг.). Лингвистические проблемы машинного перевода (50–60-е гг.). Развитие вычислительной техники и машинный перевод (70–80-е гг.). Машинный перевод сегодня (конец 90-х гг. – начало XXI в.). Перспективы машинного перевода.

Тема 2. Общая характеристика средств машинного перевода (2 ч.)

Электронные словари и системы машинного перевода как основные средства компьютерного перевода иноязычных текстов. Понятие «системные требования», возможности установки электронных словарей и систем машинного перевода в соответствии с их системными требованиями. Особенности инсталляции электронных словарей и систем машинного перевода, установка в полной и неполной компоновке. Лингвистическое, математико-алгоритмическое, программное, информационное и логическое обеспечение систем машинного перевода.

Тема 3. Электронные словари (2 ч.)

Электронные словари: классификация, особенности интерфейса и компонентного состава. Нерезидентные, резидентные, многооконные и многофункциональные системы электронных словарей. Сравнительный анализ существующих электронных словарей в соответствии с функциональным назначением (объем вокабуляра, специализированный словарь, система поиска; пользовательский словарь, интерфейс). Решение задач по переводу иноязычных текстов на основе использования электронных словарей.

Тема 4. Системы машинного перевода (4 ч.)

Алгоритмы анализа и синтеза систем машинного перевода. Диалоговый, пакетный и фоновый режимы перевода. Особенности интерфейса и компонентного состава систем машинного перевода. Возможности настроек систем машинного перевода. Семантико-синтаксический перевод. Система ЭТАП. Фразеологический перевод. Система RETRANS. Взаимодействие систем машинного перевода и приложений Microsoft Office, а также систем различных производителей. Мультимедийные особенности систем машинного перевода. Сетевые возможности систем машинного перевода. Машинный перевод в Интернете. Возможности систем машинного перевода по статейному переводу и сравнительная характеристика систем машинного перевода различных производителей. Системы машинного перевода как часть систем искусственного интеллекта. Составление общего алгоритма перевода иноязычных текстов и реализация его с помощью систем машинного перевода. Проведение сравнительного анализа областей применения электронных словарей и систем машинного перевода посредством решения задач по переводу иноязычных текстов.

Заключение (2 ч.)

Итоговые выводы о целесообразности, сферах применения машинного перевода и улучшения качества выходного текста.

По окончании курса *студент должен знать:*

- основные способы сочетаемости лексических единиц и основные словообразовательные модели;
- знать формы деловой переписки, иметь представление о форме договоров, контрактов, патента;
- нормы официально-деловой письменной речи, международные и национальные стандарты видов и разновидностей служебных документов;
- характерные способы и приемы отбора языкового материала в соответствии с различными видами речевого общения и перевода;
- основную иноязычную терминологию специальности, русские эквиваленты основных слов и выражений профессиональной речи.

студент должен уметь:

- выполнять устный, письменный и машинный переводы текстов, участвуя в редакционно-издательском процессе;
- переводить и редактировать текст, ориентированный на ту или иную форму речевого общения;
- работать с оригинальной литературой по специальности.

студент должен иметь навыки:

- работы со словарем (читать транскрипцию, различать прямое и переносное значение слов, находить перевод фразеологических единиц);
- реферирования и аннотирования литературы, опираясь на результаты машинного перевода текстов.

Введение

I. Слово «перевод» имеет два значения. В одном значении «перевод» – это продукт деятельности переводчика, т.е. текст, созданный им в устной или письменной форме. В другом значении слово «перевод» означает процесс создания этого продукта, т.е. деятельность переводчика, создающего текст.

II. «Переводной текст является производным от исходного текста. Он создается на основе определенных характеристик последнего и воспроизводит их средствами иного языка в условиях другой культуры (Л.К. Латышев)». Другими словами, созданный текст в идеале должен выступать в качестве полноценной коммуникативной замены оригинала и отождествляться с ним в функциональном, структурном и содержательном отношении.

Функциональное отождествление оригинала и перевода заключается в том, что перевод как бы приписывается автору оригинала, публикуется под его именем, обсуждается, цитируется и пр. так, как будто он и есть оригинал, только на другом языке.

Содержательное отождествление оригинала и перевода заключается в том, что перевод полностью воспроизводит содержание оригинала, что в нем передается то же содержание средствами иного языка.

Структурное отождествление перевода с оригиналом заключается в том, что перевод воспроизводит оригинал не только в целом, но и в частностях. Предполагается, что переводчик точно передает структуру и порядок изложения содержания в оригинале, не позволяет себе что-либо изменить, исключить или добавить от себя. Количество и содержание разделов и других подразделений текста в оригинале и переводе должно совпадать. Если в оригинале какая-то мысль высказана в начале второго раздела, то и в переводе она должна быть обнаружена на том же месте и т.д. Если переводчик и позволяет себе какие-то отступления в отношении частных деталей структуры текста, то лишь для того, чтобы точнее передать содержание оригинала.

III. Переводной текст оценивается в свете теории перевода с точки зрения категории соответствия, которая представлена такими понятиями как «эквивалентность» и «адекватность». Разграничение данных понятий носит принципиальный характер, так как они обуславливают установление критериев точности перевода, который трактуется не как «простое механическое воспроизведение всей совокупности элементов подлинника, а сложный сознательный отбор различных возможностей их передачи» (А.В. Федоров). Обобщая мнения исследователей по поводу природы эквивалентности и адекватности перевода, можно сформулировать в общем виде их основные критерии при разграничении данных понятий.

Эквивалентность в известной мере нацелена на идеальный эталон, далеко не всегда достижимый (хотя, в принципе, возможный). Она подразумевает исчерпывающую передачу содержания оригинала на всех семиотических уровнях и в полном объеме его функций (полная эквивалентность), или, по крайней мере – применительно к тому или иному семиотическому уровню или той или иной функции (частичная эквивалентность). Однако стоит отметить, что случаи полной эквивалентности наблюдаются, как правило, в относительно несложных коммуникативных условиях, в текстах со сравнительно узким диапазоном функциональных характеристик.

Адекватность же представляет собой категорию с иным онтологическим статусом. Она нацелена на лучшее из того, что возможно в данных условиях. Отсюда следует, что перевод может быть адекватным даже тогда, когда он не отвечает (или отвечает частично) критериям эквивалентности. Однако любое отступление от эквивалентности должно быть продиктовано объективной необходимостью, а не произволом переводчика. В противном случае речь может идти лишь о вольном переводе, не отвечающем требованиям адекватности.

IV. Существование различных видов перевода обусловлена рядом факторов: 1) переводческая деятельность осуществляется переводчиками в различных условиях, 2) переводимые тексты весьма разнообразны по тематике, жанровой принадлежности, 3) переводы выполняются в разной форме: устной и письменной, а значит требования, предъявляемые в от-

ношении точности и полноты перевода различны и т.д. В переводоведении различают следующие виды перевода:

1. Переводы, выделяемые по соотношению типов языка перевода и языка оригинала:

- **внутриязыковой перевод** – истолкование словесных знаков посредством знаков того же языка;
- **диахронический (исторический) перевод** (перевод на современный язык исторического текста, написанного на языке предшествующей эпохи);
- **транспозиция** (перевод текста одного жанра или функционального стиля в другой жанр или функциональный стиль);
- **межъязыковой перевод** – преобразование сообщения, выраженного средствами какой-либо одной знаковой системы, в сообщение, выраженное средствами другой знаковой системы;
- **бинарный перевод** (перевод с одного естественного языка на другой);
- **интерсемиотический перевод** (перевод с естественного языка на искусственный или наоборот);
- **трансмутация** (перевод с какого-либо искусственного языка на другой искусственный язык).

2. Переводы, выделяемые по общей характеристике субъекта переводческой деятельности и по его отношению к автору переводимого текста:

- **перевод, выполненный или выполняемый переводчиком, не являющимся одновременно автором переводимого текста;**
- **авторский (авто-) перевод** (перевод, выполненный автором оригинального текста);
- **авторизованный перевод** (перевод оригинального текста, апробированный автором);
- **машинный (автоматический) перевод** – перевод, выполненный или выполняемый компьютером;
- **смешанный перевод** – перевод с использованием значительной доли традиционной (или машинной) переработки текста.

3. Переводы, выполняемые по типу переводческой сегментации текста и по используемым единицам перевода:

- **по морфемный перевод** – перевод, выполняемый на уровне отдельных морфем без учета их структурных связей;
- **пословный перевод** – перевод, выполняемый на уровне отдельных слов без учета смысловых, синтаксических и стилистических связей между словами;
- **по фразовый перевод** – перевод, выполняемый на уровне отдельных предложений или фраз, переводимых последовательно одно за другим;
- **абзацно-фразовый перевод** – перевод, осуществляемый на уровне отдельных предложений или абзацев, переводимых последовательно одно за другим;
- **цельнотекстный перевод** – перевод целого текста, без выделения в качестве отдельных единиц перевода отдельных слов, предложений или абзацев.

4. Переводы, выделяемые по признаку формы презентации текста перевода и текста оригинала:

- *письменный перевод письменного текста* (перевод письменного текста, выполненный в письменной форме);
- *письменный перевод устного текста* (перевод устного текста, выполненный в письменной форме).
- устный перевод – перевод, выполненный в устной форме:
- *устный перевод устного текста* (перевод устного текста, выполненный в устной форме):
 - *синхронный перевод* (устный перевод, осуществляемый практически одновременно с произнесением текста оригинала);
 - *последовательный перевод* (разновидность устного перевода, осуществляемого после прослушивания определенной единицы текста, в паузах между этими единицами);
 - *односторонний перевод* (устный перевод, осуществляемый только в одном направлении, т.е. с данного языка на какой-либо другой язык);
 - *двусторонний перевод* (последовательный устный перевод беседы, осуществляемый с одного языка на другой и обратно);
- *устный перевод письменного текста* (перевод письменного текста, выполненный в устной форме).

5. Переводы, выделяемые по признаку характера и качества соответствия текста перевода тексту оригинала:

- *вольный (свободный) перевод* (перевод, воспроизводящий основную информацию оригинала с возможными отклонениями – добавлениями, пропусками и т.п.; осуществляется на уровне текста, поэтому для него оказываются нерелевантными категории эквивалентности языковых единиц);
- *интерпретация* (вид перевода, основанный на обращении к внеязыковой деятельности, в отличие от собственно перевода, осуществляемого по заданным правилам перехода от средств выражения, принадлежащих одной языковой системе, к средствам выражения, принадлежащим другой языковой системе);
- *адекватный перевод* (перевод, соответствующий оригиналу и выражающий те же коммуникативные установки, что и оригинал);
- *точный (правильный) перевод* (перевод, характеризующийся свойством семантической точности, т.е. семантически полно и правильно передающий план содержания оригинала);
- *аутентичный перевод* (перевод официального документа, имеющий одинаковую юридическую силу с оригиналом; согласно международному праву текст договора может быть выработан и принят на одном языке, но его аутентичность установлена на двух и более языках);
- *заверенный перевод* (перевод, соответствие которого оригиналу подтверждается юридически).

6. Переводы, выделяемые по признаку жанрово-стилистической характеристики переводимого материала и жанровой принадлежности:

- *научно-технический перевод* (перевод научно-технических текстов и документации);
- *общественно-политический перевод* (перевод общественно-политических текстов);
- *художественный перевод* (перевод художественных текстов);
- *военный перевод* (перевод текстов по военной тематике);
- *юридический перевод* (перевод текстов юридического характера);
- *бытовой перевод* (перевод текстов разговорно-бытового характера).

7. Переводы, выделяемые по признакам полноты и способа передачи смыслового содержания оригинала:

- *полный (слошной) перевод* – перевод, передающий смысловое содержание оригинала без пропусков и сокращений.
- *неполный перевод* – перевод, передающий смысловое содержание оригинала с пропусками и сокращениями:
- *сокращенный перевод* (перевод, передающий смысловое содержание текста в свернутом виде, т.е. с сокращением);
- *фрагментарный перевод* (перевод не целого текста, а лишь отдельного отрывка или отрывков);
- *аспектный перевод* (перевод лишь части текста в соответствии с каким-либо заданным признаком отбора (аспектом));
- *аннотационный перевод* (перевод, в котором отражаются лишь главная тема, предмет и назначение переводимого текста);
- *реферативный перевод* (перевод, в котором содержатся относительно подробные сведения о реферируемом документе – его назначении, тематике, методах исследования, полученных результатах).

8. Переводы, выполняемые по признаку основной прагматической функции:

- *практический перевод* – перевод, предназначенный для практического использования в качестве источника информации;
- *рабочий (информационный) перевод* (неотредактированный перевод, в основном пригодный для практического использования, но не оформленный для опубликования);
- *консультативный перевод* (вид информационного перевода, осуществляется обычно в устной форме, включает элементы аннотирования, реферирования и выборочного перевода с листа, выполняется, как правило, в присутствии заказчика, уточняющего по ходу перевода интересующие его аспекты содержания текста оригинала);
- *издательский (печатный) перевод* (письменный перевод, тиражированный при помощи средств массового размножения и предназначенный для широкого распространения);
- *опубликованный перевод* (практический или учебный перевод, тиражированный при помощи средств массового размножения);

- учебный перевод – перевод, используемый в учебном процессе для подготовки переводчиков или как один из приемов обучения иностранному языку;
 - экспериментальный перевод – перевод, выполненный с исследовательской целью;
 - эталонный перевод – образцовый перевод, используемый для сравнения с квалифицируемым переводом.
9. Переводы, выделяемые по признаку первичности/непервичности текста оригинала:
- прямой (первичный, непосредственный) перевод – перевод, выполненный непосредственно с оригинала;
 - косвенный (вторичный, непрямой) перевод – перевод, осуществленный не непосредственно с текста оригинала, а с его перевода на какой-либо другой язык;
 - обратный перевод – экспериментальный или учебный перевод уже переведенного текста на исходный язык.
10. Переводы, выделяемые по типу адекватности:
- семантико-стилистический адекватный перевод – семантически полный и точный и стилистически эквивалентный перевод, соответствующий функционально-стилистическим нормам языка перевода;
 - прагматический (функционально) адекватный перевод – перевод, правильно передающий основную (доминирующую) коммуникативную функцию оригинала;
 - дезигнативно-адекватный перевод – перевод, полно и правильно отвечающий на информационный запрос потребителя и не обязательно передающий полное смысловое содержание и ведущую коммуникативную функцию оригинала.

V. Машинный, точнее компьютерный перевод это письменный перевод, поскольку в результате мы получаем письменный текст. Однако осуществляет его не переводчик, а особая компьютерная программа. Современные компьютерные программы достаточно совершенны, но они до сих пор не могут разрешить самую сложную задачу процесса перевода: выбор контекстуально необходимого варианта, который в каждом тексте обусловлен многими причинами. В настоящее время результат этого вида перевода может быть использован как черновой вариант будущего текста, который будет редактировать переводчик, а также как средство, чтобы в крайней ситуации отсутствия переводчика получить общее представление о теме и содержании текста.

VI. Принадлежность текстов оригинала и перевода к определенному функциональному стилю предъявляет особые требования к переводчику и оказывает влияние на ход и результат переводческого процесса. Специфика определенного вида перевода зависит не только от языковых особенностей, которые обнаруживаются в соответствующем стиле каждого из языков, участвующих в переводе, но, главным образом, тем, как

соотносятся эти особенности между собой, насколько совпадают стилистические характеристики данного типа материалов в обоих языках. Если какие-то особенности обнаруживаются только в одном из языков, то при переводе происходит своеобразная стилистическая адаптация: специфические средства изложения в оригинале заменяются языковыми средствами, отвечающими требованиям данного стиля в ПЯ.

Специфическими особенностями, оказывающими влияние на процесс перевода, обладает и газетно-информационный стиль. Основная задача материалов этого стиля заключается в сообщении определенных сведений с определенных позиций; тем самым в достижении желаемого воздействия на реципиента. Содержание газетно-информационных сообщений характеризуется тем, что здесь речь идет о явлениях, доступных для понимания широким слоям неспециалистов, прямо или косвенно связанных их жизнью и интересами. Поскольку задача заключается в сообщении каких-то фактов, то возникает необходимость точного обозначения понятий и явлений. Отсюда важная роль терминов, имен и названий, однозначно указывающих на предмет мысли.

В газетно-информационных материалах нередко встречаются многозначные термины, термины-синонимы, сокращенные термины и названия. Например, термин *state* в политической терминологии США может значить как государство, так и штат: *Both the state and Federal authority are bent on establishing a police state*. В первом случае термин *state* стоит в одном ряду с определением «федеральный» и несомненно обозначает правительства штатов в отличие от правительства всей страны. Во втором случае *state* употреблено в значении государство.

Один и тот же термин может получать разное значение зависимости от идейной направленности текста, в котором он использован. Термин *idealism* может использоваться в философском смысле как название мировоззрения, противостоящего материализму, и иметь положительный или отрицательный смысл в зависимости от идейной позиции автора. Но еще чаще он используется в положительном смысле, непосредственно соотносясь с понятием *ideals* – идеалы и означая «служение (приверженность) высоким идеалам (или принципам)», например: *The Foreign Secretary's most elaborate and numerous speeches seem to prove that idealism is his guiding star*.

Широкое использование в газетно-информационном стиле имен и названий делает сообщение конкретным и относит передаваемые сведения к определенным лицам, учреждениям или районам. Это предполагает значительные предварительные (фоновые) знания у реципиента, позволяющие ему связать название с называемым объектом. Так, англичанину вне контекста хорошо известно, что *Park Lane* – это улица, *Piccadilly Circus* – площадь, а *Columbia Pictures* – кинокомпания. Названия и имена нередко используются в газетно-информационных материалах в сокращенной форме. Нередко эти сокращения могут быть неизвестны широ-

кому читателю и их значение тут же расшифровывается в самой заметке или сообщении. Но существует немало таких сокращенных названий, к которым читатели газеты давно привыкли и которые поэтому не нуждаются в разъяснениях. Обилие сокращений является характерной чертой газетно-информационного стиля современного английского языка. Ср. названия партий, профсоюзов, различного рода организаций и должностей: *GOP = Grand Old (Republican) Party; DD = Defense Department; DA = District Attorney*; фамилии или фамильярные прозвища известных политических или общественных деятелей: *JFK = John F. Kennedy; Rocky = Rockefeller; Ike = Eisenhower; RLS = Robert Louis Stevenson*; географические названия: *N.J. = New Jersey; Mo. = Missouri; SF = San Francisco; S.P. = South Pacific; E-W = East-West, etc.*

Характерная особенность английского газетно-информационного стиля заключается в стилистической разноплановости лексики. Наряду с книжной лексикой здесь широко используются разговорные и поэтические слова и сочетания:

Instead of answering the Minister took the line of "you're another", that other West German Ministries and the police had still more ex-nazis in them than his own ministry.

В области фразеологии газетно-информационный стиль отличает широкое использование «готовых формул» или клише. Здесь мы находим как многочисленные вводные обороты, указывающие на источник информации (*it is reported; it is claimed; our correspondent reports from; according to well-informed sources*), устойчивые сочетания со стержней образностью (*to set the tone; to throw light; to lay the corner-stone; to give the lie*), так и целый ряд политических штампов типа: *government reshuffle; vested interests; an unnamed Power; generation gap; a foregone conclusion, etc.*

В газетно-информационных материалах отмечаются и некоторые особенности синтаксической организации текста: наличие кратких самостоятельных сообщений (1–3 высказывания), состоящих из длинных предложений со сложной структурой (*Marooned by a gale on a skeleton of a fire-gutted Wyle light-house in Morecombe Bay, with their dingey swamped, nine workmen last night decided to risk the two-mile journey back over the sands to Fleetwood*), максимальное дробление текста на абзацы, когда почти каждое предложение начинается с новой строки, наличие подзаголовков в корпусе текста для повышения интереса читателей, частое использование многочисленных атрибутивных групп (*Paris underground and bus transport services were stopped today by a 24-hour warning strike called by the CGT (French TUC) with the support of other unions*). Особенно четко лексико-грамматическая специфика газетно-информационного стиля проявляется в газетных заголовках.

В области лексики для заголовков английских газет характерно частое использование небольшого числа специальных слов, составляющих своего рода «заголовочный жаргон»: *ban, bid, claim, crack, crash, cut,*

dash, hit, move, pact, plea, probe, quit, quiz, rap, rush, slash и др. Отличительной особенностью такой «заголовочной лексики» является не только частота их употребления, но и универсальный характер их семантики. Слово *pact* в заголовке может означать не только пакт, но и договор, соглашение, сделку и т.п. Глагол *hit* может быть употреблен в связи с любым критическим выступлением. *Red* может означать и коммунистический, и социалистический; *bid* подразумевает и призыв, и приглашение, и попытку достичь определенной цели и т.д.:

National Gallery Launches Bid to Buy the Titian. – Национальная галерея пытается приобрести картину Тициана. *Bid to Stop New Police Powers.* – Призыв не допустить расширения прав полиции.

В газетных заголовках особенно широко используются жаргонизмы и другие лексические элементы разговорного стиля: *Report Raps Lack of Law Reform; Hits Gapers Housing Stand; Diet Lends JFK a Helping Hand*, etc. Даже если в самой статье какая-либо ситуация описывается в более сдержанном стиле, заголовок часто носит более разговорный характер. Ср. начало заметки в английской газете: *A leading Chinese diplomat has been accused of responsibility for violence against foreign embassies с ее заголовком: China Blames Diplomat for Embassy Rows.*

Газетные заголовки обладают и рядом грамматических особенностей. В английских и американских газетах преобладают глагольные заголовки типа: *Roads Hit Scotland; William Faulkner Is Dead; Exports to Russia Are Rising.* Глагольность обычно сохраняется также в заголовках, состоящих из вопросительного предложения: *Will There Be Another Major Slump Next Year?* Специфическая особенность английского заголовка заключается в возможности опустить подлежащее: *Hues Teen-Agers as Scabs; Want No War Hysteria in Toronto Schools; His Arrests of Peace Campaigners*, etc.

Существенные отличия от других функциональных стилей современного английского языка отмечаются в характере использования в заголовках глагольных временных форм. Английские и американские газеты, как правило, используют в заголовках неперфектные формы глагола. Когда речь идет о событиях, происшедших в недавнем прошлом, обычно используется настоящее неопределенное время: *Richard Aldington Dies 70; Concorde Lands at Heathrow.* Это самый распространенный тип заголовков; употребление настоящего неопределенного времени придает им живость, приближает события к читателю, делает его как бы участником этих событий и тем самым усиливает его интерес к публикуемому материалу. The Past Indefinite Tense употребляется в заголовках, относящихся к прошлым событиям, в основном, в тех случаях, когда в заголовке есть обстоятельство времени, либо если читателю известно, что описываемое событие произошло в определенный момент в прошлом: *Husband Disappeared Two Years Ago; Why Rockefeller Couldn't Buy a Landslide Victory?*, etc.

Для обозначения будущего времени в заголовках широко используется инфинитив: *America To Resume Testing; Laundry Workers To Vote on New Contract; World Unions To Fight Monopoly, etc.*

Важной особенностью английских газетных заголовков является распространенность в них эллиптической формы пассивного залога с опущением вспомогательного глагола *to be* для описания событий как в прошедшем, так и в настоящем времени: *Paris Protest March Staged by Students; 8-Year-Old Boy Kidnapped in Miami; All Piers Paralysed on East Coast, etc.*

Тема 1. История машинного перевода (2 ч.)

Началом истории машинного перевода принято считать 40-е гг. XX в. Директор отделения естественных наук Рокфеллеровского фонда Уоррен Уивер впервые сформулировал концепцию машинного перевода (МП).

В 50-е гг. прошлого столетия проводилось множество исследований в области МП. Однако отсутствие теоретической базы и низкий уровень технических средств послужили причиной невысоких результатов этих работ. Исторически машинный перевод является первой попыткой использования компьютеров для решения невычислительных задач (знаменитый Джорджтаунский эксперимент в США в 1954 г.; работы по машинному переводу в СССР, начавшиеся в 1954 г.). Развитие электронной техники, рост объема памяти и производительности компьютеров создавали иллюзию быстрого решения этой задачи. Идея захватила воображение ученых и администраторов. Практическая цель была простой: загрузить в память компьютера максимально возможный словарь и с его помощью из иноязычных текстов получать текст на родном языке в удобочитаемом виде. Однако первоначальная эйфория по поводу того, что столь трудоемкую работу можно поручить ЭВМ, сменилась разочарованием в связи с абсолютной непригодностью получаемых текстов. Приведем в качестве примера результаты работы одной из современных коммерческих систем перевода. Предложим ей перевести народное английское стихотворение, известное нам в переводе «Робин-Бобин» (текст этот очень простой, московские дети изучают его в начальной школе):

Robin, Robin, what a man!
He eats as much as no one can.
He ate a lot of fish, he ate a lot of
meat.
He ate a lot of ice-cream and a
sweet.
He ate a lot of porridge and ten eggs
And all the cookies Mother had.
He drank a lot of juice, he ate a cake
Then said: «I have a stomach-ache»

Малиновка, Малиновка, какой человек!
Он ест насколько никто не может.
Он съел много рыб, он съел много мяс.
Он съел много ледяных-сливки и слад-
кий.
Он съел много каша и десять яйцо
И вся Мать повары имела.
Он пил много соков, он съел торт
Затем сказал: «У меня есть желудок –
боль»

Сравним с художественным переводом К. Чуковского:

Робин Бобин Барабек
Скушал сорок человек.
И корову, и быка,
И кривого мясника,
И телегу, и дугу,
И метлу, и кочергу.
Скушал церковь, скушал дом,
И кузницу с кузнецом,
А потом и говорит:
– У меня живот болит!

Следующий пример показывает неустойчивость системы машинного перевода при обработке неоднозначностей. Два предложения по отдельности «Flyer flies» и «Flyers fly» переводятся «Летчик летает» и «Летчики летают», если же из тех же словосочетаний составить одно предложение «Flyer flies and flyers fly» получаем «Летчик летает и муха летчиков».

В 60-е гг. к исследовательской деятельности в области МП присоединяется СССР (лаборатория математической лингвистики Ленинградского государственного университета).

70–80-е гг. усиливается развитие вычислительной техники, что способствует большей успешности исследовательских работ в области МП.

С начала 90-х гг. активизируется развитие рынка ПК, информационных технологий и сети Интернет, что сделало возможным дальнейшее развитие систем МП.

Тема 2. Общая характеристика средств машинного перевода (2 ч.)

Машинный перевод, перспективы и границы его использования.

Принципы и методика машинного перевода. Организация интерактивного режима работы пользователя с программной средой в условиях машинного перевода текстовой информации.

Необходимо отметить различия подходов к переводу человека и машины. Квалифицированный переводчик понимает смысл текста и пересказывает его на другом языке словами и стилем, максимально близкими к оригиналу. Для компьютера этот путь выливается в решение двух задач: 1) перевод текста в некоторое внутреннее семантическое представление и 2) генерация по этому представлению текста на другом языке. Поскольку не только не решена сама по себе ни одна из этих задач, а нет даже общепринятой концепции семантического представления текстов, при автоматическом переводе приходится фактически делать «подстрочник», заменяя по отдельности слова одного языка на слова другого и пытаясь после этого придать получившемуся предложению некоторую синтаксическую согласованность. Смысл при этом может быть искажен или безвозвратно утерян. Конечно, системы, настроенные на определенную

предметную область, дают гораздо более приемлемые результаты. Однако в этом случае системы перевода получаются очень узко ориентированными, и попытка использовать их даже в смежных предметных областях дает совершенно непредсказуемые результаты.

Актуальной для систем МП является проблема согласованности. Согласованность важна в рамках одного документа: один и тот же термин, даже если его без потери смысла можно перевести несколькими словосочетаниями, должен переводиться одинаково на протяжении всего документа. Однако еще более важной является согласованность в широком смысле – разработка и применение единой концепции интерпретации одного и того же термина на разных языках (скажем, американский разработчик программного обеспечения может быть недоволен, что термин *dialog box* переводится на итальянский как *finestra* (окно) и как *boite* (коробка, ящик) на французский). Ошибки, возникающие вследствие нарушения согласованности, являются серьезной проблемой, так как, имея только текст-результат перевода, уже невозможно установить, какие термины в оригинале были одинаковыми, а теперь переведены по-разному (в отличие от орфографических ошибок, которые исправить никогда не поздно).

Наряду с установленными правилами построения предложения в каждом языке существуют и свои неписанные законы, которые иногда называются красотами языка. Например, предложение на английском языке «This is my book» дословно переводится «Это есть моя книга», и формально это будет правильным, но по-русски так не говорят. В данном случае можно сказать, что предложение «написано так, будто его составил иностранец». Конечно, приведенный пример является простейшим, и возможность исключения слова «is» очень просто отражается в программе машинного перевода. Но на практике получившийся перевод похож на текст, написанный иностранцем.

Текст также может содержать слова, которые нужно понимать в контексте образа жизни людей в конкретной стране. Например, под словом «демократ» в США подразумеваются политики, выступающие за большее вмешательство государства в экономику, а в России те, кто выступает за большую свободу рынка. Это разные понятия.

Если компьютер используется для перевода литературных текстов, то получается черновой вариант текста, так называемый подстрочник, который превращается в произведение искусства человеком, слабо владеющим языком оригинала, но являющимся хорошим литературным редактором. Если же речь идет о переводе технических текстов, то здесь при правильном выборе словаря по специальности, в рамках которой написан текст, получается вполне удовлетворительный результат, иногда не требующий последующего вмешательства. Вообще необходимость редактирования компьютерного перевода очень часто возникает в связи с проблемами, перечисленными выше. Для этого системы машинного перевода обязательно имеют средства редактирования текстов.

Тема 3. Электронные словари (2 ч.)

В зависимости от того или иного стиля и назначения текста одно и то же слово нередко имеет разные значения. В какой-то мере эта особенность учитывается в системах машинного перевода: предусмотрены сменные словари, иногда для каждого вида текста предусмотрен свой словарь. Если лексики одного машинного словаря не хватает и применяются несколько словарей одновременно, можно указать системе, из какого словаря нужно брать слово, если есть несколько вариантов его перевода. Наконец, программа сама может предлагать на выбор пользователю несколько вариантов перевода, и он выбирает подходящий вариант, так сказать, вручную. Могут возникнуть и проблемы с переводом слов в устойчивых словосочетаниях и фразеологизмах, но это вполне по силам компьютеру.

Для качественного перевода очень важно, чтобы практически все слова исходного текста легко было найти и в словаре системы. А те из них, которых в нем нет, переносятся в текст непереуведенными уже на выходе из системы, и их впоследствии переводят вручную при редактировании результатов перевода. Такие слова могут повлиять на качество перевода предложения. Дело в том, что для определения, к какой части речи относится рассматриваемое слово, система производит анализ всего предложения в целом. При этом имитируется мыслительная деятельность человека (такую систему принято называть системой с элементами искусственного интеллекта). Если значение хотя бы одного слова в предложении не определено, то это может исказить анализ всего предложения, а иногда и результаты всего перевода.

Специализированные словари (принцип фразеологического перевода)

Для систем фразеологического перевода необходимы машинные словари большого объема. Такие словари должны создаваться прежде всего на основе автоматизированной обработки двуязычных текстов, являющихся переводами друг друга, и в процессе функционирования систем перевода. Наряду с основным политематическим словарем большого объема, в системах фразеологического машинного перевода целесообразно использовать также набор небольших по объему дополнительных тематических словарей. Дополнительные словари должны содержать только ту информацию, которая отсутствует в основном словаре (например, информацию о приоритетных переводных эквивалентах словосочетаний и слов для различных предметных областей, если эти эквиваленты не совпадают с приоритетными переводными эквивалентами основного словаря).

Тема 4. Системы машинного перевода (4 ч.)

В рамках данной темы будут рассмотрены принципы построения интерфейса современных пакетов машинного перевода и типовые процедуры работы в программной среде. А также современные программные средства машинного перевода, их использование в редакционно-издательском процессе. Особенности машинного перевода текстов разной тематической направленности.

В настоящее время несколько десятков компаний занимаются разработкой систем МП: Trident Software, Cross Language, Ata Software, Transparent Language, Lingvistica b.v. и др.

Есть возможность воспользоваться услугами автоматических переводчиков в сети Интернет: PROMT's online Translator, LogoMedia.net и др.

Существуют также системы автоматизированного перевода – программы, которые не берут на себя полностью весь перевод, а лишь помогают человеку-переводчику справиться с некоторыми трудностями (Computer Aided Translation). Одним из примеров таких систем является EuroLang Optimizer. Его можно рассматривать как нечто переходное между компьютерным словарем и программой-переводчиком, как некий набор предметно-ориентированных глоссариев, снабженный интерфейсом для удобства переводчика: предлагается несколько вариантов перевода, выделенные разными цветами в зависимости от условий применимости; переводчик может с помощью меню определенным образом настраивать словари для более быстрого и правильного выбора нужного эквивалента.

Подобные программные средства могут помочь в решении проблем, связанных с терминологией и вообще со знаниями переводчика о предметной области: одни и те же слова могут по-разному переводиться в зависимости от того, о каком предмете идет речь.

В последнее время также появляются автоматизированные системы «доперевода» или «перевода изменений». Их возникновение связано с тем, что большинство технических текстов (описания, инструкции) не являются целиком новыми (как и явления, продукты, механизмы и т.п., ими описываемые), а содержат в себе лишь некоторые изменения, связанные, например, с усовершенствованием конструкции. Система «доперевода» извлекает из памяти знакомые предложения, а новые куски предлагает переводчику. Заметим, что такой человек-машинный способ генерации новых текстов также помогает согласованности в стиле и терминологии при переходе от одной версии к другой.

Развитием систем подобного вида можно считать канадскую (Канада – двуязычная страна, постоянно сталкивающаяся с проблемой перевода на государственном уровне) систему генерации прогнозов погоды Forecast Generator (FOG). Можно считать, что в ней перевод полностью заменен генерацией текстов. В памяти системы хранится 20 миллионов слов и словосочетаний, связанных с прогнозами погоды, что позволяет генерировать как английский, так и французский вариант непосредственно из базы данных. Конечно, успешная работа этой системы в значительной мере объясняется ограниченной природой текстов: сообщения о погоде являются классическим примером подъязыка. Ограниченность словаря, грамматики и семантики дает возможность достичь отличных результатов сравнительно простыми методами.

Концепция фразеологического машинного перевода

Концепция фразеологического МП была впервые предложена профессором Г.Г. Белоноговым в 1975 г. Главным тезисом этой концеп-

ции является утверждение, что наименования понятий в текстах чаще всего обозначаются не отдельными словами, а словосочетаниями; и смысл словосочетаний, как правило, не сводим или не полностью сводим к смыслу составляющих их слов. Поэтому основными единицами языка и речи, которые следует, прежде всего включать в машинные словари, должны быть фразеологические единицы (словосочетания и фразы). Отдельные слова также могут включаться в словари, но они должны использоваться только в тех случаях, когда не удастся осуществить перевод, опираясь на фразеологические единицы.

Реальные тексты, независимо от их принадлежности к той или иной тематической области, обычно бывают политематическими, если они имеют достаточно большой объем. И отличаются они друг от друга не столько словарным составом, сколько распределениями вероятностей появления в них различных слов и словосочетаний из общенационального словарного фонда. Поэтому машинный словарь, предназначенный для перевода текстов даже только по одной тематике, должен быть политематическим, а для перевода текстов из различных предметных областей тем более.

Основным средством разрешения полисемии (многозначности) слов в системах фразеологического машинного перевода является их использование в составе фразеологических словосочетаний, которые переводятся как целостные единицы смысла. Дополнительным – аппарат дополнительных тематических словарей, где для каждого многозначного слова или словосочетания указывается его приоритетный переводной эквивалент, специфичный для рассматриваемой предметной области.

Важную роль в системах фразеологического машинного перевода должны играть процедуры морфологического и синтаксического анализа и синтеза текстов, построенные на основе принципа аналогии. Такие процедуры позволяют отказаться от хранения в памяти ЭВМ «грамматических» словарей большого объема и делают систему перевода открытой – способной обрабатывать тексты с «новой» лексикой и с «новыми» синтаксическими структурами.

Наряду с переводом текстов в автоматическом режиме, целесообразно в системах фразеологического машинного перевода предусмотреть и интерактивный режим их работы. В этом режиме пользователь должен иметь возможность вмешиваться в процесс перевода и настраивать дополнительные машинные словари на тематику переводимых текстов. При этом он может вводить в систему любые новые слова и словосочетания в любой грамматической форме и должен следить только за смыслом вводимых им элементов, а не за их грамматическими характеристиками. Эти характеристики должны определяться автоматически.

Системы МП, основанные на статическом анализе.

В последнее десятилетие наметилась тенденция к использованию систем машинного перевода, основанных на статистическом анализе.

Безусловно, можно добиться выигрыша благодаря применению статистических данных и четкой вероятностной модели того, в чем состоит качественный анализ или передача текста, на любом из этапов, показанных на рисунке. Но под понятием «статистического машинного перевода» подразумевается общий подход к решению проблемы перевода, который основан на поиске наиболее вероятного перевода предложения с использованием данных, полученных из двуязычной совокупности текстов. В качестве примера двуязычной совокупности текстов можно назвать парламентские отчеты, которые представляют собой протоколы дебатов в парламенте. Двуязычные парламентские отчеты издаются в Канаде, Гонконге и других странах; официальные документы Европейского экономического сообщества издаются на 11 языках; а Организация объединенных наций публикует документы на нескольких языках. Как оказалось, эти материалы представляют собой бесценные ресурсы для статистического машинного перевода.



Схематическое изображение вариантов организации систем машинного перевода. Схема начинается с английского текста, показанного в левой нижней части. Система с промежуточным языком следует по сплошным линиям, выполняя синтаксический анализ английского текста и преобразуя его вначале в синтаксическую форму, затем в семантическую форму представления и в форму представления на промежуточном языке, после этого выполняет этапы преобразования в семантическую, синтаксическую и лексическую форму на французском языке. В системе на основе передачи в качестве сокращенных путей используются пунктирные линии. В различных системах передача осуществляется на разных уровнях, причем в некоторых системах она происходит одновременно на нескольких уровнях

Практические задания

Объясните, чем обусловлены семантико-структурные различия оригинального и переводного текстов. Отредактируйте перевод.

People, politics ease China's disaster evacuation efforts

BEIJING, China (CNN)— As Typhoon Krosa barreled toward eastern China over the weekend, every level of the Chinese government was whipped into action, evacuating 1.4 million people from the two coastal provinces on which Krosa set its sights.

Krosa stormed across Taiwan, reportedly killing five people, before hammering Zhejiang and Fujian provinces with ferocious winds, rain, flooding and general devastation. Roads were blocked with debris, residents were left without power, businesses closed, schools shut down — but there were no official reports of deaths in China.

Analysts and observers said China's ability to move hundreds of thousands of people out of harm's way speaks to the country's exceptional ability to mobilize resources and people to deal with disasters.

«We do a decent job in responding to earthquakes», said a Chinese official, requesting anonymity. «We do a better job in dealing with natural disasters that we can anticipate, like typhoons. We can often avoid or minimize deaths if not destruction».

THERE is no doubt that JESSICA SIMPSON has been looking in great shape recently.

She proved she could still rock a pair of Daisy Duke hotpants and turned heads during New York fashion week.

So it was no surprise to see her exercising again — though this time as part of her new film role in Major Movie Star.

Люди, политика ослабляет усилия по эвакуации бедствия Китая

ПЕКИН, Китай (Си-Эн-Эн) — Как Тайфун Кроса, разлитый к восточному Китаю до конца недели, каждый уровень китайского правительства хлестали в действие, эвакуируя 1.4 миллиона человек из двух прибрежных областей, на которых Кроса устанавливали его достопримечательности.

Кроса штурмовал поперек Тайваня, по сообщениям убивая пять человек, перед стучащими областями Чжэцзяна и Фуцзяни со свирепыми ветрами, дождем, затопляя и общим опустошением. Дороги были блокированы с развалинами, жителей оставляли без власти, закрытые фирмы, закрытые школы — но не было никаких официальных сообщений смертельных случаев в Китае.

Аналитики и наблюдатели сказали, что способность Китая перемещать сотни тысяч людей из пути вреда говорит с исключительной способностью страны мобилизовать ресурсы и людей, чтобы иметь дело с бедствиями.

«Мы делаем приличную работу в отечении на землетрясения», сказал китайский чиновник, прося анонимность. «Мы делаем лучшую работу имея дело с стихийными бедствиями, которые мы можем ожидать, как тайфуны. Мы можем часто избегать или минимизировать смертельные случаи если не разрушение».

Имеется без сомнения, что ДЖЕССИКА SIMPSON смотрела в большой форме недавно.

Она доказала, что она могла все еще качать пару Герцога Маргаритки hotpants и направила главы в течение Нью-Йоркской моды неделя.

Так что это не было никакое удивление (неожиданность), чтобы видеть ее осуществление снова — хотя на сей раз как часть ее новой роли фильма в Главной Кинозвезде.

<p>Jess, who plays a fallen starlet who ends up joining the army, was presumably filming one of her boot camp scenes here.</p> <p>Whatever the reason, it certainly looks good on her...</p>	<p>Дже, кто играет купающуюся звездочку, кто заканчивает соединять армию, возможно, снимал одну из ее сцен лагеря ботинка здесь.</p> <p>Безотносительно причины, это конечно выглядит хорошим на ее ...</p>
--	---

Библиографический список

1. Андреева Е.В. История машинного перевода [Электронный ресурс]. <http://www.laugust.ru/etc/history.shtml>
2. Латышев Л.К. Технология перевода. М., 2002. 280 с.
3. Мухачев П.А. Машинный перевод в XX веке. Как это было. Как этого нет // Мир Интернет. 2001. №2. С. 14–23.
4. Паршин А. Теория и практика перевода. М., 2005. 213 с.
5. Системы машинного перевода и словари: Обзор продуктов // Компьютер пресс. 2001. №9. С. 73–81.
6. Система перевода текста PROMT Internet: Руководство пользователя. СПб., 1999.
7. Система электронных словарей LINGVO для DOS и Windows. Версия 4.0: Руководство пользователя. М., 1995.
8. Федоров А.В. Основы общей теории перевода. М., 2005. 325 с.