

**МІНІСТЕРСТВО ОСВІТИ І НАУКИ, МОЛОДІ ТА СПОРТУ УКРАЇНИ
НАЦІОНАЛЬНА МЕТАЛУРГІЙНА АКАДЕМІЯ УКРАЇНИ**

Н.В. НЕЧУХАСВА

КОМП'ЮТЕРНА ЛЕКСИКОЛОГІЯ ТА ПЕРЕКЛАД
ДЛЯ ГРУП ФІ
(БАКАЛАВРАТ ТА СПЕЦІАЛІСТ)

Навчальний посібник

Дніпропетровськ НМетАУ
2013

ЛЕКЦИЯ 1

ЛЕКСИКОЛОГИЯ И ЛЕКСИКОГРАФИЯ

Лексикология – раздел языкознания, изучающий слово как единицу словарного состава языка (лексики) и всю лексическую систему (лексику) языка.

Лексикография – теория и практика составления словарей.

Слово как единица лексики. Значение слова

Слово – это языковая единица, которая служит для обозначения (наименования) предметов и признаков (действий, отношений, качеств, количеств).

Под **значением** слова понимают отображение в слове реалий действительности (предметов, признаков, отношений, процессов и т. д.). Основная функция слова – называние реалий действительности (так называемая номинативная функция). Но слово не только называет, но и обобщает: называя *столом* предметы разного размера, формы, назначения, мы отвлекаемся от индивидуальных особенностей конкретных предметов и сосредотачиваем внимание на их основных чертах, отражающих не все, а лишь существенные признаки реалии, позволяющие отличать одно явление от другого (например, предметы, называемые словом *стол*, от предметов, называемых словом *табурет*). В значение слова могут входить эмоционально-оценочные компоненты. Так, например, слова *книга* и *книжонка* называют один и тот же предмет, но второе слово обладает также дополнительным эмоционально-оценочным компонентом значения – выражает пренебрежительную оценку говорящего.

Значения слов фиксируются в толковых словарях.

Словарь – книга, содержащая перечень слов или других языковых единиц (морфем, словосочетаний, фразеологизмов), размещенных в определенном порядке, чаще всего алфавитном. Словари бывают энциклопедические и лингвистические.

Энциклопедические словари описывают сам объект внеязыковой действительности и стараются представить максимальное количество признаков этого объекта.

Лингвистические словари описывают языковые единицы. Лингвистические словари можно разделить на две группы: толковые, описывающие лексическое значение слов (попутно также отражающие его написание, ударение, часть речи, отдельные грамматические формы), и аспектные, описывающие слова с точки зрения их написания (орфографические), произношения (орфоэпические), морфемного состава (морфемные), словообразовательной производности (словообразовательные), грамматических форм (грамматические), происхождения (этимологические, иностранных слов), а также с точки зрения их взаимоотношений с другими словами (словари синонимов, антонимов, сочетаемости и др.).

Однозначные и многозначные слова. Прямое и переносное значения слова. Типы переносных значений

Слова бывают **однозначные** (имеющие одно лексическое значение) и **многозначные** (имеющие несколько значений).

Однозначными словами является большинство терминов, некоторые названия инструментов, профессий, разновидностей деревьев и пр. Однозначны, например, слова *табурет, сахарница, огромный, суффикс*.

Большое количество слов имеет несколько (два и более) значений. Так, например, для слова *голова* в Толковом словаре С. И. Ожегова приводится 6 значений:

- 1) часть тела человека или животного,
- 2) ум, рассудок (*Он человек с головой*),
- 3) человек как носитель идей (*Он голова!*),
- 4) передняя часть чего-либо (*голова колонны, поезда*),
- 5) единица счета скота (*стадо в 100 голов*),
- 6) пищевой продукт в форме шара, конуса (*голова сыра, сахара*).

Все значения многозначного слова связаны между собой (хоть иногда и не все сразу).

В многозначном слове выделяются главное (исходное, первичное) значение слова и производные от него значения. Новые значения возникают у слова в результате переноса наименования (внешней оболочки слова – звуковой и буквенной последовательности) с одного объекта действительности на другие объекты.

Существуют два типа переноса наименования: 1) по сходству (метафора), 2) по смежности – реальной связи объектов (метонимия).

Охарактеризуем эти типы переноса.

Перенос по сходству (метафора).

Чем-то похожие друг на друга объекты начинают называться одним словом.

Сходство между предметами может быть

- 1) внешнее:
 - а) форма: *лента дороги, пузатый чайник*;
 - б) цвет: *медные волосы, собирать лисички*;
 - в) расположение: *горло залива, цепь гор*;
 - г) размер, количество: *море слез, гора вещей*;
 - д) степень плотности: *стена дождя, кисель дорог*;
 - е) степень подвижности: *быстрый ум, машина ползет*;
 - ж) характер звучания: *дождь барабанит, скрипучий голос*;
- 2) функциональное: *дворники машины, брачные оковы*;
- 3) в восприятии человеком: *холодный взгляд, кислое выражение лица*,

Перенос по смежности (метонимия).

Два явления, реально связанных друг с другом (пространственно, ситуативно, логически и т. д.), получают одно наименование, называются одним словом. Связь явлений бывает:

- 1) пространственная – помещение и люди, находящиеся в нем: *класс опоздал, зал аплодировал*;

2) временная – действие и предмет – результат этого действия: *подарочное издание, набор инструментов*;

3) логическая:

а) действие и место этого действия: *вход, остановка*;

б) действие и люди, его производящие: *защита, нападение* (защитники, нападающие);

в) материал и изделие из этого материала: *носить золото, меха, выиграть золото, серебро, бронзу*;

г) автор и его произведения: *ставить Чехова, пользоваться Ожеговым*.

Разновидностью этого типа переноса является перенос наименования с части на целое и с целого на часть (синекдоха):

1) с части на целое: *в семье прибавился лишний рот* (= человек), *стадо в сто голов скота* (= животных), *номер в гостинице* (= комната), *первая скрипка, первая ракетка*;

2) с целого на часть: *соседи купили машину* (= автомобиль); но *машиной* называется и целый класс механизмов: стиральные, швейные и др. машины.

Степень образности и распространенности в языке у разных переносных по происхождению значений различна.

Часть переносных употреблений не используется нами в речи и существуют лишь в тексте определенного автора. Это индивидуальные (авторские) метафоры и метонимии: *медные кишочки автомобиля* (И. Ильф и Е. Петров), *скула яблока* (Ю. Олеся). Образность их максимальна, употребление связано лишь с авторским текстом; в словарях эти переносные значения не отражаются.

Другие метафоры и метонимии являются общеупотребительными, они не «привязаны» к определенному тексту и употребляются носителями языка в разных ситуациях (обычно в разговорной речи): *стена леса, море слез, звезда экрана, пилить* (ругать). Их образность меньше, чем у авторских, но отчетливо ощущается носителями языка; они отражаются в словаре и имеют помету *переносное*.

Третий тип переносных по происхождению значений характеризуется тем, что их образность не ощущается (в лингвистике их называют «сухими»): *ножка стула, шляпка гриба, гусеницы трактора, солнце зашло, часы идут*. Эти значения являются основными, словарными наименованиями определенного предмета, признака или действия: переносные по происхождению, они являются прямыми с точки зрения функционирования в современном языке и не имеют в словаре пометы *переносное*.

Таким образом, в слове может быть несколько прямых значений – исходное и не имеющие образности переносные по происхождению («сухие»).

Рассмотрим в качестве примера структуру значений приведенного выше слова *голова*: 1) часть тела человека или животного, 2) ум, рассудок (*Он человек с головой*), 3) человек как носитель идей (*Он голова!*), 4) передняя часть чего-либо (*голова колонны, поезда*), 5) единица счета скота (*стадо в 100 голов*), 6) пищевой продукт в форме шара, конуса (*голова сыра, сахара*).

Исходным и прямым у этого слова является значение ‘часть тела человека или животного’ (в лингвистике знаком ‘’ обозначается значение языковой единицы).

На основании сходства (метафорический перенос) образованы значения 4 и 6: ‘передняя

часть чего-либо' (сходство по расположению) и 'пищевой продукт в форме шара, конуса' (сходство по форме). Причем оба эти значения также являются прямыми и используются в нейтральном стиле речи. На основании смежности (метонимический перенос) на базе исходного значения образовано образное значение 2 'ум, рассудок'. На базе значения 2 образовано переносное значение 3 'человек как носитель идей' – название целого по части (синекдоха). На базе 1 исходного значения образовано прямое значение 5 'единица счета скота' – также перенос с части на целое (синекдоха). Таким образом, у слова *голова* 4 прямых значения, не имеющих образности, – значения 1, 4, 5, 6, и 2 образных – значения 2 и 3.

Как мы видим, понятие метафорического переноса в языкознании шире, чем в литературоведении, где такие языковые («сухие») метафоры, выполняющие только функцию называния и не обладающие образностью, не рассматриваются в качестве метафор.

Разные значения одного слова помещаются в одну словарную статью толкового словаря.

Омонимы

Омонимия – это совпадение в звучании и написании языковых единиц, значения которых не связаны друг с другом.

Основным видом омонимов являются **лексические омонимы** – слова одной и той же части речи, имеющие одинаковое звучание, написание и грамматическое оформление, но разное значение. Если между значениями многозначного слова прослеживаются смысловые связи, основанные на разных типах переноса наименования, то у омонимов значения не связаны между собой, у них нет общих смысловых компонентов (в отличие от разных значений многозначного слова). Омонимы являются разными словами.

Лексические омонимы могут иметь одинаковое или разное количество грамматических форм; первые называются полными, вторые – неполными. Так, например, полными омонимами являются слова *гриф* ('птица'), *гриф* ('печать') и *гриф* ('часть струнного инструмента'); *мешать* ('перемешивать') и *мешать* ('быть помехой'). Неполными омонимами являются слова *лук* 'оружие' (I) и *лук* 'растение' (II): у *лук I* есть формы и единственного, и множественного числа, у *лук II* формы множественного числа отсутствуют, но в единственном числе их формы полностью совпадают.

Возникновение омонимов может быть связано со случайным совпадением в языке исконно русского и заимствованного слова (*клуб дыма* – рус. и *клуб* 'общество' – англ.) или нескольких заимствований из разных языков (*фокус* 'трюк' – нем. и *фокус оптический* – лат.), а также с особенностями словообразования (*критический* от *кризис* и *критический* от *критика*) и с распадом многозначности (*слог* 'часть слова' и *слог* 'стиль').

В отличие от значений многозначного слова, которые помещаются в толковых словарях в одной словарной статье, омонимы, будучи разными словами, выделяются в разные словарные статьи.

От лексической омонимии следует отличать смежные с нею явления – фонетическую, грамматическую и графическую омонимию.

Фонетические омонимы (омофо́ны) – это слова, различно пишущиеся, но одинаково

произносящиеся (за счет редукции и оглушения/озвончения), например, *код – кот, пруд – прут, обессилеть – обессилить, пребывать – прибывать*.

ЛЕКЦИЯ 2

Что такое Национальный корпус языка?

Корпус – это собрание текстов в электронной форме, в котором можно осуществлять поиск слов, словосочетаний, грамматических форм, значений слов с помощью определенной поисковой системы. Существует много разновидностей корпусов – например, корпус одного автора, корпус одной книги (в частности, первые корпуса делались для Библии). **Национальный корпус** некоторого языка представляет данный язык на определенном этапе его существования и во всём много-образии жанров, стилей, территориальных и социальных вариантов и т. п. Национальный корпус русского языка функционирует в Интернете с 2003 года и в данный момент включает самые разные русские тексты общим объемом 140 миллионов словоупотреблений. Национальный корпус любого языка имеет две важные особенности:

Во-первых, он **представителен и сбалансирован**, т. е. содержит все типы письменных и устных текстов, представленных в данном языке (художественные разных жанров, публицистические, учебные, научные, деловые, разговорные, диалектные и т.п.), и все эти тексты входят в корпус по возможности пропорционально их доле в языке соответствующего периода.

Во-вторых, корпус содержит особую дополнительную информацию о свойствах входящих в него текстов (так называемую **разметку**, или **аннотацию**). Разметка — главная характеристика корпуса; она отличает корпус от простых коллекций (или «библиотек») текстов, в изобилии представленных в современном Интернете. Чем богаче и разнообразнее разметка, тем выше научная и учебная ценность корпуса. В Национальном корпусе русского языка в настоящее время широко используются три типа разметки: метатекстовая (она характеризует текст в целом, с точки зрения его автора, жанровых и других особенностей), морфологическая и семантическая (морфологическая и семантическая разметки характеризуют не целый текст, а отдельное слово).

Зачем нужен национальный корпус?

Национальный корпус предназначен в первую очередь для научных исследований лексики и грамматики языка. Другая задача корпуса — предоставление всевозможных справок, относящихся к указанным областям (лексика, грамматика, акцентология, история языка). Раньше специалист мог лишь просматривать тексты и вручную выписывать из них нужные примеры; эта деятельность была очень трудоемкой и не позволяла обрабатывать большие массивы материала. Теперь ограничений на объем анализируемого материала и скорость поиска информации нет, а это означает, что в распоряжении ученого и преподавателя оказываются колоссальные массивы текстов самого разного типа.

Основными пользователями национальных корпусов являются, конечно, исследователи-лингвисты самого разного профиля. Однако ими круг пользователей корпуса вовсе не ограничивается! Надежные статистические данные о языке определенной эпохи или определенного автора могут интересовать литературоведов, историков и представителей многих других областей гуманитарного знания. Важнейшее значение национальные корпуса имеют и для преподавания языка

в качестве родного или иностранного; всё больше учебников и учебных программ в наше время оказываются ориентированы на корпус. Быстро и эффективно проверить с помощью корпуса особенности употребления незнакомого слова или грамматической формы у авторитетных авторов сможет и иностранец, и школьник, и учитель, и журналист, и писатель.

 Два основных способа деления корпусов на классы:

1) это противопоставление корпусов, относящихся ко всему языку (часто к языку определенного периода), корпусам, относящимся к какому-либо подязыку (жанр, стиль, язык определенной возрастной или социальной группы, язык писателя или ученого и т.п.);

2) разделение корпусов по типу лингвистической разметки. Несмотря на наличие множества типов разметки, большинство реально существующих корпусов относится к корпусам морфологического либо синтаксического типа (treebanks, «банки синтаксических структур»).

Для анализа и сравнения корпусов необходим набор признаков-характеристик. В частности:

| | |
|--|---|
| <ul style="list-style-type: none"> ➤ Тип данных ➤ Язык текстов ➤ «Параллельность» ➤ «Литературность» ➤ Специфичность ➤ Жанр ➤ Доступность ➤ Назначение | <ul style="list-style-type: none"> ➤ Динамичность ➤ Разметка ➤ Характер разметки ➤ Объем текстов ➤ Хронологический аспект ➤ «Общность» ➤ Структура |
|--|---|

Сайты национальных корпусов:

| |
|---|
| <p>Национальный корпус русского языка http://ruscorpora.ru</p> |
| <p>Компьютерный корпус текстов русских газет конца XX-го века http://www.philol.msu.ru/~lex/corpus</p> |
| <p>Корпус русского языка ХАНКО (Хельсинский университет) http://www.ling.helsinki.fi/projects/hanco/</p> |
| <p>Корпуса русских текстов на сайте Университета в Лидсе, Великобритания http://corpus.leeds.ac.uk</p> |
| <p>Русские корпуса Тюбингенского Университета http://www.sfb441.uni-tuebingen.de/b1/en/korpora.html</p> |
| <p>Словарь-корпус языка А.С. Грибоедова http://www.inforeg.ru/electron/concord/concord.htm</p> |

ЛЕКЦИЯ 3

Создание корпусного словаря с помощью Баз Данных

Как известно, традиционный вид сбора данных для словаря – это каталожная карточка, в которой указываются следующие параметры: описываемое слово, пример употребления, источник примера, автор, сфера функционирования. Этот трудоемкий и длительный процесс может быть упрощен благодаря использованию компьютерных технологий, которые позволяют максимально облегчить сбор и хранение информации, используя для этого расширенную базу данных (база данных – это совокупность определенным образом упорядоченных сведений о некоторых объектах [8, с. 146]). В отличие от обычной бумажной картотеки записи базы данных позволяют легко вносить изменения, автоматически отбирать нужные примеры, сортировать данные по выбранным параметрам. Однако с развитием объема информации возросла сложность его обработки. На данный момент в компьютерной лингвистике не существует специализированных лексикографических баз данных, однако такие базы данных, как ACCESS , FOX - Base , PARADOX , D - Base могут быть использованы для ведения словарных картотек, а также для их редактирования [8].

Решение задач, связанных с поиском примеров, иллюстрирующих то или иное языковое явление, а также с уточнением употребления отдельных лексем привело к появлению нового направления в прикладной лингвистике – к *корпусной лингвистике* [7], т.е. науке, которая занимается созданием корпусов текстов и их использованием.

За последнее десятилетие корпусная лингвистика быстро приняла тот опыт, который был накоплен в других достаточно развитых областях, объединяющихся термином АОТ (автоматическая обработка текста), или же АПТ (автоматическое понимание текста)

Информационный анализ реализуется в системах типа ИПС (информационно-поисковые системы), а лингвистический анализ стремится интерпретировать, выявить смысл текста, при этом лингвистический анализ характеризуется локальностью, т.е. проходит в рамках одного предложения. Информационный анализ, наоборот, “видит” текст как целое. Эти два направления развивались почти параллельно, практически никак не соприкасаясь, вследствие чего не получили достаточного развития. Лингвистический анализ не дал корректного машинного перевода, т.к. одна неправильно интерпретированная фраза может в контексте приобретать совершенно другое значение.

В настоящее время в научном словаре лингвистов появляются очень близкие понятия: “электронные библиотеки”, “массив текстов”, “коллекция текстов”, “электронный архив”, “полнотекстовая база данных”. Среди них можно выделить лингвистические корпуса, или языковые корпуса.

Корпус текстов - это вид корпуса данных, единицами которого являются тексты или их достаточно значительные фрагменты, включающие, например, какие-то полные фрагменты макроструктуры текстов данной проблемной области

Корпус текстов характеризуется **четырьмя** основными параметрами:

- во-первых, он должен быть достаточно большого объема;
- во-вторых, корпус должен быть структурированным или размеченным;
- в-третьих, тексты, составляющие определенный корпус должны быть в электронном варианте;

- в четвертых, в понятие “электронный корпус” входит, как правило, специальное программное обеспечение для работы с этим корпусом.

Ценность корпуса нам видится в следующем:

- однажды сделанный корпус может многократно использоваться;
- корпус показывает языковые данные в их реальном окружении, что позволяет исследовать лексическую и грамматическую структуру языка, а также непрерывные процессы языковых изменений, происходящие в языке на протяжении определенного отрезка времени;
- корпус характеризуется представительностью, или сбалансированным составом текстов, что позволяет использовать его для тестирования поисковых машин, машинных морфологий, систем перевода, а также использовать его в различных лингвистических исследованиях;
- корпус имеет важное значение для преподавания русского языка, так как с помощью корпуса можно быстро и эффективно проверить особенности употребления незнакомого слова или грамматической формы.

Работа с корпусами, то есть с массивами текстов, представленными в электронном виде, давно уже стала одним из основных методов лингвистических исследований. Так, еще в 1960-е годы создавался Брауновский корпус (США), который включает 1 млн. слов; в 1970-е : LOB корпус (Великобритания, Норвегия), который также включает 1 млн. слов. В 1980-е годы начали создаваться такие корпуса, как: Машинный Фонд русского языка, Уппсальский корпус русского языка (Швеция), 1 млн. слов, The Bank of English , Birmingham , 20 млн. слов. В 1990-е создан British National Corpus , который включает 100 млн. слов, а также другие национальные корпуса (венгерский, итальянский, хорватский, чешский, японский) объемом 100 млн. слов. В это же время создавался The Bank of English , Birmingham , включающий 600 млн. слов. В начале XXI в. создавались такие корпуса ,как American National Corpus , 100 млн. слов и Gigaword corpora (английский, арабский, китайский), включающий 1 млрд. слов. Особое внимание следует обратить на русский национальный корпус [10], над которым работают лингвисты Москвы и Санкт-Петербурга.

Существование корпусов текстов дает возможность значительно расширить и автоматизировать анализ языкового материала, который является важнейшей базой любого лингвистического исследования. Чем больше материалов анализируется, тем выше значимость выводов и уровень их достоверности.

Современные компьютерные программы позволяют находить нужные примеры из корпусов текстов, которые хранятся в электронном виде на компьютере. Это экономит значительное количество времени по сравнению с традиционной технологией сбора примеров вручную.

Отметим, что именно аннотация, или разметка, - главная характеристика корпуса, которая и отличает его от электронных коллекций, библиотек, энциклопедий, широко представленных в современном Интернете [4]. Разметка текста – это приписывание тексту определенной информации для более удобного анализа.

Существуют различные типы разметки:

- метатекстовая разметка (автор, название, дата создания, объем, тематика текста и т.д.), которая характеризует текст в целом;
- структурная разметка представляет собой информацию о структуре текста, которая позволяет отделить одно слово от другого, выделить границы словосочетания, предложения, текста;
- лингвистическая разметка заключается в приписывании единицам текста определенной лингвистической информации (отрицательное предложение или вопросительное, управление или примыкание и т.д.).

Как известно, чем богаче и разнообразнее разметка, тем выше научная и учебная ценность корпуса.

Пространство электронных текстовых корпусов дало возможность результативного использования *электронных конкордансов*, которые открывают перспективы моделирования языковой картины мира. Конкорданс является специализированной лингвистической прикладной программой, с помощью которой осуществляется автоматическая выборка заданных языковых единиц из электронных текстов. Функцию конкорданса можно сравнить с функцией поиска в текстовом редакторе, однако возможности конкорданса шире – он анализирует не один, а сразу несколько текстов или корпусов электронных текстов, при этом конкорданс выводит на экран информацию о контексте использования заданных языковых единиц. В зависимости от технических возможностей конкорданс может предоставлять информацию о частотности употребления и сочетаемости той или иной языковой единицы, а также дает возможность обращаться к конкретному тексту, в котором был найден пример.

Корпусная лингвистика ярко и четко проявляет себя в одном из перспективных направлений – *компьютерной лексикографии*. Электронные словари и энциклопедии разрабатываются сегодня как автономные и сетевые программные продукты. Не вызывает сомнения тот факт, что электронные словари предоставляют пользователю множество дополнительных возможностей по сравнению с печатными аналогами:

- позволяют хранить большой объем информации за счет использования гиперссылок;
- электронный словарь имеет эффективную систему поиска (полнотекстовый поиск, одновременный поиск в нескольких словарях, скорость поиска);
- в электронных словарях могут применяться средства мультимедиа: озвучивание заголовочных слов, введение иллюстративного материала с фотографиями, анимацией, видеофрагментами;
- возможность использование словарей в локальной и глобальной сетях;
- значительная экономия времени и материальных ресурсов при создании компьютерных словарей.

Значительный опыт исследований в компьютерной лексикографии позволяет расширить возможности изучения смысловой структуры политического слова. Так, сегодня в научных проектах активно разрабатывается идея, связанная с организацией концептуального пространства языка политики с позиции гипертекстового единства [3].

Говоря о специфике содержания такого словаря, отметим, что слово в электронном варианте реализуется в семантическом макрополе, состоящем из нескольких полей, выстроенных в определенную систему:

- 1) *поле индексации* служит для удобного поиска в базе данных;
- 2) *поле заголовочного слова*;
- 3) *поле грамматической информации*, прагматическая нагрузка которой связана с конкретным грамматическим статусом в текстах;
- 4) *поле этимологических справок*;
- 5) *поле авторского толкования* строится не по всей семантической парадигме слова, а лишь по ее основной части, тематически связанной со специфической сферой употребления (например, географической, политической,...);
- б) *поле иллюстраций* выявляет семантику политического слова в определенном контекстуальном окружении. Важно отметить, что в словаре иллюстрации из текстов различных жанров используются и в традиционном виде как оправдательный контекст (это особенно важно для метафорических значений) и как лексикографический прием, помогающий создать пространственную, временную и социально-историческую перспективы;
- 7) *поле фразеологии*;
- 8) *поле энциклопедических сведений* отражает основные этапы становления данной дефиниции, эволюцию его доминантных признаков.

Компьютерное обеспечение словаря толково-справочного типа складывается из двух составляющих:

- база данных словаря определяется системой управления базами данных Access ;
- интерфейс разрабатывается специально для данного словаря. Базой для его реализации послужил язык программирования Delphi , который позволяет осуществлять задачи автоматического поиска слов, а также удобной навигации внутри словаря.

ЛЕКЦИЯ 4

Здесь приводится лексикографический виртуальный портрет слова «эскалация» .

Используются три типа словарей : толковых, энциклопедических терминов и иллюстраций.

Инструкция по заполнению базы данных

«Словарь политических терминов»

1 Пример заполнения БД Словарь политических терминов

The screenshot displays a web-based database interface for a political terminology dictionary. The main window, titled "База данных", shows the word "Эскалация" in the search field. Below the search field are three tabs: "Поле иллюстраций", "Поле толковых словарей", and "Поле энциклопедических сведений?". A sidebar on the left contains a list of terms, with "Эскалация" selected. The "Заголовочное слово" section provides a definition: "Эскалация, и, ж. [англ. escalation букв. восхождение с помощью лестницы]. Постепенное усиление, увеличение, расширение, распространение чего-нибудь. Эскалация войны." Three smaller windows are overlaid on the main interface, each displaying content from a different source:

- Поле иллюстраций:** Contains three news snippets. The first discusses price escalation in St. Petersburg. The second mentions the Ukrainian financial system's vulnerability to political crisis escalation. The third discusses terrorism escalation in Russia.
- Поле толковых словарей:** Shows definitions from several dictionaries, including a brief political dictionary and the Oshchepkov dictionary, defining escalation as a gradual increase or intensification.
- Поле энциклопедических сведений:** Provides an encyclopedic definition of escalation as a process of increasing intensity and spread in various contexts, including armed conflicts.

2 Порядок заполнения БД

- Для ввода новых терминов в поле **Слово** (нижнее) нажать клавишу **Insert** и против * ввести термин.
- Для удаления термина из списка нажать сочетание [**Ctrl+Del**]
- Поиск термина в списке производится *по первым символам* (в отличие от использования шаблона !)
- Отмена ввода термина производится нажатием клавиши **Назад** (отменяет ВСЕ предыдущие действия пользователя)
- После окончания ввода в любое из полей нажать клавишу **Принять**

Рекомендуемая литература:

- 1) Берков В. П. Двухязычная лексикография. – СПб.: Изд-во С.-Петербур. ун-та, 1996. – 248 с.
- 2) Герд А. С. Научно-техническая лексикография / Прикладное языкознание. – СПб.: Изд-во С.-Петербур. ун-та, 1996. – С. 287 – 307.
- 3) Карпова О. М. Английская лексикография. – М.: Академия, 2010. – 176 с.
- 4) Семёнов А. Л. Современные информационные технологии и пе-ревод. – М.: Академия, 2008. – 224 с.
- 5) Соловьёва А. В. Профессиональный перевод с помощью ком-пьютера. – СПб.: Питер, 2008. – 160 с.
- 6) Убин И. И. ЭВМ и словарь. – М.: Всероссийский центр перево-дов, 1992. – 180 с.
- 7) Чепик Е. Ю. Компьютерная лексикография как одно из направ-лений современной прикладной лингвистики // Ученые записки Таврического национального университета им. В.И. Вернадско-го. Серия «Филология». Том 19 (58). – 2006. – № 2. – С. 274-279.
- 8) Krishnamurthy R. Corpus-driven lexicography // International Journal of Lexicography. – Vol. 21 No. 3. – Oxford University Press, 2008. – pp. 231-242.
- 9) Schryver, de Gilles-Maurice. Lexicographer’s dreams in the elec-tronic-dictionary age // International Journal of Lexicography, Vol. 16 No. 2. – Oxford University Press, 2003. – pp. 143-198.

Электронные ресурсы:

- 1) Большаков И. А. Кросслексика – большой электронный словарь сочетаний и смысловых связей русского языка // Международ-ная конференция по компьютерной лингвистике «Диалог 2009». – URL: <http://www.dialog-21.ru/dialog2009/materials/html/08.htm>
- 2) Воскресенский А. Л., Хахалин Г. К. Мультимедийный толковый словарь русского

- жестового языка // Международная конференция по компьютерной лингвистике «Диалог 2007». – URL: <http://www.dialog-21.ru/dialog2007/materials/html/18.htm>
- 3) Ляшевская О.Н., Шаров С.А. Частотный словарь национального корпуса русского языка: концепция и технология создания // Международная конференция по компьютерной лингвистике «Диалог 2008». – URL: <http://www.dialog-21.ru/dialog2008/materials/html/53.htm>
- 4) Остапова И. В. Лексикографическая структура этимологического словаря и его представление в цифровой среде // Международная конференция по компьютерной лингвистике «Диалог 2009». – URL: <http://www.dialog-21.ru/dialog2009/materials/html/55.htm>
- 5) Селегей В. Электронные словари и компьютерная лексикография / Ассоциация лексикографов Lingvo, 2004. – URL: http://www.lingvoda.ru/transforum/articles/pdf/selegey_a1.pdf