

СИСТЕМА ЗАДАНИЙ ПО РАБОТЕ С НКРЯ ДЛЯ СТУДЕНТОВ, ИЗУЧАЮЩИХ РУССКИЙ КАК ИНОСТРАННЫЙ

Национальный корпус русского языка (НКРЯ) [1] — постоянно пополняемый представительный лингвистический ресурс с интуитивно понятным интерфейсом, круг потенциальных пользователей которого включает, помимо профессиональных лингвистов, и тех, кто изучает русский как иностранный (РКИ). Перспективные направления использования НКРЯ в преподавании РКИ определены в работе Н.Р. Добрушиной [2] (там же дана библиография работ по этой теме); опыт создания интерактивных ресурсов по РКИ на базе НКРЯ описан в [3].

В данной работе приведены комментированные примеры заданий, выполняемых при изучении темы «Расширенный поиск в НКРЯ» в рамках дисциплин «Основы информационных технологий» (для филологов) и «Информационные технологии в филологии» в группах студентов, изучающих РКИ. Цель, преследуемая в этом случае, комплексная: обучить студентов работе с НКРЯ как ресурсом и показать возможности НКРЯ для изучения РКИ.

Теоретический блок, предваряющий работу на практических, содержит сведения о синтаксисе запросов к НКРЯ. Краткая справка (в виде комментированных запросов-образцов) выдается студентам для самостоятельного изучения.

Запросы-образцы содержат небольшое число простых запросов, и (преимущественно) запросы расширенные. Простой запрос определяется выбором значений параметров групп «Грамматические признаки», «Семантические признаки», «Дополнительные признаки» и / или заполнением поля «Слово». Расширенный запрос предполагает использование масок и операторов. В масках, помимо буквенной части, используется знак * со значением '0 и более символов'. Запись поискового термина в кавычках используется для поиска указанной словоформы. Используются обозначенный знаком минуса оператор «не» и обозначенный вертикальным слешем оператор «или». Маски и операторы могут быть использованы не только при составлении запроса, но и при определении подкорпуса (последнее не выносится на изучение). Подробно о расширенных запросах см. [4].

В качестве зоны поиска на занятиях используется основной корпус; при поиске трансформаций фразеологизмов — основной и газетный корпусы. Заметим, что при работе с НКРЯ в группах с уровнем владения языком ниже порогового (о чем не идет речь в данной статье) допустимо использовать только обучающий корпус, содержащий тексты, нормализованные по орфографии, а в области морфологии — размеченные в соответствии с нормативными учебными пособиями. Обучающий корпус эффективно применим для выполнения так называемых «пропусковых» упражнений и упражнений на раскрытие скобок (примеры от разработчиков см. в [5]).

Типичными ошибками при составлении НКРЯ-запросов являются:

- 1) помещение в поле «Слово» незакавыченной косвенной словоформы;
- 2) помещение в поле «Слово» неоднословной конструкции (например, фразеологизма или фрагмента фразы);
- 3) использование неверного синтаксиса вида **x**; ср. верный синтаксис: *x** (например: *солн** для поиска контекстов со словами *солнце, солнцестояние, солнечный* и т.п.); **x* (например, **логический* для поиска контекстов со словами *логический, психологический* и т.п.); *x₁*x₂* (например, *по*ски* для поиска контекстов со словами *по-русски, по-человечески* и т.п.).

Для предотвращения типичных ошибок и проверки усвоения теоретического материала в начале практического блока предусмотрен разбор запросов-примеров: необходимо найти ошибочные и корректные запросы, для ошибочных — прокомментировать ошибку, для корректных — по возможности описать ожидаемую выдачу.

Ниже приведены примеры решаемых на практике заданий, сгруппированные по темам.

А. Исторические изменения семантики

Разбор таких заданий, чрезвычайно интересный и полезный для носителей языка, является крайне сложной задачей для изучающих РКИ и уместен скорее на факультативе. Приведем несколько видов упражнений (по возрастанию сложности): 1) даны ЛЗ слова и список контекстов, задача — сопоставить ЛЗ слова контекстам; 2) даны ЛЗ слова (в том числе и устаревшие варианты), задача — подобрать в НКРЯ контексты на каждое ЛЗ; 3) дан фрагмент научно-популярной статьи по этимологии слова (например, из [6, 7]), задача — подобрать в НКРЯ контексты на устаревшее ЛЗ слова по выбору; 4) исследовать изменения в семантике слова, данного преподавателем (см. примеры корпусных исследований семантики по НКРЯ в [8]).

В упрощенной форме (см. ниже) такие задания могут быть даны как домашние.

ЗАДАНИЕ. Что означает слово *мотор* в предложении *Бери мотор и езжай по этому адресу* (С. Довлатов)? При необходимости обратитесь к «Большому толковому словарю» на сайте Грамота.ру. Найдите в основном корпусе НКРЯ другие примеры использования слова *мотор* в этом значении. Подсказка: создайте запрос на поиск сочетаний *брать / взять мотор, на моторе* (по аналогии с *на такси*).

Комментарий. Запросы-решения таковы:

А-1) **брать | взять - bmark**

на расстоянии **1** от **мотор**

А-2) **на - bmark**

на расстоянии **1** от **"моторе"**

Запросы содержат «отминусованный» тег **bmark**, выбираемый в группе «Дополнительные признаки» и обозначающий ‘слово перед любым знаком препинания’; так определяется бóльшая синтаксическая связность компонентов.

Выдача содержит: контексты с искомым значением; контексты, где слово *мотор* имеет значение ‘двигатель’; контексты с не зафиксированным в БТС значением ‘мотолодка’ (*Но и эта трудность была с успехом преодолена: паводковая вода позволила подобраться к объектам работ на моторе* (Д.И. Саврасов) и нек. др.).

Б. Культурно-специфичное употребление слов

ЗАДАНИЕ. В чем особенность использования слова *бабушка* в контексте *Давайте помогу Вам, бабушка!* (при обращении к пожилой женщине-неродственнице)? Какие из следующих слов могут аналогично функционировать: а) *отец, папа, папаша*; б) *дедушка, прадедушка, прабабушка*; в) *брат, сестра, племянник*. Подсказка: вспомните, каковы формальные приметы конструкции обращения.

Комментарий. В поле «Слово» с использованием оператора «или» вводится исследуемая тройка слов, например: **отец | папа | папаша** (!определение наименований лиц по родству через тег в группе «Семантические признаки» дает сильный шум в выдаче).

Для определения формальных признаков обращения, находящегося в центре предложения, в группе «Грамматические признаки» выставляется тег **nom** (‘именительный падеж’), в группе «Дополнительные признаки» выставляются теги **bcomma**, **acomma** (‘слово перед запятой’, ‘слово после запятой’).

Для группировки выдачи по опорному слову используется доступная в интерфейсе НКРЯ сортировка: Настройки — Упорядочить по левому / правому контексту учитывая форму искомого слова.

В. Фразеология, сочетаемость

ЗАДАНИЕ. Пользуясь основным и газетным корпусами, найдите трансформации фразеологизма *метод **научного тыка*** (позиция трансформации выделена).

Комментарий. Запрос-решение:

метод - bmark

на расстоянии 1 от - **научный - bmark**

на расстоянии 1 от «**тыка**»

На позиции трансформации будут найдены слова *слепой, случайный, среднепотолочный*. Далее задание можно расширить за счет изучения контекстов с сочетаниями *слепой случай, взять с потолка, взять с неба*.

ЗАДАНИЕ. Пользуясь основным и газетным корпусами, составьте предложения с конструкциями вида *кроткий, как + название животного*.

Запрос-решение:

кроткий bcomma

на расстоянии 1 от **как**

на расстоянии 1 от **S r:concr & t:animal**

При создании запроса выставляются теги: **bcomma** (‘слово перед запятой’) в группе «Дополнительные признаки», **t:animal** (‘животные’) в группе «Семантические признаки».

Г. Морфология

ЗАДАНИЕ. Пользуясь основным корпусом со снятой омонимией, выясните, какой частью речи могут быть слова, в начальной форме заканчивающиеся на *-чь*.

Комментарий. Финаль исследуется прямым перебором значений частей речи в группе «Грамматические признаки» либо серией запросов на исключение: **чь -V -S* (исключаем из выдачи глаголы и существительные), **чь -V -S -ADV* (исключаем из выдачи глаголы, существительные, наречия) и т.д.

ЗАДАНИЕ. Пользуясь Основным корпусом НКРЯ, подберите к словам *держатель*, *вычислитель* примеры, где это слово обозначает: а) человека по роду занятий; б) аппарат или устройство.

Комментарий. Морфологической меткой слов групп (а) и (б) будет значение одушевленности / неодушевленности (теги *anim* и *inan* группы «Грамматические признаки»); однако подбор контекстов по корпусу с неснятой омонимией является достаточно сложной задачей для изучающих РКИ.

Д. Словообразование

ЗАДАНИЕ. Пользуясь НКРЯ, найдите слова, в начальной форме завершающихся на *...логический* (!но не слово *логический*).

Комментарий. Запрос-решение: **логический -логический*.

ЗАДАНИЕ. Обратившись к корпусу, выясните, образуются ли существительные, обозначающие лица, от слов *отважный*, *храбрый*, *смелый*. Субстантиваты не учитывать.

Комментарий. Задание решается серией из трех запросов, где в поле «Слово» вводится строка вида *основа прилагательного**, в группе «Грамматические признаки» вводится тег *S* («существительное»), в группе «Семантические признаки» вводится тег *t:hum* («лица»).

В заключение заметим, что студенты должны понимать: неснятая омонимия в корпусе является его свойством, а не ошибкой. Осознанное обращение к корпусу как справочно-информационному ресурсу позволяет сделать изучение РКИ более эффективным.

ЛИТЕРАТУРА

1. Национальный корпус русского языка [Электронный ресурс]. — М., 2003–2017. — Режим доступа: <http://ruscorpora.ru/>. — Дата доступа: 12.03.2017.
2. Добрушина, Н.Р. Корпусные методики обучения русскому языку // Национальный корпус русского языка: 2006–2008. Новые результаты и перспективы. — СПб., 2009. — С. 335–352.
3. Сибирцева, В.Г., Хоменко, А.Ю., Баранова, Ю.Н. Национальный корпус русского языка как основа новаторских электронных учебников // Образовательные технологии и общество. — 2013. — Вып. 3. — Т. 16. — С. 508–520.
4. Станкевич, А.Ю. *Расширенный поиск в корпусе региональной и зарубежной прессы НКРЯ // Контрастивные исследования и прикладная лингвистика: материалы междунар. науч. конф., Минск, 29–30 окт. 2014 г.: В 2 ч. / отв. ред.: А.В. Зубов, Т.П. Карпилович.* — Минск, 2014. — Ч. 2. — С. 145–149.
5. Савчук, С.О., Сичинава, Д.В. Обучающий корпус русского языка и его использование в преподавательской практике // Национальный корпус русского языка: 2006–2008. Новые результаты и перспективы. — СПб., 2009. — С. 317–334.

6. Голь, Н.М. Жизнь замечательных слов, или Беллетризованная этимологическая малая энциклопедия. — М., 2015.
7. Гребенщикова, Н.С. Историко-лингвистический комментарий на уроках словесности: пособ. для учителей рус. яз. и лит. — Минск, 1998.
8. Два века в двадцати словах / М.К. Данова, Н.Р. Добрушина, А.С. Опачанова и др.; отв. ред. Н.Р. Добрушина, М.А. Даниэль; НИУ ВШЭ. — М., 2016.