

С. О. Савчук,
Д. В. Сичинава

Обучающий корпус русского языка и его использование в преподавательской практике



удобность в создании Обучающего корпуса русского языка («ОК, русский язык!», или окря) возникла уже спустя год-два после возникновения

«большого» Национального корпуса русского языка. Дело в том, что 2005–2007 годы стали этапом решающего «прорыва» и даже, можно сказать, «бума» в использовании Корпуса как средства обучения русскому языку в высшей и средней школе. Впечатляющие творческие результаты в самых разных областях обучения русскому языку изложены в статье [Добрушина 2005], где данная проблематика, насколько мы понимаем, была представлена научному и образовательному сообществу впервые. При помощи Национального корпуса можно буквально в мгновение ока составлять самые разные задания для студентов, во многих отношениях превосходящие устаревшие и по необходимости «высосанные из пальца» упражнения и примеры современных пособий. Как употребляется в современной журналистике архаичная лексика? А чем одушев-

лённый *персонаж* отличается от неодушевлённого? Мы не говорим уже о создании простейших упражнений типа «расставить запятые» или «подобрать однокоренные слова». Оказывается, что при помощи нашего Корпуса можно делать даже экспертизу существующих учебных пособий, например, найти ответ на вопрос о том, надо ли заставлять школьников учить правило расстановки запятых в такой-то конструкции или не надо (поскольку такая конструкция в современных текстах просто не встречается ни разу). А это всего лишь немногие примеры из статьи Н. Р. Добрушиной.

В 2007 году в Высшей школе экономики состоялась конференция «Национальный корпус русского языка и проблемы гуманитарного образования», был проведён методический семинар для учителей из регионов. По итогам конференции был издан одноименный сборник¹, демонстрирующий самые разные подходы к корпусу как средству образования: здесь и изучение иностранными студентами тонкостей употребления русских слов *зачем* и *почему* [Бьяджини 2007], и преподавание риторики [Левинзон 2007], и анализ концептов-мифологем на университетских занятиях [Абьякая 2007], и обучение школьника деловому стилю [Кыркунова 2007] — в общем, многое такое, о чём мы при создании Корпуса специально не думали или вообще не догадывались.

При этом выяснилось, что создание для обучающихся целей подкорпуса текстов со снятой омонимией, специальным образом отобранного и отредактированного, выглядит самостоятельной насущной задачей (разумеется, не отменяющей работы учителя и ученика также с «большим» НКРЯ, а для ряда сложных заданий в вузах — и требующей её).

При всех своих достоинствах Национальный корпус русского языка не всегда может безоговорочно использоваться в качестве корпуса обучающего, особенно ориентированного на среднюю общеобразовательную школу. Причин тому несколько. Национальный корпус русского языка создавался с ориентацией на решение самого широкого круга задач. Предполагалось и предполагается, что к нему могут обращаться все, кому интересны те или иные явления, связанные с языком — лингвисты, литературоведы, про-

¹ Национальный корпус русского языка и проблемы гуманитарного образования. М., ТЕИС, 2007.

граммисты, работающие над автоматическим анализом естественного языка, редакторы и другие. Режим подачи текстов и разметки был соответственно разработан максимально «демократический», не ограничивающий какую-то группу пользователей в решении своих задач.

Так, нормализация орфографии и особенно пунктуации в Национальном корпусе минимальна: тексты сохранены в виде, наиболее близком к тому, в котором они функционируют реально, что может представлять большую ценность для лингвиста (редактора и т. п.), интересующегося не кодифицированной нормой, а реальным узусом, частотностью тех или иных вариантов и т. п. Понятно, что такой подход неприемлем для обучающего корпуса, где орфография и пунктуация должна быть приближена к норме, изучаемой в общеобразовательной школе.

Далее, значительный объём Национального корпуса диктует то, что омонимия слов с различными грамматическими разборами (попадающиеся на каждом шагу случаи вроде *печь* — существительное и *печь* — глагол, *большой* — именительный и винительный падеж мужского рода, родительный, дательный, творительный и предложный женского рода) снята (автоматически с ручной коррекцией) лишь в нескольких процентах текстов. Изучение языка в непрофильном институте и тем более в школе — это прежде всего изучение ригористической литературной грамматической нормы и «школьной» грамматической модели (в последние годы размывающейся в разнообразных пособиях и у разных преподавателей, но в основных чертах довольно устойчивой). Школьного и даже вузовского преподавателя могут смутить и затруднить такие черты нкря, как неснятая омонимия в большинстве текстов, отсутствие помет словоизменительных категорий (первое, второе и третье склонение), а в других случаях — чрезмерная для нужд непрофессионала «дробность» частеречных и словоизменительных помет. В Обучающем корпусе (имеющем сравнительно небольшой объём) таких решений, которые могут поставить ученика и учителя в тупик, быть не должно; во всяком случае, без труда разграничиваемые случаи вроде вышеприведённого должны быть приведены к однозначному разбору. Наконец не очень подходит для Обучающего корпуса столь же предельный «демократизм», принятый в подборе текстов для

корпуса; тексты, предлагаемые для обучения в школе, нуждаются в более тщательном отборе с дополнительным привлечением нескольких специфических критериев. С точки зрения состава текстов имеет, в частности, смысл поиск баланса между произведениями школьной программы по литературе, научными и художественными текстами, жанры которых изучаются в современной школе.

Созданный согласно этим задачам Обучающий корпус русского языка с ноября 2007 г. доступен в Интернете по адресу: <http://ruscorpora.ru/search-school.html>.

Ниже мы изложим ряд конкретных задач, появившихся при создании Обучающего корпуса, и принятых в связи с этим решений.

1. Состав корпуса

1.1. Сбалансированность и проблема определения функционального стиля

Поскольку школьный курс русского языка предполагает обучение учащегося всем функциональным разновидностям русской речи, корпус должен включать в себя тексты, принадлежащие к различным сферам употребления языка: художественная литература, публицистика, деловая, научная, разговорная речь. Таким образом, возможна постановка такой задачи, как анализ употребления слов, словосочетаний, грамматических форм в том или ином функциональном стиле русского языка.

На данном этапе имеется методическая проблема, связанная с отнесением конкретного текста к тому или иному функциональному стилю (например, интервью учёного по разным формальным признакам, в том числе и упоминаемым в педагогической литературе, может быть отнесено и к учебно-научной, и к публицистической, и к устной речи).

Кроме того, очевидно, что тот принцип пропорционального балансирования текстов различных жанров, который принят в Национальном корпусе русского языка, — где, например, публицистические и научные тексты заметно превосходят художественные, — прямо на обучающий корпус перенести нельзя. «Литературоцентричность» школьного обучения продолжает оставаться выраженной, и объём художественных текстов должен составлять всё же не менее двух

третьей от корпуса. Критерий литературного регистра, о котором ниже, препятствует и широкому включению в корпус записей устной речи. При этом публицистика и учебно-научные тексты в корпусе представлены довольно широко и разнообразно. Специфически методическую роль играют критико-филологические тексты, принимая во внимание роль школьного курса словесности в подготовке к написанию сочинения, — это классики критического жанра, такие как Белинский и Гончаров, и современные филологи. Но в корпус включаются, разумеется, также и иные научные тексты, а также образцы официально-делового стиля. Навыки в создании такого рода текстов являются первоочередными для каждого грамотного человека. Ср. опыт обучения деловому стилю при помощи Национального корпуса, представленный в [Кыркунова 2007].

Соображения баланса указывают, что тексты одного автора или схожей тематики не должны занимать в корпусе более 3% от его общего объёма (в данном случае—20 тысяч слов от 630 тысяч совокупного объёма). Большие по объёму тексты было решено включить не целиком, а в виде фрагментов—нескольких глав, частей.

1.2. Проблема просторечия и нестандартных форм

Так как школьные и вузовские курсы представляют собой курс русского литературного (стандартного в зарубежном понимании) языка, то корпус должен включать в себя лучшие образцы русской литературной речи, а также нейтральные образцы разговорной речи (в том числе письменной, такие, как письмо и дневник). Тексты, активно использующие просторечие, при включении в корпус должны избегаться, однако практика показывает, что разумно делать исключения для таких культурно значимых текстов, как, например, «Недоросль» Фонвизина, где просторечие используется для создания речевой маски некоторых героев. Это же относится и к встречающимся в литературе XVIII–XX вв. устаревшим грамматическим формам (которые даются со специальной пометой «аномальная форма»).

1.3. Подбор текстов и школьная программа

Выбор образцовых художественных текстов сам по себе представляет сложную и довольно субъективную проблему, но примени-

тельно к школьному преподаванию эта канонизирующая роль уже выполнена существующей школьной программой, которая задаёт хрестоматийный канон как имён, так и текстов. Подкорпус художественной литературы в Обучающем корпусе включает в себя преимущественно те произведения, которые входят в школьный курс и/или в программу для абитуриентов по русской литературе XVIII–XX вв., или (в небольшом числе случаев) другие произведения тех же авторов. Для общеязыкового фона добавлены также в небольших количествах произведения ряда авторов, не входящих в школьную программу, но язык которых показателен для русского языка середины—второй половины XX в. (И. Грекова, Татьяна Толстая и другие).

1.4. Подбор текстов и орфографическая выверенность

Тексты, входящие в корпус, исправлены с точки зрения соответствия их орфографических и пунктуационных характеристик требованиям русской орфографии и пунктуации. Нужно отметить, что реальные тексты, включая даже тексты художественной литературы, в ряде случаев от этого идеала отклоняются. Особенно это относится к текстам XIX–XX в., если речь идёт о художественной литературе (хорошо известная каждому школьному учителю проблема «авторской орфографии»); здесь мы просто следуем устоявшейся практике школьных изданий, в большинстве случаев нормализующей орфографию, а в ряде случаев и пунктуацию классиков. Что касается бытовых текстов, текстов публицистики (восходящих к электронным версиям газет) и иных нехудожественных жанров, то здесь составители корпуса по мере обнаружения недочётов выполняют функцию корректора (там, где публикации не было или же она оказалась почему-либо неудовлетворительной с точки зрения целей Обучающего корпуса).

1.5. Состав и структура Обучающего корпуса

На ноябрь 2007 г. корпус содержит 216 текстов совокупным объёмом 630 тысяч словоупотреблений. Распределение текстов по основным разделам (соответствующим одной или нескольким функциональным сферам текста) таково:

| | | |
|---|---------|------|
| Художественные тексты | 385 259 | 61% |
| Официально-деловые тексты | 24 658 | 4% |
| Публицистические тексты (включая ме- муары) | 143 133 | 23% |
| Разговорные тексты (устная публичная, непубличная речь, бытовая переписка) | 25 032 | 4% |
| Учебно-научные тексты | 52 162 | 8% |
| Всего | 630 244 | 100% |

В работе над определением состава Корпуса принимали участие Е. В. Рахилина, С. О. Савчук и Д. В. Сичинава. Консультацию и экспертизу осуществляли Н. Р. Добрушина и А. И. Левинзон.

1.5.1. Художественные произведения

Среди художественных произведений на ноябрь 2007 г. представлены тексты следующих авторов:

XVIII век—Д. И. Фонвизин.

XIX век—С. Т. Аксаков, В. М. Гаршин, А. И. Герцен, Н. В. Гоголь, И. А. Гончаров, Ф. М. Достоевский, М. Ю. Лермонтов, Н. С. Лесков, Д. Н. Мамин-Сибиряк, А. С. Пушкин, М. Е. Салтыков-Щедрин, Л. Н. Толстой, И. С. Тургенев, А. П. Чехов.

XX век—В. П. Аксенов, Л. Н. Андреев, Г. Я. Бакланов, М. А. Булгаков, И. А. Бунин, Г. Газданов, И. Грекова, Ю. О. Домбровский, Э. Г. Казакевич, В. Г. Короленко, А. И. Куприн, В. Г. Распутин, Т. Н. Толстая, В. Т. Шаламов, В. М. Шукшин.

Представлены следующие **типы** текстов: роман, повесть, рассказ, сказка, очерк, киноповесть, пьеса.

Хронотоп (место и время происходящих событий) большинства произведений связан с Россией соответствующей эпохи, однако представлен и фантастический мир (в сказке), и европейский хронотоп разных веков («Ася» Тургенева, «Обезьяна...» Домбровского).

Преобладают произведения, выдержанные в нейтральном стиле своей эпохи. Представлены также немногочисленные тексты с тем или иным отпечатком индивидуально-авторского стиля (Н. С. Лесков, Т. Н. Толстая) либо просторечия (Д. И. Фонвизин, В. М. Шукшин).

1.5.2. Публицистика

Отдел публицистики включает в себя, помимо публикаций и заметок из прессы XX—начала XXI в., также критические тексты В. Г. Белинского и И. А. Гончарова, очерково-мемуарную прозу А. Ф. Кони, Даниила Гранина, Ирины Архиповой, Игоря Кио.

Основной принцип отбора текстов заключался в том, чтобы как можно полнее представить палитру современной прессы. Поэтому в корпус включены тексты как центральных газет и журналов («Известия», «Труд», «Аргументы и факты», «Итоги», «Еженедельный журнал», «Октябрь», «Юность»), так и местных («Вечерняя Москва», «Московский комсомолец», «Время МН», «Новороссийский рабочий», «МК в Саранске»). Наряду с периодикой общественно-политической тематики включены специализированные издания, ориентированные на определенный круг тем—например, «Computerworld» (компьютеры); «Бизнес-журнал» (бизнес, экономика); «Поиск» и «Знание—сила» (наука); «Экран и сцена», «Народное творчество» (искусство); «100% здоровья» (медицина и здоровье); «За рулем», «Автопилот» (автомобили); «Домовой», «Сад своими руками» (досуг, домоводство). При отборе текстов учитывалась и целевая аудитория изданий: в частности, в корпус включены тексты из детских, молодежных («Мурзилка», «Трамвай», «Столица») и женских журналов («Даша», «Лиза»).

Жанровый состав текстов разнообразен и отражает требования школьной программы: новостное сообщение, статья, интервью, репортаж, очерк, отчет, рецензия, эссе. Разнообразие типов изданий, собранных в корпусе, позволит получить представление о модификациях основных жанровых типов в зависимости от тематики и направленности издания (например, интервью в деловой газете, в журнале для семейного досуга и в молодежном журнале при соблюдении общих принципов будут отличаться по отбору тем, стратегии диалога и составу языковых средств).

1.5.3. Учебно-научные тексты

Раздел учебно-научных текстов включает как предназначенные для квалифицированного читателя статьи из научных журналов, так и (в большинстве) научно-популярные статьи и публикации из

такой прессы, как «Знание—сила», «Первое сентября» и др. Представлены следующие типы текста: аннотация к научной статье, рецензия, статья.

Наряду с академическими собственно научными текстами (научная статья, рецензия, аннотация, реферат) включены заметки, интервью из научно-популярных журналов. Среди авторов учебно-научных статей и интервью—филологи М. М. Бахтин, Ю. М. Лотман, психолог А. Н. Леонтьев, математик В. А. Успенский.

1.5.4. Разговорные тексты

В этом небольшом разделе объединены такие тексты, как записи устной речи (интервью, дискуссия, беседа, телефонный разговор), а также непринуждённые письменные тексты, относящиеся к бытовой сфере (дневники и письма).

1.5.5. Официально-деловые тексты

В данном разделе представлены подлинные законы, договоры, характеристики, деловые письма, доверенности и некоторые другие типы текстов, характерные для официально-деловой письменности. Документы относятся к разным сторонам жизни—общественно-политической, бизнесу, праву, науке. Деловая речь представлена всеми жанрами, изучаемыми в вузе и школе: закон, договор, приказ, протокол, автобиография, резюме, заявление, расписка, деловое письмо и пр.

2. Усовершенствование морфологической разметки

Принципы адаптации морфологической лингвистической разметки к стандартным нормативным пособиям по русскому языку.

Как уже было сказано выше, Обучающий корпус базируется на Национальном корпусе русского языка. Поскольку Национальный корпус ориентируется прежде всего на интересы профессиональных лингвистов, использующих современную научную терминологию, то эта внутренняя разметка (морфологическая и семантическая) в Обучающем корпусе несколько упрощена и в целом адаптирована к требованиям школьной программы. При работе использовались следующие издания: Современный русский язык.

М., 2002; *Программа по русскому языку. 5–9-й классы*. Под ред. ак. РАО А. А. Леонтьева.

Основой лингвистической разметки в Корпусе является морфологическая разметка, указывающая значения всех основных грамматических категорий русских слов—т.е., в школьных терминах, дающая «грамматический разбор» каждого слова. Конечно, «грамматический разбор» в школьном понимании зачастую далек от современных научных представлений о том, как должны выглядеть грамматические характеристики исследуемого слова. Современная научная грамматика выделяет большее число падежей у русского существительного, чем это принято в школе, иначе описывает противопоставление залоговых форм у глагола, более сложным образом трактует проблемы грамматической омонимии и разных форм слова и т.п. Однако для нужд школы в той части Корпуса, которая ориентирована на задачи школьного преподавания и практическую помощь учителю-словеснику, должен присутствовать именно традиционный, упрощенный грамматический разбор.

Ещё одна особенность школьного преподавания—ориентация на орфографию. Во многом именно с необходимостью правильного написания слов и связан сам факт преподавания морфологии, «правописания приставок и суффиксов» в школе (при морфологическом принципе русского правописания это, действительно, неизбежная вещь). Знание о «первом и третьем склонении» полезно для того, чтобы не написать **в блокади* и **в тетраде*; о корне—для поиска «проверочных слов», о приставке—для написания *рас-* и *раз-*, о падеже—для верного написания согласуемых адъективов (выбор между, допустим, *-щем* и *-щим*). Иногда это приводит и к прямому упрощению сущностей: например, в школе заучивают список глаголов II спряжения с гласным инфинитива *а* или *е* (*знать, дышать, смотреть и видеть...*), но в этот список входят только такие глаголы, окончание которых оказывается без ударения и в написании которого возникает орфографическая проблема; над тем, что глаголы *кричать* или *скрипеть* также II спряжения, в школе, как правило, не задумываются, поскольку их окончание всегда под ударением (при создании Обучающего корпуса полезно обратить внимание школьников на это обстоятельство).

Отсюда вытекают следующие решения, принятые при составлении и обработке Обучающего корпуса русского языка.

Из Обучающего корпуса исключены некоторые дополнительные категории, используемые в современном научном описании русского языка: второй компаратив (на *по-*), счётная форма (*два часа́*), звательная форма, местоимение-предикатив (*некого, нечего*). В то же время предусмотрена система дополнительного поиска тех форм, которые охватываются данными пометами. Особенности использования системы дополнительного поиска разъяснены в инструкции для учителя.

С другой стороны, в Обучающий корпус внесены дополнительные пометы на морфологическую классификацию лексики, используемую в школе (см. следующий подраздел).

В Обучающем корпусе сводится к минимуму неоднозначность вида «краткие прилагательные—наречия—слова категории состояния (предикативные наречия)»; «союзы—союзные слова (местоимения, местоименные наречия)—частицы». Все такие места специально проверялись (прежде всего Г. И. Кустовой).

Расширенный морфологический стандарт.

Морфологическая разметка Корпуса для нужд окря была пополнена новыми параметрами.

Это прежде всего информация о типе словоизменения, включающая в себя:

- Стандартный инвентарь трёх склонений (I склонение *голова—головы*, II склонение *стол—стола* и III склонение *тетрадь—тетради*) пополнен также изучаемыми в школе особыми подтипами II склонения на *-ий, -ие* и I склонения на *-ия* (*санаторий, призвание, история*), разносклоняемыми словами с *-ен-*, а также особыми морфологическими типами, такими, как адъективное склонение («склоняется как прилагательное»: *военный, дежурная*), склонение фамилий, слова только единственного и только множественного числа.
- Типы спряжений: первое (*болеть*), второе (*гасить*), глаголы разноспрягаемые (*хотеть, бежать*) и неправильные (*есть, дать*). Слова (в том числе отсутствующие в ранее использовавшемся грамматическом словаре корпуса) были распределены по типам скло-

нений и спряжений программно при помощи эвристики, ориентирующейся на представленные в корпусе со снятой омонимией словоформы, с последующей ручной коррекцией получившегося грамматического словаря.

В корпус внесена семантическая информация о лексико-грамматических разрядах частей речи, изучаемых в школе:

- существительных (конкретные, абстрактные, вещественные, собирательные);
- прилагательных (относительные, притяжательные, качественные);
- местоимений (личные, возвратное и т. п.).

Для этой цели привлечён семантический словарь, использовавшийся при работе над Национальным корпусом русского языка (о нём подробнее см. [Кустова и др. 2005]), что потребовало в ряде случаев редактирования словаря и изменения программы наложения признаков.

В процессе работы над окря большую роль в усовершенствовании системы грамматических помет сыграли А. А. Аброскин, Н. В. Григорьев, Г. И. Кустова и О. Н. Ляшевская.

Дополнительная грамматическая разметка

Тексты, отобранные для Обучающего корпуса русского языка, подвергались морфологической разметке разработанными коллективом программными средствами, с последующим редактированием результатов морфологической аннотации (общие сведения о данном—едином для Корпуса—технологическом процессе, включая инструкцию разметчика, отражены в публикациях [Поляков 2005], [Ляшевская, Плунгян, Сичинава 2005]). Ряд задач по дополнительной разметке текста и словаря (определение спряжения и склонения слов, разрядов части слов) выполнен алгоритмически, дополнительная корректировка материала—вручную. Принципиальным моментом работы является отсутствие в Обучающем корпусе текстов с неснятой грамматической омонимией; двойной разбор допускается лишь в случае реальной неоднозначности трактовки той или иной словоформы.

Большинство текстов корпуса (за исключением небольшого количества художественных и официально-деловых текстов) прошли

полуавтоматическую акцентную разметку (расстановка ударений и буквы ё) на базе словаря с последующей ручной коррекцией омонимичных случаев. Акцентуации не подвергались имена собственные, аномальные формы (поэтому, как правило, не кодифицированные с орфоэпической точки зрения) и слова, отсутствующие в Грамматическом словаре А. А. Зализняка.

В работе над разметкой Обучающего корпуса русского языка принимали основное участие О. Л. Бирюк, Г. И. Кустова, Е. А. Пшехоцкая и Д. В. Сичинава.

3. СОЗДАНИЕ ИНТЕРФЕЙСА ПОЛЬЗОВАТЕЛЯ И ИНСТРУКЦИИ ПО ПОЛЬЗОВАНИЮ КОРПУСОМ

Указанные усовершенствования и изменения режима разметки получают поддержку на уровне поискового интерфейса корпуса. Обучающий корпус организован в виде отдельного модуля в составе Национального корпуса русского языка; для пользования им разработан самостоятельный поисковый интерфейс, учитывающий потребности преподавателей и учащихся и уровень их подготовки.

В целях облегчения работы с Обучающим корпусом русского языка создана (Е. А. Гришиной) инструкция для потенциальных пользователей корпуса, написанная в доступной и наглядной форме (в настоящее время она размещена по адресу <http://ruscorpora.ru/instruction-school.pdf>).

4. ПРИМЕРЫ ОБРАЗОВАТЕЛЬНЫХ ЗАДАЧ, РЕШАЕМЫХ ПРИ ПОМОЩИ ОБУЧАЮЩЕГО КОРПУСА

При помощи Обучающего корпуса русского языка можно строить упражнения различного типа, в том числе недоступные при помощи основного корпуса. Для создания упражнений достаточно провести минимальный отбор среди полученных в выдаче результатов; в ряде случаев наиболее эффективным оказывается поиск по сочетанию нескольких параметров (например, окончание слова на *еть + II склонение + III лицо). Вот несколько примеров.

1. Заполните пропуски (упражнение на типы склонения). При подготовке этого упражнения искали существительные второго склонения в предложном падеже; отдельно искался подтип скло-

нения на *-ий* и регулярный тип:

- По их *окончани...* найденные на полях сражений останки фронтовиков со всеми почестями будут преданы земле на новгородских воинских мемориалах.
- 17 декабря в саранском профессиональном *лице...* №21 состоялась научно-методическая конференция.
- Только что она в радостном *нетерпени...* вспоминала пальцами клавиши, переворачивала воображаемые ноты.
- В Европе в античные времена и в *Средневековье...* аконит был известен только как яд.
- Он отметил, что взрыв произошёл в северной *провинци...* КНДР Янгандо, рядом с границей с Китаем.
- Они проводили опыты в Нью-Йорке, в местном *дельфинари...*
- Родились эти люди в снах, вышли из снов и прочнейшим образом обосновались в моей *кель...*

2. Заполните пропуски (упражнение на типы спряжения). Вёлся поиск разноспрягаемых глаголов и глаголов второго спряжения на *-еть, -ать* как в отклоняющейся форме (2 и 3 лицо настоящего времени), так и в форме прошедшего времени, где гласный у этой формы регулярный:

- Ни ГК РФ, ни законодательство о страховании не *содерж...т* термин «характер ответственности».
- Обращается особое внимание на то, что на территории соседнего Китая никто не *вид...л* и не *слыш...л* никаких взрывов.
- Двигает рисунок ближе к себе, *верт...т* головой.
- Фестиваль очень важен, потому что даёт новые впечатления, зрители *вид...т* новые формы.
- У нас собрались мастерицы, прошедшие курсы лоскутного шитья, которые умеют и (*хотеть*) шить современную одежду в народном стиле, отталкиваясь от русской традиции.
- «Полированная»—это когда нормальные шкурки стригут, *бре...т*, полируют по спецтехнологии и потом уже шьют пальто.
- Все пришедшие просили у Филиппа Филипповича, все *завис...ли* от его ответа.
- Кстати, это отчасти объясняет большую устойчивость США

и Англии к популярным в своё время идеям экспроприации, потому что собственность в этих странах *выгляд...т* не как нечто единое, что можно забрать, а как запутанный социальный институт, в котором ничего невозможно сделать в один ход.

- Дано самое невинное и безотказное кокетство—во всем: в том, как сидит, как расправляет трен или *держ...т* чашку.
- Кузов кареты и бричка начинают подпрыгивать по неровной дороге, и берёзы большой аллеи одна за другой (*бежать*) мимо нас.
- Выбежал из лесу—до будки сто сажен, не больше, осталось, *слыш...т*—на фабрике гудок загудел.
- Так вот, в 86 году я закончил школу и покорно пошёл на физфак, потому что мама и папа сказали: «Режиссёров всех *разгон...т*, а на военном заводе кусок хлеба будет всегда!»
- Как правило, результат по судебному делу *завис...т* от множества факторов.

3. В каких примерах слова употреблены в конкретном, а в каких—в отвлечённом значении? Чтобы получить материал этого упражнения, нужно найти в корпусе лексемы, имеющие оба разбора—*concr, abstr*—семантическая неоднозначность не снята—а затем подобрать примеры из поиска на каждое слово:

- Работа поддержана РГНФ и Программой поддержки научных *школ*.
- В 2002 г. мексиканские исследователи Карлос Веласко Масиас и Мануэль Неварес-де-лос-Рейес обнаружили не просто новый, никем ранее не виданный и не описанный кактус—они нашли растение, невообразимо отличающееся от всех, до сих пор известных. Настоящую *химеру*, сочетающую отличительные признаки нескольких наиболее ярких представителей семейства.
- Наверное, и погода (до этого дня холодная) ждала дня, когда у нас в Тверской *области* возродится древний праздник—День Ивана Купала.
- Мне было важно показать, как совесть превращается в *химеру*.
- В 1937 году, в *школу*, где я учился, нагрянула съёмочная группа, набиравшая актёров, персонажей для фильма.
- Встаёт вопрос о необходимости более чёткого обоснования выбора типовых методик для получения реальной картины уровня

развития умственной работоспособности, следует также уточнить *область* применения методик и правил служебного пользования тестовым материалом психолога.

4. Найдите наречия причины и цели, подчеркните их по-разному (ср. [Бьяджини 2007] об обучении иностранцев схожей проблеме). Чтобы получить материал этого упражнения, нужно задать поиск «наречия причины» и «наречия цели»:

- Как это? Всем генпрокуратурам *назло*?
- *Недаром* бытует расхожее мнение, что критиками становятся неудавшиеся режиссёры, сценаристы и телеведущие.
- Мой папаша был мужик, идиот, ничего не понимал, меня не учил, а только бил *спьяна*, и всё палкой.
- Так будет правильно для домика, где нет места для светских разговоров, в котором *поневоле* оказываешься так близко к собеседнику, что ничего не значащие реплики превращаются...
- Дом Обломовых был когда-то богат и знаменит в своей стороне, но потом, бог знает *отчего*, все беднел, мельчал и наконец незаметно потерялся между не старыми дворянскими домами.
- Ведь *неспроста* во всех крупных театрах и концертных залах издавна существовали царские, теперь правительственные, ложи.
- — Ира? Ты *что* это дома сегодня? — Да я на работу не пошла, дома работаю.
- Ничего не было видно, как в подземелье, и Маргарита *невольно* уцепилась за плащ Азazelло, опасаясь споткнуться.
- Может быть, и бакенбардами своими он дорожил *потому*, что видел в детстве своём много старых слуг с этим старинным, аристократическим украшением.

5. Раскройте скобки (задание на правописание неопределённых местоимений). В корпусе ищутся неопределённые местоимения; для разрывно пишущихся—частица *кое*; для примеров с раздельным *не*, которые к неопределённым местоимениям не относятся—поиск «не + кто» и т. п. В примере *не какой-нибудь* эти запросы сочетаются; первая скобка не относится к правописанию сложных местоимений, а вторая относится.

- Кого теперь удивить «Волгой»; «самый быстроходный в мире трактор», с усмешкой говорят о ней, но тогда «Волга» кое(что) значила и кое(что) о хозяйне говорила.
- Враг этот не(кто) другой, как наш северный мороз, хотя, впрочем, и говорят, что он очень здоров.
- Без фокусов остаться с собою наедине—и спросить себя кое(о) (чём), не боясь огорчить правдивым ответом.
- И кое(с)(кем) начинал о чем(то) заговаривать, причём, сколько я понимаю, в тумане и дыму получал твёрдые отказы.
- Не(кто) Ивлев ехал однажды в начале июня в дальний край своего уезда.
- Он не(какой)(нибудь) мелкий исполнитель чужой, готовой мысли; он сам творец и сам исполнитель своих идей.

Таким образом, Обучающий корпус русского языка—новое средство, позволяющее, с одной стороны, «рафинировать» с точки зрения школьной программы и надёжности результатов возможности «большого» Национального корпуса русского языка, а с другой стороны, расширить их тем, что отсутствовало в прежней версии (это касается, прежде всего, новых возможностей морфологического поиска).

ЛИТЕРАТУРА

- Абьякая О. В. Актуализация признаков концептов-мифологем (на материале Национального корпуса русского языка) // Национальный корпус русского языка и проблемы гуманитарного образования: учебно-методическое пособие. М., 2007, с. 129–135.
- Бьяджины Ф. Составление упражнений на употребление *зачем* и *почему* для студентов-иностранцев при помощи *нкря* // Национальный корпус русского языка и проблемы гуманитарного образования: учебно-методическое пособие. М., 2007, с. 101–108.
- Добрушина Н. Р. Как использовать Национальный корпус русского языка в образовании? // *нкря* 2003–2005. Результаты и перспективы. М., 2005, с. 308–329.
- Кустова Г. И., Ляшевская О. Н., Падучева Е. В., Рахилина Е. В. Семантическая разметка лексики в Национальном корпусе русского языка: принципы, проблемы, перспективы // *нкря* 2003–2005. Результаты и перспективы. М., 2005, с. 155–174.
- Кыркунова Л. Г. Дидактико-методические аспекты использования «Национального корпуса русского языка» в речеведческих курсах (на примере курса «Деловой русский язык») // Национальный корпус русского языка и проблемы гуманитарного образования: учебно-методическое пособие. М., 2007, с. 192–200.
- Левинзон А. И. Использование *нкря* в преподавании дисциплины «риторика» в вузе // Национальный корпус русского языка и проблемы гуманитарного образования: учебно-методическое пособие. М., 2007, с. 160–173.
- Ляшевская О. Н., Плунгян В. А., Сичинава Д. В. О морфологическом стандарте Национального корпуса русского языка // *нкря* 2003–2005. Результаты и перспективы. М., 2005, с. 111–154.
- Поляков А. Е. Технология подготовки информации в Национальном корпусе русского языка // *нкря* 2003–2005. Результаты и перспективы. М., 2005, с. 175–192.