

тывать это при составлении алгоритма. Для обработки оборотов, разорванных другими членами предложения, необходимы данные синтаксического анализа.

При машинном переводе текстов, содержащих идиоматические выражения, фразеологические единицы включаются в машинный словарь, как основные единицы языка и речи, а перевод отдельных слов используется только, когда не удастся осуществить перевод, опираясь на фразеологические единицы. Так же учитывается, что тексты, имеющие большой объем, обычно бывают политематическими и отличаются друг от друга распределениями вероятностей появления в них различных слов из общенационального словарного фонда. Поэтому машинный словарь, предназначенный для перевода текстов из одной тематической области, как и из различных предметных областей, является политематическим. В системах фразеологического машинного перевода наряду с основным политематическим словарем большого объема используются наборы небольших по объему дополнительных тематических словарей. Такие словари содержат информацию, которая отсутствует в основном словаре. Например, информацию о приоритетных переводных эквивалентах словосочетаний и слов для различных предметных областей, когда они не совпадают с приоритетными переводными эквивалентами основного словаря. В системах фразеологического машинного перевода используется и интерактивный режим работы. В этом режиме пользователь имеет возможность вмешиваться в процесс перевода и настраивать дополнительные машинные словари на тематику переводимых текстов. Для получения качественного машинного перевода текстов, особенно политематических, необходимы машинные словари большого объема, которые могут создаваться на основе автоматизированной обработки двуязычных параллельных текстов и в процессе функционирования систем перевода.

А. В. Зубов

ФОРМАЛИЗАЦИЯ СОДЕРЖАНИЯ РУССКОГО ЛИРИЧЕСКОГО СТИХОТВОРЕНИЯ

Под стихотворением понимается разбитые на строфы последовательности семантически связанных предложений. Поль Валери заметил, что «стихотворение создается не из идей и не из чувств, оно создается из слов». Естественно, что без идеи и чувств создать стихотворение невозможно, но когда речь идет об анализе уже готового стихотворения главным становится слово.

Известно, что каждый текст выражает индивидуальные лингвистические и экстралингвистические знания автора. Лингвистические знания связаны со спецификой употребляемых автором слов, словосочетаний, типов предложений и их последовательностей (абзацев, строф). Экстралингвистические

особенности автора проявляются в выборе манеры изложения, в способе представления ситуаций окружающей действительности, выражающемся структурной организацией абзаца или строфы, порядке следования в тексте этих укрупненных единиц и т.п.

Выделяя в содержании каждого текста две составляющие (статическую и динамическую), можно определенным образом смоделировать заложенные в текст лингвистические и экстралингвистические знания автора текста.

Статическая составляющая текста может быть представлена в виде таблицы основного статического содержания, в которую путем статистического анализа текста включены наиболее важные для данного текста знаменательные слова – главные и второстепенные опорные слова. Они, с учетом всех словарных и контекстуальных синонимов текста, отражают главные действующие лица текста, место и время их действия, зафиксированные в исследуемом тексте.

Динамическая составляющая каждого текста может быть описана некоторой логико-семантической формулой этого текста. Если говорить о тексте стихотворения, то предполагается, что каждая строфа содержит описание некоторой микроситуации, т.е. определенного числа субъектов, объектов реального мира с их свойствами и взаимоотношениями.

Основной смысловой единицей динамической составляющей содержания стихотворения является предметно-логическое содержание строфы. Под предметно-логическим содержанием строфы понимается дифференцированное по типам главных и второстепенных субъектов, объектов, мест и времени действия ситуации, представленной в стихотворении, перечисление фактов объективной действительности (явлений, событий, состояний, признаков и т.д.), связанных с содержанием всего стихотворения. Например: «Констатация некоторого действия главного субъекта и описание природы». Или «Описание состояния и действий автора» и т.п. В текстах каждого автора можно выделить конечное число таких предметно-логических составляющих. Давая им определенные буквенные обозначения (например, M001, M002 и т.д.), можно построить логико-семантические формулы таких текстов.

Например: $T_{01} = T001 \& T012 \& T006 \& T021 \& T005$.

Предлагаемый подход к описанию содержания текстов демонстрируется в сообщении на примере более 100 русских стихотворений.

А. В. Зубов

СТРУКТУРА СЛОВАРЕЙ ДЛЯ АВТОМАТИЧЕСКОГО ПЕРЕВОДА ТЕКСТОВ

В наши дни автоматический перевод текстов проводится только для перевода научно-технических текстов, в составе которых преобладает, в основном, терминологическая лексика. Существуют различные подходы к классификации терминологических автоматических словарей.