

## НАУЧНОМУ ЯЗЫКУ - НАУЧНЫЕ МЕТОДЫ ИЗУЧЕНИЯ

**Захаров В.П.**

Санкт-Петербургский государственный университет, Россия, 199034, Санкт-Петербург,  
Университетская наб., 7-9, +7-911-9371769, v.zakharov@spbu.ru

Основная тема доклада — моделирование источниковедческой базы для исследования языков для специальных целей (ЯСЦ). Под моделированием источниковедческой базы мы понимаем создание корпусов специальных текстов в соответствии с критериями корпусной лингвистики и исследовательскими задачами. ЯСЦ занимают важное место в лингвистике, в системах научной коммуникации, в информационных технологиях, в обучении специалистов иностранным языкам. формальным характеристикам. Прежде всего это научно-технические (а также юридические, инструктивно-методические и т.п.) тексты, насыщенные специфической лексикой, отсутствующей в общих словарях и обладающей особой сочетаемостью. В настоящее время огромное развитие получили корпуса текстов как инструменты лингвистического анализа. Отсюда проистекает задача – изучение и описание ЯСЦ корпусно-ориентированными методами.

Понятие «язык для специальных целей» используется для обозначения функциональной разновидности языка, призванной обеспечить адекватное и эффективное общение специалистов на всех уровнях научно-технической (и иной) деятельности в определенной предметной области. С лингвистической точки зрения ЯСЦ понимаются как ограниченный набор лексических и грамматических конструкций, имеющих, с другой стороны, свои особенности по лексике, синтаксису, семантике. Для полноценного изучения языков для специальных целей необходим эмпирический материал, а именно, репрезентативные корпуса текстов ЯСЦ.

В докладе освещаются средства и методы создания таких корпусов, а также их использование для исследования языков и решения практических задач. Рассмотрены основные параметры специальных корпусов текста, возможности их варьирования и взаимосвязь с типами лингвистических задач. Сформулированы основные признаки научного термина на уровне формальных параметров, фиксируемых на разных уровнях описания языковой структуры: статистические, статистико-комбинаторные, синтаксические, морфологические. Отмечается влияние на терминосистемы русского языка лексической системы английского языка.

На основе дистрибутивно-статистического анализа данных специальных корпусов разработаны автоматизированные процедуры выделения терминов и терминологических словосочетаний и построения лексико-семантических полей. Созданы корпуса по 31 тематическому направлению общим объемом более 20 млн. словоупотреблений. Проведены эксперименты по автоматизированному выделению терминов и терминологических сочетаний. Обсуждаются проблемы оценки автоматизированных методов в лингвистике. Нам представляется, что полученные результаты имеют свою теоретическую значимость и практическую ценность.