

УДК 81'37

ГЛАГОЛЬНАЯ ЛЕКСИКА В ЛЕКСИКОГРАФИЧЕСКИХ БАЗАХ ДАННЫХ: ОБЗОР ОСНОВНЫХ РЕСУРСОВ

А.М. Галиева

НИИ «Прикладная семиотика» Академии наук Республики Татарстан
amgalieva@gmail.ru

Аннотация

Дан краткий обзор электронных лексикографических ресурсов и баз данных, представляющих семантику глагола. Глагол как одна из самых сложных, семантически ёмких и грамматически содержательных частей речи в любом языке характеризуется разветвлённой системой значений и грамматических форм. Семантическая структура глагола – это комплекс онтологических и реляционных семантических компонентов, которые могут получать свое формальное выражение на разных уровнях языковой структуры. При фиксации глаголов в электронных лексикографических ресурсах разработчики исходят из различных методологических установок и отдают приоритет различным аспектам семантической организации глагольной лексики.

Ключевые слова: лексикографические ресурсы и базы данных, семантика, глагол, семантические классы слов.

ВВЕДЕНИЕ

В последние десятилетия появление поисковых систем и аннотированных лингвистических корпусов значительно расширило арсенал лингвистических исследований. Созданы и продолжают разрабатываться специальные лексикографические ресурсы, позволяющие получить актуальную информацию о семантике и распределении слов отдельных частей речи, о моделях управления лексемы, вариативности лексических единиц разных классов, посмотреть иллюстративный материал на примерах из реальных текстов и т. п.

Компьютерные словари глаголов являются ключевыми в процессе обработки естественного языка, нацеленного на интерпретацию данных. Данная статья представляет собой краткий обзор основных англоязычных электронных лексикографических ресурсов, разработанных для представления семантики глагола.

Идеографический словарь является специфическим объектом, который дает возможность исследовать системные свойства лексики языка, разнообразные сигнификативные и логико-семантические отношения, проявления реальных связей данной лексемы с другими, разнообразные реляционные свойства.

СЕМАНТИКА ГЛАГОЛА: КЛЮЧЕВЫЕ ОСОБЕННОСТИ

Для различных типов слов в лингвистической литературе предлагаются различные форматы представления значения. Глагол – одна из самых сложных, семантически ёмких и грамматически содержательных частей речи с чрезвычайно разветвлённой системой значений и грамматических форм. Семантическая структура глагола – это комплекс семантических компонентов, которые могут получать свое формальное выражение на разных уровнях языковой структуры. Подразделение всей глагольной лексики конкретного языка на семантические классы и семантические подклассы, определение типов отношений между лексическими единицами имеет важное значение не только в теоретическом, но и практическом отношении. Полное семантическое описание глагола должно, с одной стороны, основываться на определении места лексемы в системе языка, представлении его парадигматических отношений с другими словами, с другой стороны, учитывать синтаксический потенциал и типичные коллокации глагольной единицы.

Значение слов с предметным значением кардинально отличается от манифестации содержательной составляющей признаков – глагольных и адъективных – лексем. Если базовое общекатегориальное значение предметности у существительных проистекают из эссенциалистских установок (философская категория сущности), то в глагольных лексемах отражаются процессуальные признаки, действия и состояния.

Семантика существительных основывается на общекатегориальном значении предметности и «носит абсолютный, самодостаточный характер», в то время как семантика глагольных и адъективных лексем «носит реляционный характер»

в силу того, что в ее основе лежит понятие признака, который «имплицитно несет некую субстанцию, предмет, которым он должен и может быть придан» [1]. Глагол задает минимальные логико-семантические модели: субъект и его действие или состояние; действие и объект, над которым осуществляется это действие и т. п. Сами эти модели вытекают из семантической валентности — детерминированной лексическим значением необходимости сочетания глагола с другими словами. Следствием этого является то, что семантика глагольного слова — «не элементарна, а комплексна в том смысле, что она отображает не законченное, полное понятие о классе предметов, как это имеет место в предметных именах, а минимальные дискретные «кусочки действительности», приближающиеся к элементарным ситуациям и событиям» [1]. Поэтому «универсальным свойством лексического значения глагольных лексем является то, что каждый полнозначный глагол представляет собой потенциальную синтагму; в содержательном плане признаковое имя формирует свое значение: 1) в акте знакообразования, в номинации с учетом носителя (субъекта, объекта) данного признака; 2) при функционировании в речи, где оно дополнительно уточняется, конкретизируется и формирует тем самым круг сочетающихся с ним предметных имен» [1].

Поскольку семантика глагола имеет множество измерений, классификационные схемы, предложенные разными исследователями, могут в корне отличаться. Имеются разнообразные семантические классификации глаголов, выполненные на материале различных языков. Наиболее часто в литературе упоминается классификация В. Levin, которая классифицирует английские глаголы (около 3100 единиц); основываясь на сходстве их значений и синтаксических свойств; исследователь выделяет семантические классы, принимая в расчет широкий спектр возможных синтаксических преобразований лексем (преимущественно мены диктанта, локативности и пр.), которые отражают значение глагола [2]. Работа В. Levin, основанная на идеях генеративной лингвистики, показывает, что семантические признаки глагола имеют явно выраженную корреляцию с его синтаксическими свойствами и интерпретацией его аргументов. Несмотря на то, что полнота и пригодность классификационной сетки, предложенной В. Levin, для описания реальных языковых фактов вызывает некоторые вопросы (см. об этом ниже), семантические классы, выделенные В. Levin, стали основой для многих других

классификаций и лексикографических баз данных.

ПРОЕКТ FRAMENET

Проект FrameNet (<https://framenet.icsi.berkeley.edu>) представляет собой практическую реализацию семантики фреймов (Frame Semantics) – лингвистической концепции, предложенной Ч. Филлмором [3–6]. Семантика фреймов является своего рода продолжением падежной грамматики и представляет собой подход к формализованному описанию деятельности человека в контексте ситуации. Основную идею семантики фреймов можно определить следующим образом: без подключения энциклопедических знаний о деятельности человека нельзя понять и описать значение слова. Слово активирует рамки семантических знаний, относящихся к конкретному понятию, с которым оно соотносится. Таким образом, в проекте FrameNet слова группируются согласно концептуальным структурам (фреймам), лежащим в их основе, и их сочетаемость выводится индуктивным путем на основе корпусных данных.

Работа над проектом FrameNet была начата в 1997 году в Международном институте информатики (International Computer Science Institute) в Беркли, Калифорния. FrameNet основан на таких терминах семантики фреймов, как семантический фрейм, элементы фрейма, взаимоотношения между фреймами.

Основание для сравнения	FrameNet	Levin 1993
Группировки	230 семантических фреймов	193 глагольных класса
Основание	лексическая семантика	синтаксис аргументов
Источники данных	корпусы	литература по лингвистике
Покрытие	2100 сущ., 1700 глаголов (включая аналитические конструкции), 460 прил.	3100 глаголов
Результаты	описания фреймов и аннотированные примеры	глагольные классы и чередования (преимущественно с описанием)

Таблица 1. Сопоставление проекта FrameNet и классификации глаголов В. Levin

Основными особенностями данного проекта являются опора на корпусные данные при выведении семантических и синтаксических обобщений, а также представление валентностей целевых слов, в которых смысловая составляющая выводится через семантику фреймов [4].

В статье [7] представлено сопоставление классификации глаголов В. Levin с представлением глаголов в проекте FrameNet. Для наглядности воспользуемся таблицей из данной работы (количественные данные по FrameNet отражают состояние проекта на 2002 г., когда была опубликована эта статья).

Так как количество категорий в обоих ресурсах сопоставимо (230 фреймов против 193 глагольных классов), можно предположить, что значительная часть фреймов будет так или иначе соотноситься с глагольными классами, выделенными В. Levin. В действительности между анализируемыми группировками имеются значительные расхождения. В проекте FrameNet отнесение предикатов к фрейму обусловлено лишь общностью их семантики, и сходства в синтаксическом поведении лексемы не требуется, соответственно, в один фрейм могут быть объединены глаголы с разными типами мены диатезы и другими альтернативами. В то время как модели альтернатив для В. Levin — решающий критерий, лексемы с разным типом синтаксического поведения в один класс объединены быть не могут. Классификация FrameNet строится на реальных корпусных данных, а классификация В. Levin во многом носит умозрительный характер; авторы статьи [7] отмечают, что во многих классах у В. Levin имеются примеры гипотетических конструкций, которые не подкрепляются корпусными свидетельствами. Кроме того, основной тезис В. Levin о том, что группировка слов согласно их синтаксическому поведению дает семантически релевантные классы, не всегда подтверждается, так как слова, близкие по значению, могут попасть в разные классы и, наоборот, слова, довольно далекие по значению, могут быть объединены в один класс. Кроме того, семантические фреймы, выделенные в проекте FrameNet, позволяют описывать не только глаголы, но и существительные и прилагательные [7].

На рис. 1 представлены результаты на запрос по глаголу *to see* 'видеть' в проекте FrameNet.

FRAME NET DATA SEARCH FOR SEE

Frame search results: Closest match is see
 See_through, Seeking, Seeking_to_achieve

Lexical unit search results: Closest match is see

Lexical Unit	Frame	LU Status	Lexical Entry Report	Annotation Report
see (through).v	See_through	Created	LE	
see eye to eye.v	Be_in_agreement_on_assessment	Created	LE	
see.n	Relational_political_locales	Created	LE	
see.v	Perception_experience	Needs_SCs	LE	Anno
see.v	Grasp	Finished_Initial	LE	Anno
see.v	Categorization	Created	LE	Anno
see.v	Touring	Insufficient_Attestations	LE	Anno
see.v	Reference_text	Needs_SCs	LE	Anno
see.v	Causation	Created	LE	Anno
see.v	Eventive_affecting	Created	LE	
see.v	Condition_symptom_relation	Add_Annotation	LE	Anno
seed.v	Filling	Finished_Initial	LE	Anno
seed.v	Emptying	Created	LE	
seeing.v	Personal_relationship	Finished_Initial	LE	Anno

Рис. 1. Фреймы с глаголом *to see* в проекте FrameNet

Основополагающие принципы англоязычного FrameNet'а успешно используются для анализа и описания типологически разных языков: испанского, немецкого, шведского, корейского, японского, китайского и других [8].

В настоящее время в России разрабатывается проект общедоступного электронного словаря глагольных конструкций FrameBank (<http://framebank.ru>). Разработчики позиционируют свой проект как «создание русского фреймнет-ориентированного ресурса, спроектированного с учетом традиций отечественной лексической семантики и специфики русского языка, где информация о предложно-падежной реализации управления предикатов и поверхностно-синтаксических свойствах других конструкций имеет особую ценность» [9].

FrameBank не является русскоязычной копией FrameNet'а, а имеет свою специфику. Если центральным элементом FrameNet'а являются фреймы — типовые ситуации с известным набором участников и расписанными ролями, то русский FrameBank строится вокруг конструкций конкретных лексем [10].

FrameBank описывает:

- русскую лексическую систему, структуру лексико-семантических групп и

полисемии в русском языке (разработчики англоязычного проекта FrameNet исходят из идеи, что сеть фреймов универсальна для всех языках);

- парадигматические отношения между значениями многозначных слов (как они отражаются в системе связанных с этими значениями лексических конструкций);

- лексико-семантические ограничения на слоты конструкций;

- грамматические особенности русского языка (порядок слов, особенности использования падежей, согласования и т. п.) (10).

Ядро системы FrameBank составляют 2200 ключевых русских глаголов и ассоциированных с ними конструкций и корпусных примеров, для каждого глагола указывается семантический класс. Каждая конструкция представлена в виде шаблона, в котором указаны: а) морфосинтаксические характеристики элементов конструкции; б) синтаксический ранг участника; в) экспликация (роль) участника; г) основные семантические ограничения на заполнение слота [10]. На рис. 2 представлены конструкции с глаголом *видеть* из списка конструкций FrameBank.

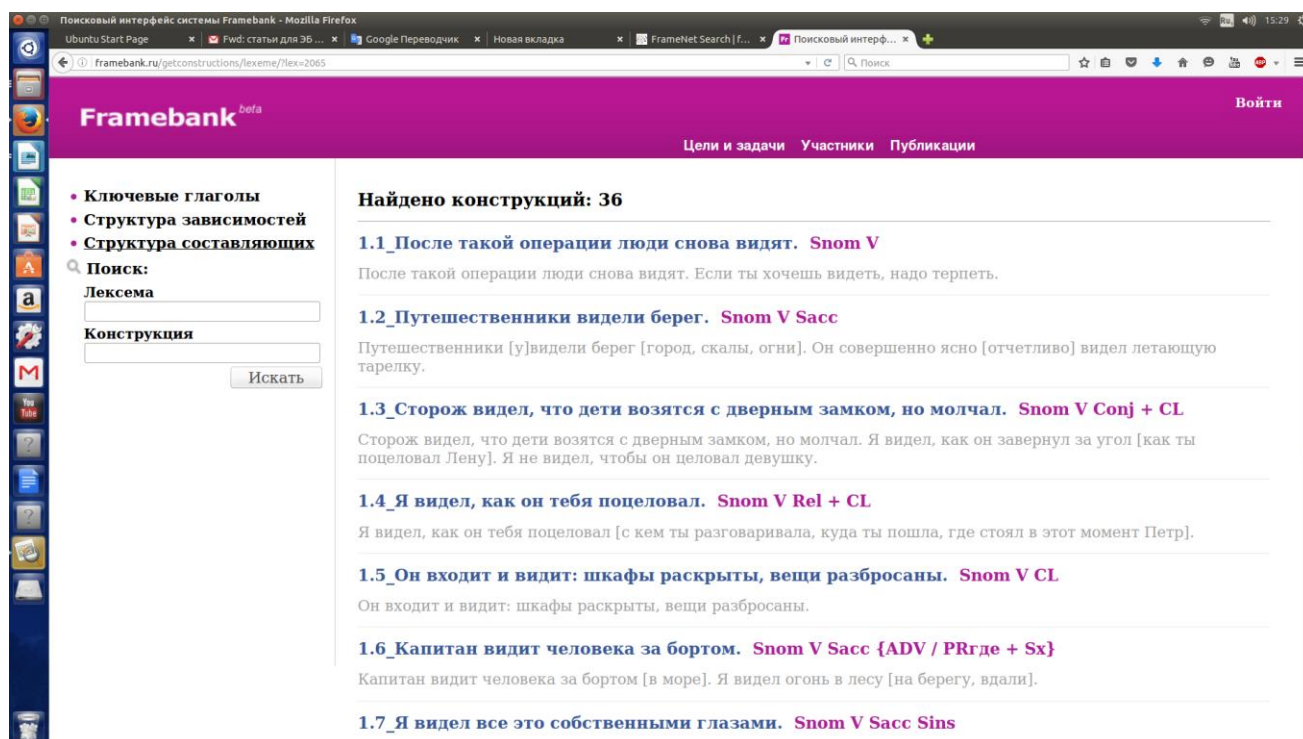


Рис. 2. Конструкции, содержащие глагол *видеть*, в проекте FrameBank

ПРОЕКТ VERBNET

Как уже отмечалось, семантические классы, выделенные В. Levin, стали основой для многих других классификаций и лексикографических баз данных, в частности, проекта VerbNet [11], который, по мнению разработчиков, в настоящее время является самым большим глагольным компьютерным лексикографическим ресурсом. В данном тезаурусе представлен более расширенный, по сравнению с классификацией В. Levin, вариант семантических классов; для каждого класса описаны тематические роли, представлены ограничения на отбор аргументов, а также синтаксические фреймы. VerbNet представляет собой иерархический тезаурус, не привязанный к конкретной предметной области. В VerbNet каждый класс имеет синтаксическое описание, отображающее возможные поверхностные реализации структуры аргумента в типах конструкций, содержащих переходные и непереходные глаголы, сочетания глагола с предлогами, а также большой набор залоговых чередований. Семантические ограничения (например, живое существо, человек, организация) использованы для того, чтобы отобразить типы семантических ролей. Первоначальная версия VerbNet, как уже отмечалось, была основана на классах, выделенных В. Levin, и применение данного ресурса для обработки естественного языка сталкивалось со значительными трудностями. Поэтому исследователи предлагают способы автоматического расширения данного ресурса за счет корпусных данных [11, 12]. VerbNet предполагает отображение связей между глаголами в данном ресурсе и отдельными значениями глаголов, представленными в WordNet, а также фреймами FrameNet.

Рис. 3 представляет фреймы с глаголом *to see* в проекте VerbNet.

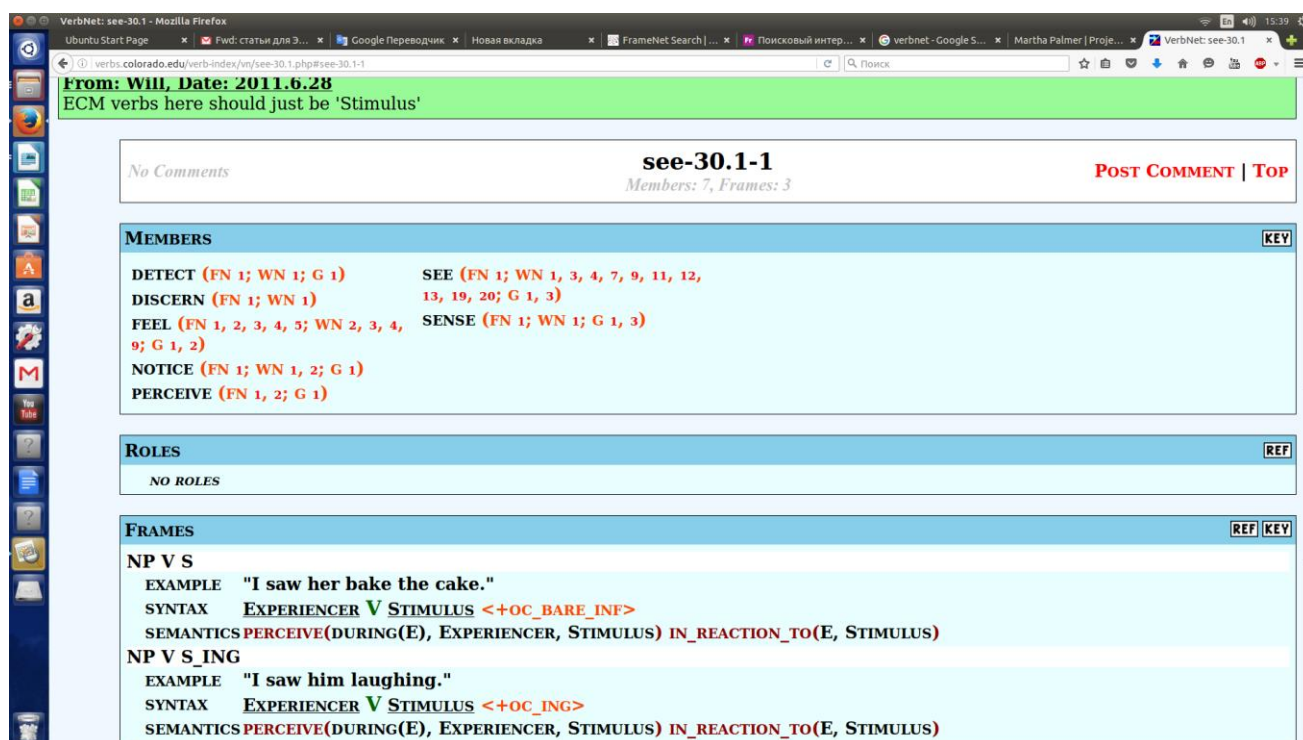


Рис. 3. Фреймы, содержащие глагол *see*, в проекте VerbNet

ПРОЕКТ PROPBANK

Еще одним англоязычным ресурсом, нацеленным на представление глагольной семантики, является PropBank (<http://propbank.github.io>) который представляет «банк пропозиций» английских глаголов [13]. Изначально проект разрабатывался как корпус текстов с аннотацией важнейших семантических пропозиций. Отношения «предикат – аргументы» были добавлены к синтаксическим деревьям проекта Penn Treebank (<http://www.cis.upenn.edu/~treebank>). Penn Treebank позволяет приписывать некоторые семантические тэги (например, время или локативность) для ряда конструкций, но не определяет семантические роли грамматического субъекта или объекта. Так как один и тот же глагол в сходном синтаксическом окружении может иметь аргументы с разными семантическими ролями, данные роли не могут правильно определяться автоматически. PropBank нацелен на создание независящего от предметной области корпуса с ручной разметкой семантических ролей. Для каждого глагола определяется набор семантических ролей его аргументов, и таким образом аннотируется каждый пример из Penn Treebank [13].

Рис. 4 показывает представление предиката to see в проекте PropBank.

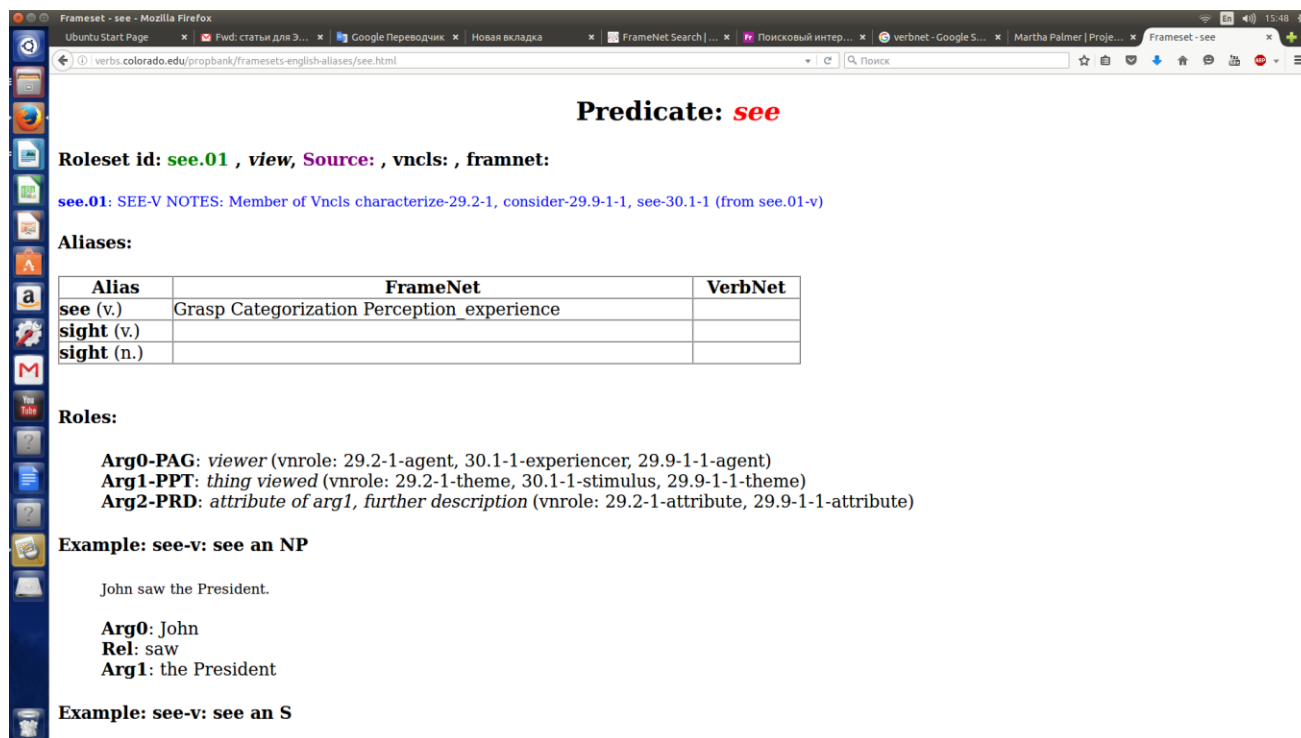


Рис. 4. Глагол to see в проекте PropBank

Отметим основные отличия PropBank от FrameNet. PropBank представляет собой специальный ресурс, разработанный для глагольной лексики, в то время как FrameNet описывает значение глаголов (как и других частей речи) в рамках более общей семантики фреймов. PropBank представляет глаголы из конкретного корпуса, а FrameNet выбирает наборы примеров предложений из большого корпуса, и только в некоторых случаях имеются аннотированные отрезки непрерывного текста. Формат аннотирования в PropBank тесно связан с синтаксическим уровнем (требуется, чтобы все аргументы глагола имели синтаксическое выражение; считается, что значения слов отличаются только в том случае, если имеются различия в аргументах), а аннотация в FrameNet в большей степени обусловлена семантикой.

В настоящее время остро встает задача установления соответствий между разными ресурсами, описывающими глагольную лексику, в частности, важное значение имеет связывание проектов VerbNet и PropBank [14]

ПРЕДСТАВЛЕНИЕ ГЛАГОЛОВ В WORDNET

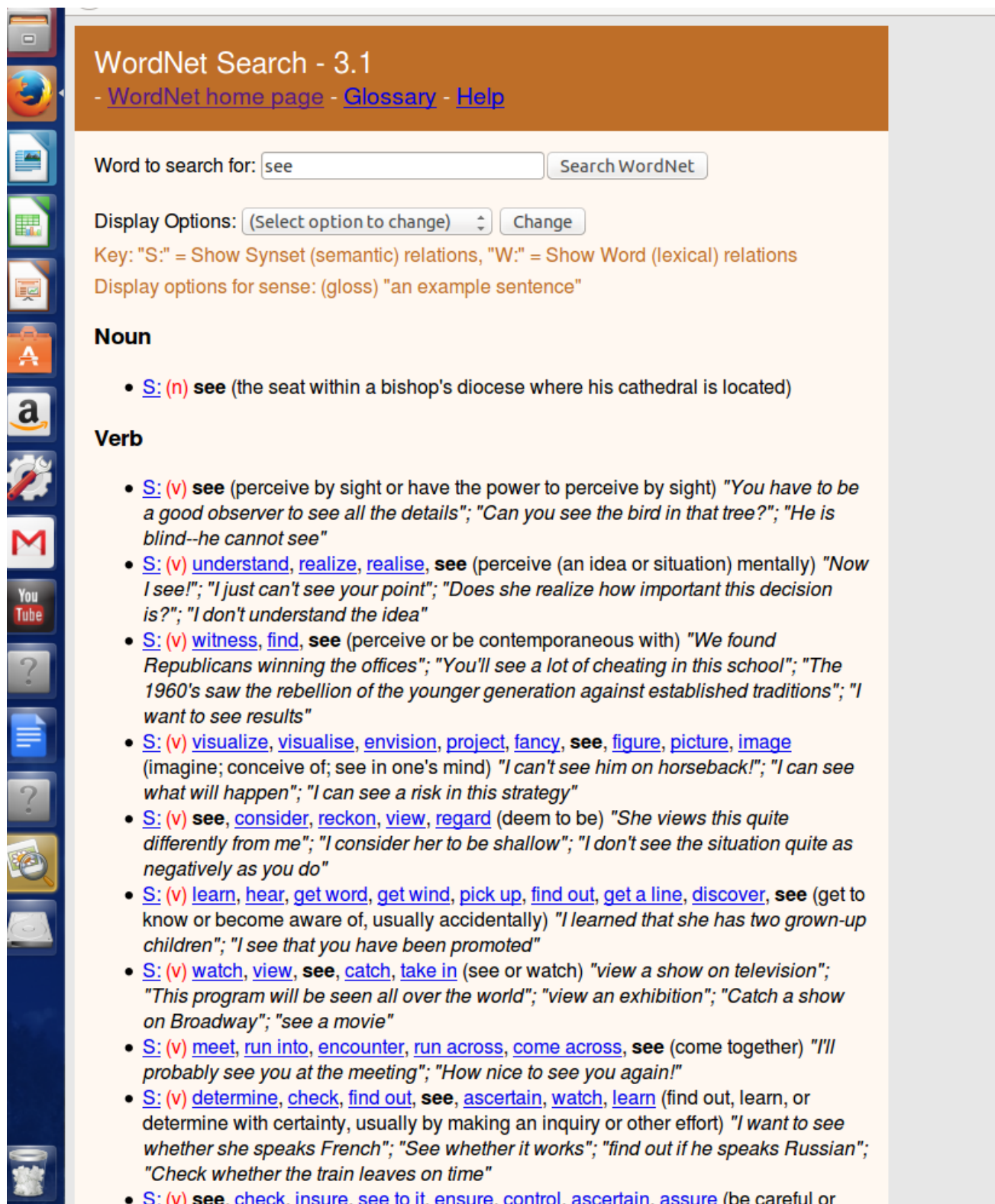
Тезаурус WordNet (<https://wordnet.princeton.edu>) является одним из наиболее известных лексикографических ресурсов в области компьютерной лингвистики и автоматической обработки текстов. Он был разработан в 1995 году в Принстонском университете. WordNet представляет собой иерархическую сеть лексикализованных понятий (синсетов). Основными единицами структуры wordnet являются синонимические ряды - синсеты, которые связаны между собой различными семантическими отношениями [15, 16]. Синонимические отношения в тезаурусе определяются не между словами, а между отдельными значениями слов.

Несмотря на то, что понятие синонимии является общепринятым, точные критерии синонимичности до сих пор являются предметом дискуссий. В рамках проектов Принстонского WordNet'a и EuroWordNet'a синонимия определяется через понятие взаимозаменяемости: так, в проекте EuroWordNet слова считаются семантически эквивалентными, когда они обозначают один и тот же ряд сущностей, независимо от морфолого-синтаксических, стилистических, диалектных различий, а также различий в прагматическом использовании слова. Кроме того, синонимы не могут быть связаны между собой другими типами семантических отношений [17].

Разработчики wordnet-тезауруса RussNet (http://project.phil.spbu.ru/RussNet/index_ru.shtml) для русского языка критерий взаимозаменяемости рассматривают как дополнительный по отношению к критерию семантической близости. Последний выявляется при дефиниционном анализе, для которого требуется установление идентичности словарных определений или взаимная отсылка в синонимических определениях [18].

В WordNet для разных частей речи используются различные типы отношений. Семантические отношения между синсетами для глаголов описываются следующим образом:

- а) отношения следования (Entailment): *идти* – *шагать*;
- б) отношения тропонимии: *сказать* – *шептать*;
- в) отношения каузативности: *есть* – *кормить*.



The screenshot shows the WordNet Search interface. At the top, the title is "WordNet Search - 3.1" with links to the home page, glossary, and help. Below the title, there is a search input field containing the word "see" and a "Search WordNet" button. Underneath, there are "Display Options" with a dropdown menu set to "(Select option to change)" and a "Change" button. A key is provided: "S:" for Synset (semantic) relations and "W:" for Word (lexical) relations. The "Display options for sense" are set to "(gloss) 'an example sentence'".

Noun

- **S: (n) see** (the seat within a bishop's diocese where his cathedral is located)

Verb

- **S: (v) see** (perceive by sight or have the power to perceive by sight) *"You have to be a good observer to see all the details"; "Can you see the bird in that tree?"; "He is blind—he cannot see"*
- **S: (v) understand, realize, realise, see** (perceive (an idea or situation) mentally) *"Now I see!"; "I just can't see your point"; "Does she realize how important this decision is?"; "I don't understand the idea"*
- **S: (v) witness, find, see** (perceive or be contemporaneous with) *"We found Republicans winning the offices"; "You'll see a lot of cheating in this school"; "The 1960's saw the rebellion of the younger generation against established traditions"; "I want to see results"*
- **S: (v) visualize, visualise, envision, project, fancy, see, figure, picture, image** (imagine; conceive of; see in one's mind) *"I can't see him on horseback!"; "I can see what will happen"; "I can see a risk in this strategy"*
- **S: (v) see, consider, reckon, view, regard** (deem to be) *"She views this quite differently from me"; "I consider her to be shallow"; "I don't see the situation quite as negatively as you do"*
- **S: (v) learn, hear, get word, get wind, pick up, find out, get a line, discover, see** (get to know or become aware of, usually accidentally) *"I learned that she has two grown-up children"; "I see that you have been promoted"*
- **S: (v) watch, view, see, catch, take in** (see or watch) *"view a show on television"; "This program will be seen all over the world"; "view an exhibition"; "Catch a show on Broadway"; "see a movie"*
- **S: (v) meet, run into, encounter, run across, come across, see** (come together) *"I'll probably see you at the meeting"; "How nice to see you again!"*
- **S: (v) determine, check, find out, see, ascertain, watch, learn** (find out, learn, or determine with certainty, usually by making an inquiry or other effort) *"I want to see whether she speaks French"; "See whether it works"; "find out if he speaks Russian"; "Check whether the train leaves on time"*
- **S: (v) see, check, insure, see to it, ensure, control, ascertain, assure** (be careful or

Рис. 5. Глагол to see в проекте WordNet

Отношение тропонимии – это особый случай отношения следования. Глагольные иерархии, представляющие отношение тропонимии, обладают более узкой, но в то же время более кустистой структурой по сравнению с существительными, число уровней в иерархии при этом обычно не превышает четырех [19].

Рис. 5 дает представление о том, как отображается глагол *to see* в принстонском проекте WordNet.

В тезаурусах типа WordNet фиксируются семантические отношения между глаголами различных классов, при этом валентность глагола и особенности его аргументов, а также типы фреймов, в которых он используется, не принимаются в расчет.

Тезаурусы типа Wordnet созданы во многих странах и успешно используются при обработке естественного языка – при информационном поиске, разрешении многозначности, анализе тональности и др. Информация о существующих и разрабатываемых в настоящее время wordnet'ах представлена на сайте Всемирной WordNet-ассоциации [20].

ЗАКЛЮЧЕНИЕ

К настоящему времени разработано значительное число специальных ресурсов для автоматической обработки текстов, большая часть из них создана для английского языка. Особой ролью глагольной лексики в языках можно объяснить то, что многие ресурсы создаются специально для представления значений глаголов. Глагол является синтаксическим ядром предложения, его значение не только номинативно, но и реляционно: именно глагол задает отношения между участниками ситуации и кодирует их в синтаксических структурах.

Многомерность глагольной семантики обуславливает разнообразие подходов к его представлению в лексикографических ресурсах, каждый из которых создавался в рамках конкретных теоретических и методологических установок и ввиду конкретных целей, стоящих перед разработчиками. Значение глагола как синтаксического ядра предложения рассматривается в проектах типа VerbNet и PropBank. FrameNet описывает глаголы в рамках семантики фреймов, а WordNet устанавливает иерархические семантические отношения между отдельными глаголами, не принимая в расчет аргументную структуру глагола.

СПИСОК ЛИТЕРАТУРЫ

1. Уфимцева А.А. Лексическое значение: принцип семиологического описания лексики. М.: Наука, 1986. 239 с.
 2. Levin B. English Verb Classes and Alternations: A Preliminary Investigation. Chicago, University of Chicago, 1993. 348 p.
 3. Fillmore C. Frame Semantics and the Nature of Language // Annals of the New York Academy of Sciences. 1976. 280 (1). P. 20–32.
 4. Baker C.F., Fillmore C.J., Lowe J.B. The Berkeley FrameNet Project // Proceedings of the 17th International Conference on Computational linguistics. 1998, August. V. 1. Association for Computational Linguistics. P. 86–90.
 5. Fillmore C.J., Baker C.F. Frame semantics for text understanding // Proceedings of WordNet and Other Lexical Resources Workshop. Pittsburgh, 2001. URL: <http://www.ccs.neu.edu/course/csg224/resources/framenet/framenet.pdf>.
 6. Fillmore C.J., Baker C.F., Sato H. FrameNet as a “Net” // Proceedings of LREC, Lisbon, 2004. V. 4. P. 1091–1094.
 7. Baker C.F., Ruppenhofer J. FrameNet's Frames vs. Levin's verb classes // Proceedings of the 28th Annual Meeting of the Berkeley Linguistics Society. 2002. P. 27–38.
 8. Boas H.C. (Ed.). Multilingual framenets in computational lexicography: Methods and Applications. 2009. V. 200. Walter de Gruyter. 352 p.
 9. FrameBank. URL: <http://framebank.ru>.
 10. Кашкин Е.В., Ляшевская О.Н. Семантические роли и сеть конструкций в системе FrameBank // Компьютерная лингвистика и интеллектуальные технологии. По материалам ежегодной международной конференции «Диалог 2013. Т. 1. С. 297–311.
 11. Kipper K., Korhonen A., Ryant N., Palmer M. Extending VerbNet with novel verb classes // Proceedings of the Fifth International Conference on Language Resources and Evaluation – LREC'06. May, 2006, Genoa, Italy: 2006. URL: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.380.5541&rep=rep1&type=pdf>.
 12. Kipper K., Korhonen A., Ryant N., Palmer M. A large-scale classification of English verbs // Language Resources and Evaluation. 2008. V. 42 (1). P. 21–40.
 13. Palmer M., Kingsbury P., Gildea D. The proposition bank: an annotated
-

corpus of semantic role // *Computational Linguistics*. 2005. V. 31 (1). P. 71–106.

14. *Loper E., Yi S.T., Palmer M.* Combining lexical resources: mapping between PropBank and VerbNet // *Proceedings of the 7th International Workshop on Computational Linguistics, Tilburg, the Netherlands*. 2007. URL: http://verbs.colorado.edu/~kipper/Papers/semlink_iwcs7.pdf.

15. *Fellbaum C.* *WordNet. An Electronic Lexical Database*. Cambridge, Mass: MIT Press, 1998. 423 p.

16. *Mille G.A.* WordNet: A lexical database for English // *Communications of the ACM*. 1995. V. 38, No 11. P. 39–41.

17. *Vossen P.* (Ed.) *EuroWordNet General Document. Version 3*. URL: <http://vossen.info/docs/2002/EWNGeneral.pdf>.

18. *Азарова И.В., Митрофанова О.А., Синопальникова А.А.* Компьютерный тезаурус русского языка типа wordnet. URL: <http://www.dialog-21.ru/Archive/2003/Azarova.htm>.

19. *Лукашевич Н.В.* Тезаурусы в задачах информационного поиска. М.: Изд-во Московского ун-та, 2011. 511 с.

20. The Global WordNet Association. URL: <http://globalwordnet.org>.

VERBAL VOCABULARY IN LEXICOGRAPHICAL DATA BASES: REVIEWING MAIN RESOURCES

A.M. Galieva

Research Institute of Applied Semiotics of Tatarstan Academy of Sciences

amgalieva@gmail.com

Abstract

This paper gives a brief review of available lexicographical resources and databases representing verbal vocabulary. The verb, being one of the most complicated, semantically intricate and grammatically sophisticated parts of speech, is characterized by multiplicity of senses and forms. Semantic structure of any verb is a complex of ontological and relational meaning components that may find a formal expression on dif-

ferent levels of the linguistic structure. Fixing verbs in electronic lexicographical resources, researchers come from different methodological orientations, and prioritize different aspects of the organization of the semantics of verbal vocabulary.

Keywords: *lexicographical resources and databases, semantics, verb, semantic classes of words.*

REFERENCES

1. *Ufimtseva A.A.* Leksicheskoye znacheniyе: printsip semiologicheskogo opisaniya leksiki M.: Nauka. 1986. 239 s.
2. *Levin B.* English Verb Classes and Alternations: A Preliminary Investigation. Chicago, University of Chicago, 1993. 348 p.
3. *Fillmore C.J.* Frame Semantics and the nature of language // Annals of the New York Academy of Sciences. 1976. 280 (1). P. 20–32.
4. *Baker C.F., Fillmore C.J., Lowe J.B.* The Berkeley FrameNet project // Proceedings of the 17th International Conference on Computational linguistics. 1998, August. V. 1. Association for Computational Linguistics. P. 86–90.
5. *Fillmore C.J., Baker C.F.* Frame semantics for text understanding // Proceedings of WordNet and Other Lexical Resources Workshop. Pittsburgh, 2001. URL: <http://www.ccs.neu.edu/course/csg224/resources/framenet/framenet.pdf>.
6. *Fillmore C.J., Baker C.F., Sato H.* FrameNet as a “Net” // Proceedings of LREC, Lisbon, 2004. V. 4. P. 1091–1094.
7. *Baker C.F., Ruppenhofer J.* FrameNet's frames vs. Levin's verb classes // Proceedings of the 28th Annual Meeting of the Berkeley Linguistics Society. 2002. P. 27–38.
8. *Boas H.C.* (Ed.). Multilingual FrameNets in Computational Lexicography: Methods and Applications. 2009. V. 200. Walter de Gruyter. 352 p.
9. FrameBank. URL: <http://framebank.ru>.
10. *Kashkin E.V., Lyashevskaya O.N.* Semanticheskiye roli i set konstruktsiy v sisteme FrameBank // Kompyuternaya lingvistika i intellektualnyye tekhnologii. Po materialam ezhegodnoy mezhdunarodnoy konferentsii “Dialog” 2013. T. 1. S. 297–311.
11. *Kipper K., Korhonen A., Ryant N., Palmer M.* Extending VerbNet with novel

verb classes // Proceedings of the Fifth International Conference on Language Resources and Evaluation – LREC'06. May, 2006, Genoa, Italy: 2006. URL: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.380.5541&rep=rep1&type=pdf>.

12. *Kipper K., Korhonen A., Ryant N., Palmer M.* A large-scale classification of English verbs// Language Resources and Evaluation. 2008. V. 42 (1). P. 21–40.

13. *Palmer M., Kingsbury P., Gildea D.* The proposition bank: an annotated corpus of semantic role // Computational Linguistics. 2005. V. 31 (1). P. 71–106.

14. *Loper E., Yi S.T., Palmer M.* Combining lexical resources: mapping between PropBank and VerbNet // Proceedings of the 7th International Workshop on Computational Linguistics, Tilburg, the Netherlands. 2007. URL: http://verbs.colorado.edu/~kipper/Papers/semlink_iwcs7.pdf.

15. *Fellbaum C.* WordNet. An electronic lexical database. Cambridge, Mass: MIT Press, 1998. 423 p.

16. *Mille G.A.* WordNet: A lexical database for English // Communications of the ACM. 1995. V. 38, No 11. P. 39–41.

17. *Vossen P.* (Ed.) EuroWordNet general document. Version 3. URL: <http://vossen.info/docs/2002/EWNGeneral.pdf>.

18. *Azarova I.V., Mitrofanova O.A., Sinopalnikova A.A.* Kompyuternyy tezaurus russkogo yazyka tipa wordnet. URL: <http://www.dialog-21.ru/Archive/2003/Azarova.htm>.

19. *Lukashevich N.V.* Tezaurusy v zadachakh informatsionnogo poiska. M.: Izd-vo Mosk. un-ta, 2011. 511 s.

20. The Global WordNet Associaton. URL: <http://globalwordnet.org>.

СВЕДЕНИЯ ОБ АВТОРЕ



ГАЛИЕВА Альфия Макаримовна – ведущий научный сотрудник Научно-исследовательского института «Прикладная семиотика» Академии наук Республики Татарстан. Сфера научных интересов: семантика, грамматика, философия языка.

email: amgalieva@gmail.com

Alfiia GALIEVA – Senior Researcher of Research Institute of Applied Semiotics of the Academy of Sciences of Republic of Tatarstan. Research interests: semantics, grammar, philosophy of language.

email: amgalieva@gmail.com

Материал поступил в редакцию 15 марта 2016 года