

Литература:

1. Абдуль-Рида, А. А. Методика обучения употреблению русских предлогов на занятиях по РКИ в арабской аудитории: автореф. дис. ... канд. пед. наук / А. А. Абдуль-Рида. – М., 2005. – 24 с.
2. Арбузова, И. И. Теоретические и методологические основы разработки курса обучения иностранцев теме «Пространственные предлоги», 2011. – Электронный ресурс. Режим доступа: <http://www.emissia.org/offline/2011/1582.htm>. – Дата доступа: 10.10.2020.
3. Литвинов, П. П. Предлоги: повышаем речевую готовность на английском / П. П. Литвинов. – М. : ВАКО, 2004. – 223 с.

НАЦИОНАЛЬНЫЙ КОРПУС РУССКОГО ЯЗЫКА: СЕМАНТИЧЕСКАЯ РАЗМЕТКА

Флянтикова Е.В.

УО «Гродненский государственный медицинский университет»
Кафедра русского и белорусского языков

Корпус – это информационно-справочная система, основанная на собрании текстов на некотором языке в электронной форме [2]. Собрание текстов в электронном виде, сопровождаемое разметкой, называется лингвистическим корпусом, независимо от его объема [1]. Национальный корпус русского языка (далее – НКРЯ) [3] – полнотекстовый языковой ресурс, в котором представлен «данный язык на определенном этапе (или этапах) его существования во всем многообразии жанров, стилей, территориальных и социальных вариантов и т. п.» [3]. То есть в НКРЯ достаточно сбалансированно представлены разнообразные жанры устных и письменных текстов.

Однако НКРЯ, как и любой лингвистический корпус, – это не просто некое собрание текстов. Одной из главных черт лингвистического корпуса является наличие в нем специальной разметки. Именно наличие разметки делает корпус пригодным и удобным как для учебных целей, так и для различного рода исследований, позволяет сократить затраты времени на обработку языкового материала. В НКРЯ есть пять видов разметки: «метатекстовая, морфологическая (словоизменительная), синтаксическая, акцентная и семантическая» [3]. Остановимся на семантической разметке. Семантическая разметка предназначена для «выявления семантических категорий, к которым относится данное слово / словосочетание, и более узких подкатегорий, специфицирующих его значение», она «предусматривает спецификацию значения слов, расширение омонимии и синонимии, категоризацию слов, выделение тематических классов, признаков каузативности, оценочных и деривационных характеристик и т. д.» [4, с. 45-48]. НКРЯ позволяет осуществлять поиск по лексико-семантическим признакам: лица, животные, растения, вещества и материалы и т. д. Следует

отметить, что в НКРЯ «используется фасетная классификация, при которой одно слово может попадать в несколько классов» [3]. Поиск по лексико-семантическим признакам возможен благодаря наличию в корпусе частичной семантической разметке текстов. Данная разметка «осуществляется автоматически с помощью программы Semmarkup» [3]. В основе данной программы лежит Семантический словарь корпуса, включающий в себя существительные, прилагательные, числительные, местоимения, глаголы и наречия. Слова в семантически размеченных текстах снабжены следующей лексико-семантической информацией: разряд, собственно лексико-семантические характеристики, деривационные характеристики. Следует отметить, что слова различных частей речи имеют свою структуру лексико-семантических помет. Например, лексема *хлеб* снабжена следующими пометами: S – имя существительное; r: consr – предметное имя; t: food – еда и напитки (таксономия), pt: aggг – множества и совокупности объектов, а имя *Аннушка* – S – имя существительное; persn – имя.

Благодаря возможностям семантической разметки можно ограничить круг контекстов с необходимой лексемой: отобрать необходимую группу по принадлежности к части речи (предметные имена, непердметные имена, имена собственные, прилагательные, числительные и т. д.), семантический класс (предметные имена: лица (этнонимы, родовые имена, сверхъестественные существа), животные, растения и т. д.; непердметные имена: движение, помещение объекта, существование, речь, звук и т. д.

Таким образом, семантическая разметка позволяет более точно формулировать запрос при поиске в НКРЯ, что делает корпус прекрасным инструментом для решения различных учебных и исследовательских задач.

Литература:

1. Кустова, Г. И. Семантические фильтры для разрешения многозначности в Национальном корпусе русского языка: глаголы / Г. И. Кустова, С. Ю. Толдова // Компьютерная лингвистика и интеллектуальные технологии: труды междунар. конф. «Диалог-2008». – М, 2008. – С. 522–529.

2. Малый академический словарь : в 4 т. [Электронный ресурс]. – Режим доступа: <https://rus-academic-dict.slovaronline.com/>. – Дата доступа: 18.10.2020.

3. Национальный корпус русского языка [Электронный ресурс]. – Режим доступа: <https://ruscorpora.ru>. – Дата доступа: 14.10.2020.

4. Рычкова, Л. В. Компьютерные языковые ресурсы на материале специальной лексики белорусского языка : моногр. / Л. В. Рычкова [и др.] ; под науч. ред. Л. В. Рычковой. – Гродно : ГрГУ, 2013. – 185 с.