

Опубликовано в: Национальный корпус русского языка и проблемы гуманитарного образования. Материалы международной научной конференции. Москва 19-20 апреля 2007. М. 2007. С. 17-19.

## **Неснятая омонимия в текстах Национального корпуса русского языка как исследовательская база для студентов-филологов на практических занятиях по морфологии**

И.Т. Вепрева

Уральский государственный университет (г. Екатеринбург)

[irina\\_vepreva@mail.ru](mailto:irina_vepreva@mail.ru)

Всемирная компьютерная сеть стала привычным источником информации и средством общения для миллионов людей. Помимо коммуникативной и информационной функций, Интернет дает возможность использовать его и в целях обучения. Перед сегодняшними преподавателями русского языка стоит задача – найти способы эффективного использования Интернета на занятиях. Самостоятельный поиск в сети, с одной стороны, является одним из факторов мотивации в обучении, поддержки интереса к предмету, с другой стороны, превращается в своеобразный компьютерный ликбез, поскольку дает навык работы с различными системами компьютерной коммуникации (что приносит пользу как преподавателю, так и студенту-филологу, решившим осваивать Интернет вслед за «продвинутыми» университетскими, например математическими, собратями).

Появившийся Национальный Корпус русского языка (НКРЯ) представляет собой полезный инструмент, который может быть пущен в дело на занятиях по курсу «Современный русский язык» по многим разделам. Обратимся к разделу «Морфология» в рамках указанного курса. Безусловно, НКРЯ может участвовать в составлении и решении грамматических задач по каждой конкретной теме в традиционном поэтапном рассмотрении частей речи по грамматическим категориям. Для преподавателя крайне заманчива разметка корпуса текстов по лексико-семантическим группам, по грамматическим признакам, по частеречной принадлежности. Заданная авторами морфологическая и лексико-семантическая информация представляется преподавателю достаточно приемлемой для процесса обучения, поскольку не апеллирует к специальным сведениям, выходящим за рамки базового багажа студента филологического факультета, кажется доступной и понятной для поиска.

Но как всякий языковой корпус, объединяющий большой объем материала, НКРЯ страдает точностью обработки. Признание современной компьютерной лингвистики факта невозможности точного и безукоризненного автоматического анализа текста делает работу студента с базой данных как основой для проверки знаний практически невозможной. И в связи со сложившимися обстоятельствами возникает идея использования

НКРЯ как дополнительного источника материала для создания эвристических ситуаций на практических занятиях по морфологии. Такой исследовательской базой может служить неснятая авторами грамматическая неоднозначность омонимичных лексем.

Одной из практических задач обучения морфологии является формирование представления о частях речи как о сложной системе языковых единиц, как о классификации, построенной на совмещении нескольких критериев. Императив, сформулированный Л.В.Щербой в 1928 году в статье «О частях речи в русском языке», классифицировать слова по частям речи не по предвзятым принципам, а «разыскивать, какая классификация настойчиво навязывается самой языковой системой» вполне приложим в качестве исследовательской задачи на основе текстов НКРЯ с неснятой омонимией. Одной из важных практических задач является наблюдение над поведением грамматических омонимов в тексте с дальнейшей формулировкой условий, на основе которых возможно разграничение знаменательных и служебных слов; грамматических омонимов внутри служебных частей речи; предикативов, наречий и кратких форм прилагательных и т.д. Остаточный грамматический фон исходной части речи, сохраняясь, мешает дифференциации слов. Этот компонент порождает противоречивость характеристик омонимичной части речи.

Первый вариант поиска может быть связан с конкретной лексемой. Так, например, пользователи НКРЯ, задав необходимые параметры (например, лексема «и» в роли частицы), получают в ответ на запрос контексты, в которых не устранена грамматическая неоднозначность:

1) *Такая страна, как Россия, может жить и развиваться в существующих границах, только если она сильная держава, – подчеркнул Президент.*

2) *Дело не только в том, что у нас десятки малоизвестных политических партий и люди легко допускают, что существует и такая, как ПСИП.*

3) *И дали срок – до Рождества Христова.*

4) *Конечно, и цена отличная - 25 рублей*

5) *Ну и обманщица же ты*

6) *Наш экономический фундамент хотя и стал заметно прочней, но все еще слаб и неустойчив.*

Перед студентом ставится задача: на определенном количестве (оговоренном преподавателем) контекстов с выбранной единицей провести классификационную работу, выявить закономерности, которые могли бы в дальнейшем помочь разобраться в статусной характеристике омонимичной части речи. Материал для анализа, взятый не из учебника, а добытый в процессе самостоятельного поиска, приобретает особую значимость. В собранных контекстах студенту видны явные случаи употребления союза «и» в функции чистого сочинения (1–2), преподаватель может указать на случаи союзного употребления в предложениях эпического, повествовательного

характера для указания на связь с предшествующим (3), в качестве выделительного союза (2). На фоне активного употребления «и» в качестве союза выделяются контексты, в которых анализируемая единица выступает и как частица (4–5).

Выборка контекстов должна быть достаточно представительной по объему, такой, которая бы позволила выделить различные значения частицы «и», синонимичные частицам *ведь, же, именно, тоже, даже*. Кроме того, фразовый материал позволяет определить контекстные условия, в которых «и» проявляет себя в роли частицы.

Второй возможный вариант поиска ориентирован не на конкретную языковую единицу, а на поиски спорной (например, слова категории состояния) или, по определению Щербы, «туманной» категории слов (междометие). На запрос о предъявлении контекстов с предикативами, поисковая система выдает не только спорные случаи, но и случаи явно ошибочные, например.

1) *Или рассказывает о собственном приключении, как он бежал быстрее всех, но споткнулся – и бух!*

2) *Всякие **скок, топ, бух** и т.п. считают детскими сокращениями от «нормальных слов».*

3) *Калуга, как **известно**, маленький Константинополь.*

4) *Депутаты **против** Исаака Ньютона.*

5) *Исход борьбы стал тем более неясен, когда **против** разрушения этой бюрократической кормушки неожиданно выступили профессиональные объединения.*

6) *Такие пятна, однако, **можно** легко удалить, **надо** только знать как.*

Пользователи НКРЯ на практике убеждаются, что объем материала и точность аннотирования являются взаимоисключающими критериями. Но ошибочные и спорные случаи являются хорошей исследовательской площадкой в учебных условиях. При бесспорном и обоснованном выделении предикативов в (6) можно обсуждать частеречную деривацию окказионального типа в (1, 2, 4), явно ошибочные контексты с четкой частеречной принадлежностью (5).

Таким образом, работа над снятием грамматической неоднозначности ставит перед студентом настоящие «взрослые» задачи поиска критериев разграничения омонимичных частей речи, поиска оснований для выделения спорных частей речи (предикативов, междометий, модальных слов и т.д.), предлагает студентам на практике почувствовать вкус исследовательской работы.