

необязательностью изучения этого предмета в школе. Следует также добавить, что углубление знаний в сфере теоретической информатики, читаемой на лекциях, требует от аудитории определенного уровня знаний инновационных технологий и компьютерной лингвистики, которого у студентов нашего университета первого курса просто нет. Необходимо учесть, что планируемое изучение специальных пакетов прикладных программ, ориентированных на будущую профессию, имеет ограничение в связи с отсутствием у студентов знаний в профессиональной сфере.

В учебных планах нашего университета желательно организовать поэтапное изучение информационных технологий: на первом курсе знакомство с инструментарием, а на старших курсах – методики использования информационных технологий для решения реальных профессиональных задач. Выглядеть это может следующим образом: первый этап – получение на первом курсе практических знания в сфере информационных технологий при знакомстве с пакетом программ общего назначения, входящего в состав Microsoft Office – MS Word, MS Excel, MS Access, MS Power Point (ибо не вызывает сомнений полезность этих знаний и навыков, которые используются в дальнейшем студентами для подготовки текстовых документов, рефератов, курсовых и дипломных проектов, научных статей, создания презентаций своих творческих работ). Второй этап: на старших курсах, когда осуществляется освоение специальных педагогических и языковых дисциплин, в образовательных стандартах МГЛУ ввести дисциплины, предполагающие изучение информационных технологий в профессиональной деятельности. Таким образом, в настоящее время для подготовки высококвалифицированных специалистов – выпускников нашего университета, способных соответствовать требованиям современного рынка труда, по нашему мнению, необходимо вывести подход по изучению информационных технологий на более высокий качественный уровень.

Н. Г. Швец

ФОРМАЛЬНОЕ ПРЕДСТАВЛЕНИЕ ОСНОВНОГО СОДЕРЖАНИЯ ТЕКСТА

Одной из характеристик текста является наличие в нем единого содержания.

Основное содержание текста – это совокупность знаменательных слов, несущих основную семантическую нагрузку в тексте, формирующих в сознании адресата то же представление об описываемом фрагменте ситуации, которое вложил в текст автор.

Текст создается как некоторое сообщение о предметах и явлениях действительности, отражающее отношение к ним говорящего и рассчитанное на оказание определенного воздействия на принимающего это сообщение. Соответственно этому в семантике сообщения исследователи выделяют план содержания, отражающий взаимосвязь явлений и предметов действительности, и план смыслов, выражающий отношение личности к этим предметам и явлениям, учитывающее мотивы восприятия.

В ходе исследований было установлено, что для адекватного формального представления основного содержания, или темы текста, необходимо оперировать не отдельными словами, а определенным набором ключевых слов. Исходя из этого, набор ключевых слов представляется в виде специальной таблицы основного содержания текста (ТОС), в которой перечислены главные опорные слова текста (ГОС) и второстепенные опорные слова текста (ВОС). В ТОС в соответствии с предметными свойствами своих референтов ГОС и ВОС образуют группы опорных слов – субъектов, объектов, признаков, действий, слов-мест и слов-времени. При этом используется статистический метод выбора опорных, или ключевых слов, характеризующих основное содержание текста. При таком подходе основными критериями выделения опорных слов текста являются абсолютная частота употребления знаменательного слова (с учетом всех его возможных синонимов и замен) и количество абзацев, в которых встретилось слово. Этот метод определения основного содержания текста взят за основу в нашем исследовании.

Слова, выражающие наиболее важные для текста понятия, как правило, употребляются в нем неоднократно. По мнению некоторых исследователей, частота употребления знаменательного слова (с учетом его семантических эквивалентов) свидетельствует о важности обозначаемого им понятия для данного текста, коррелирует с семантической (тематической) нагрузкой данного слова и может быть использована для определения темы текста.

Однако частота употребления лексической единицы в тексте не всегда может быть однозначным параметром, так как признак частотности ключевого слова не является превалирующим. В целях получения более качественного результата при выявлении основного содержания текстов рекламных объявлений мы использовали комбинированный подход, сочетающий в себе элементы статистического и позиционного методов извлечения ключевых слов из текста.

Одним из примеров позиционного метода является метод ключевых слов в заголовке. При использовании данного метода составляется список опорных слов из заголовков, исходя из предположения о том, что в них должны быть отражены основные положения документа. Заголовок занимает одну из самых «сильных» позиций в тексте, фиксирует внимание на его важнейших аспектах. В заголовке обобщается существенная информация текста и имплицитно (или эксплицитно) присутствуют ассоциативные связи с другими фрагментами текста, поэтому заголовок является информационно-значимым элементом текста. Лингвистические исследования подтверждают высокую информативную значимость заголовков, заключающуюся в способности в концентрированном виде выразить содержание текста.

Процедура выделения ключевых (опорных) слов предполагала несколько этапов.

На первом этапе для каждого исследуемого текста РО был построен алфавитно-частотный словарь. В нем приведены алфавитные списки лексических единиц текста с указанием частоты употребления каждого слова в абзаце и в тексте.

На втором этапе анализа из словаря анализируемого текста были удалены все служебные слова, а также все знаменательные лексические единицы с частотой употребления, равной 1, кроме того были объединены синонимы и грамматические формы одного и того же слова посредством суммирования частот их употребления в тексте. Оставшиеся лексические единицы текста составили словари потенциальных опорных слов.

На третьем этапе анализа для каждого слова словаря потенциальных опорных слов был подсчитан коэффициент важности слова в тексте по формуле $K_{\text{важ}} = (F_{\text{xm}}) / (N_{\text{xn}})$, где F – абсолютная частота употребления каждого слова в тексте, в которую входит суммарная частота всех типов синонимов этого слова и местоименных замен; m – количество абзацев, в которых встретилось слово; n – общее число абзацев в тексте; N – общее число слов в тексте.

Этот коэффициент является основным критерием определения «степени важности» слова для содержания всего текста, в структуре опорных слов которого имеется определенная иерархия. Наиболее частотные из них (с учетом их синонимов и местоименных замен) и встречающиеся в наибольшем числе абзацев являются *главными*, так как определяют основной предмет сообщения. *Второстепенные* опорные слова связаны с главными, но встречаются с меньшей частотой и в меньшем числе абзацев.

На четвертом этапе анализа были вычислены критические значения коэффициентов важности, которые позволили формальным способом отделить главные и второстепенные опорные слова.

На пятом этапе анализа были получены таблицы основного содержания для каждого исследуемого текста, которые были дополнены ключевыми словами из заголовков (КСЗ) соответствующих текстов рекламных объявлений. В таких таблицах опорные слова в соответствии с предметными свойствами своих референтов в общем случае могут быть разделены на следующие группы: слова-объекты; слова-признаки; слова-действия; прочие слова.

Преимущество данной методики определения основного содержания текста состоит:

- 1) в возможности классифицировать слова конкретного текста в зависимости от степени их важности для семантической структуры текста по нескольким группам,
- 2) в ее применимости к текстам с разным количеством абзацев.

A. Shevel, R. Eckhart

THE POSSIBILITIES OF ONLINE INFORMAL LEARNING OF ENGLISH FOR DIGITAL NATIVES

Online informal learning of English is one of the most important trends to follow and be aware of. Whether teachers are paying attention or not, their students are spending a lot of time on the internet watching all sorts of English-language content.