

построение частотно-алфавитного словаря для каждого анализируемого текста. Частотно-алфавитные словари по анализируемому материалу были получены с помощью программы DICT. На втором этапе происходит выделение из частотно-алфавитного словаря потенциальных опорных слов, которые служат основой для деления опорных слов на главные и второстепенные. Для осуществления данной операции была использована компьютерная программа UNIFY. В заключение, все полученные списки потенциальных опорных слов были дополнены словами схожей тематики, и создан поисковый образ по каждой тематике. Все поисковые образы были сведены в таблицы. На основании разработанных поисковых образов по каждой тематике был составлен алгоритм автоматической идентификации тематики англоязычных публицистических текстов аналитических статей.

Таким образом, на основании проведенного исследования выявлена структура анализируемых текстов, критерии определения основного содержания текста англоязычной аналитической статьи, анализируемые статьи были классифицированы по тематикам, были созданы частотно-алфавитные словари и поисковые образы по анализируемым текстам статей, разработан алгоритм определения тематики анализируемых статей.

Д. В. Степанова

ИСПОЛЬЗОВАНИЕ КОНКОРДАНСОВ ДЛЯ РЕШЕНИЯ ЛИНГВИСТИЧЕСКИХ ЗАДАЧ

Создание электронных корпусов текстов и освоение их возможностей для решения практических и исследовательских задач является одним из приоритетных направлений корпусной лингвистики. Под электронным корпусом текстов принято понимать массив устных или письменных текстов на естественном языке, представленный на машинном носителе и упорядоченный должным образом для проведения лингвистических исследований. Среди наиболее известных корпусов английского и русского языков можно выделить *Национальный корпус русского языка*, *British National Corpus*, *Bank of English*, *American National Corpus*. Корпусный подход как метод лингвистического исследования направлен на прикладное изучение языка, особенностей его функционирования в реальной среде и текстах различных жанров и стилей.

Применение репрезентативных корпусов текстов для автоматического статистического и качественного анализа лексических единиц в естественной языковой среде становится возможным благодаря специальным компьютерным программам-конкордансам. Такие программы являются одним из видов программного обеспечения корпусной лингвистики и предназначены для автоматической обработки текстов в электронном формате посредством их преобразования в лингвистические базы данных. Примером такого программного обеспечения может послужить программа *Concordance*, которая выполняет две основные функции: поисковую, т.е. осуществление

поиска языковых единиц или целых фрагментов текста, и эвристическую, т.е. приведение контекста заданной длины для отобранной словоформы. Результатом работы программы-конкорданса является создание словаря, в котором каждая словоформа текста сопровождается не только частотным показателем, но и контекстом употребления. Технические возможности программы Concordance, такие как автоматическое построение алфавитно-частотного словаря словоформ заданного текста, выделение из текстов левой и правой контактной дистрибуции выделенной словоформы по заданным параметрам, автоматическая сортировка словоформ по алфавиту или частоте употребления, позволяют создать словарь – конкорданс, который включает индексные указатели, список словоупотреблений электронного корпуса текстов в контексте левого и правого окружения (контекста) с указанием частоты употребления в текстах, а также кластеры выделенных слов.

В прикладной лингвистике конкордансы получили особое признание среди лингвистов благодаря возможностям оперативной обработки лексического материала и эффективного исследования электронных корпусов текстов в целом.

В последнее время конкордансы активно используются при создании автоматических и автоматизированных систем обучения иностранным языкам, что связано с возможностью изучения лексических и грамматических особенностей словоформ текста, выявления лексико-грамматической сочетаемости слов, выделения примеров словоупотреблений и их контекстов, определение частотности лексем и грамматических конструкций в текстах разных авторов и жанров.

Создание конкордансов и анализ лингвистических данных имеет особое значение при работе с научно-техническими текстами. Это, прежде всего, связано с высокими темпами развития ряда предметных областей (например, в сфере информационных технологий), что влечет за собой исключение из употребления некоторых терминов или изменение их семантической структуры. В то же время разрабатываются и внедряются новые технологии, требующие номинации. При этом обозначение новых сложных понятий осуществляется, как правило, с помощью терминологических словосочетаний. В результате даже самые современные словари не успевают отразить все изменения в составе терминологии той или иной предметной области и зафиксировать новые термины.

Применение программ-конкордансов для компьютерной обработки современных электронных научно-технических текстов позволяет решить следующие задачи:

- определение частоты употребления терминов в заданных текстах и выделение основных терминов некоторой предметной области;
- оперативное установление контекста употребления выделенных лексических единиц по заданным параметрам для последующего анализа;
- выявление новых терминов некоторой предметной области, которые еще не отражены в словарях, но уже активно используются специалистами.

Таким образом, конкорданс как тип словаря является современным эффективным инструментом автоматической статистической обработки и анализа электронных корпусов текстов, который может применяться для решения различных лингвистических задач в лексикологии, в лексикографии и терминоведении, в теории и практике перевода, а также в преподавании и изучении всех аспектов языка.

Н. В. Шалимо, Е. С. Василенко, Е. И. Ильина

ПРОБЛЕМЫ ПРЕПОДАВАНИЯ ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ В МГЛУ

В современном мире происходит тотальная информатизация всех сфер человеческой деятельности, и это явление требует от каждого специалиста высокой информационной культуры, т.е. специалист должен уметь пользоваться информационными технологиями и инструментами, позволяющими распространять информацию, совместно работать в коллективе над едиными проектами, обладать способностью к инновациям.

Высшие учебные заведения являются базисом любой профессиональной деятельности, поэтому данным учреждениям необходимо наполнение профессиональными образовательными программами и дисциплинами, обеспечивающими развитие у студентов соответствующих знаний, умений и навыков. С этой целью МГЛУ включает в свою программу преподавания дисциплину «Основы информационных технологий» (ОИТ).

Актуальность преподавания этой дисциплины в настоящее время не вызывает сомнений, однако несмотря на почти двадцатипятилетний опыт преподавания ОИТ, у нас до сих пор существует ряд проблем в организации учебного процесса и методике преподавания. Всю совокупность этих проблем условно можно разделить на две группы.

К первой группе относятся, так называемые традиционные проблемы, связанные с быстрым моральным старением техники и программного обеспечения, постоянным появлением новых информационных технологий; не соответствие государственных стандартов образования реалиям современных информационных технологий. Быстрое моральное старение техники и программного обеспечения – эта основная проблема любого университета, в том числе и нашего, так как закупив год назад оборудование и программное обеспечение, сотрудники далекие от информационных технологий не видят необходимости его обновления. Однако сфера IT очень быстро развивается, и один год в информационных технологиях может равняться пяти, десяти годам разработок в другой области. Поэтому используя старое программное обеспечение для обучения студентов, университет сводит к минимуму пользу от его практического применения. Разработанные курсы обучения порой используют устаревшее программное обеспечение, что делает их не совсем пригодными к реалиям современной жизни.