

# Компьютерный тезаурус русского языка типа WordNet

Азарова И.В., Митрофанова О.А., Синопальникова А.А.  
(Кафедра математической лингвистики СПбГУ)

## Введение

В настоящее время компьютерные словари, построенные по модели WordNet, являются одним из самых распространенных типов лексических ресурсов, используемых в сфере информационных технологий. Особенности этих словарей, объединяющих в себе черты справочной системы и инструмента для проведения лингвистических исследований, открывают широкие возможности их практического и теоретического применения<sup>1</sup>.

В области **информационного поиска** wordnet-словари<sup>2</sup> удобно использовать для расширения запроса пользователя за счет парадигматически и синтагматически связанных слов, например, компонентов синсета вместе с его гипонимами и согипонимами или связей типа «глагол – актанта», которые дают возможность осуществлять контекстный поиск. Данные о синтагматических отношениях слов позволяют применять wordnet-словари для решения классической задачи снятия **неоднозначности смысла слова**. Wordnet можно использовать для вычисления **смысловой близости** текстов на основе гиперонимических отношений. Wordnet-словари могут служить **лексиконом** для **формальных грамматик**, особенно в силу исчерпывающего описания валентностей глаголов, сочетаемости прилагательных и существительных. Формат wordnet является удобным формализмом для представления состава и **структуры лексики** специальных **подъязыков** (например, медицинских, экономических терминов). Wordnet-словари являются удобным инструментом для проведения **исследований** в области **лексической семантики**, например, гипонимические отношения в wordnet-словарях позволяют определять направление метонимических переносов и прогнозировать появление новых лексико-семантических вариантов.

В рамках проекта RussNet представляется возможным соединить опыт, накопленный традиционной русской лексикографией, и практические возможности стандарта wordnet-словарей, что позволит представить лексику русского языка способом, удовлетворяющим потребностям современной компьютерной лингвистики.

## История и эволюция WordNet

В 1985 г. Дж. Миллером и его коллегами из Лаборатории когнитологии Принстонского Университета (США) была разработана модель ментального лексикона человека<sup>3</sup>. Ресурс получил название WordNet и со временем стал одним из наиболее авторитетных и широко используемых стандартов для построения лексико-семантических баз данных.

Популярность и широкое распространение WordNet обусловлены, прежде всего, его существенными содержательными и структурными характеристиками. Принстонский WordNet и все последующие варианты для других языков направлены на отображение состава и структуры лексической системы языка в целом, а не отдельных тематических областей. Например, настоящая версия WordNet охватывает общеупотребительную лексику современного английского языка – более 120 тысяч слов. Словарь состоит из 4 отдельных файлов для основных

---

<sup>1</sup> См. *Gonzalo J., Verdejo F., Peters C., Calzolari N. Applying EuroWordNet to Cross-Language Text Retrieval // Vossen P. EuroWordNet: A Multilingual Database with Lexical Semantic Networks. Dordrecht, 1998; Voorhees E. M. Using WordNet to Disambiguate Word Senses // Proceedings of the 17<sup>th</sup> International ACM SIGIR Conference on Research and Development in Information Retrieval. Pittsburg, 1994; Loupy C. de, El-Beze M. Managing Synonymy and Polysemy in a Document Retrieval System Using WordNet // Proceedings of the Workshop on Using Semantics for Information Retrieval and Filtering. LREC'2002. Las Palmas, 2002.*

<sup>2</sup> Прочно вошедшее в современную лингвистическую терминологию слово wordnet уже стало нарицательным и в настоящее время употребляется применительно к многочисленным ресурсам, построенным с опорой на принципы, сформулированные разработчиками первого словаря WordNet.

<sup>3</sup> *Fellbaum C. WordNet: An Electronic Lexical Database. Cambridge, 1998. Miller G. et al. Five Papers on WordNet. CSL-Report, vol. 43. Princeton University, 1990.*

знаменательных частей речи: существительных, глаголов, прилагательных и наречий. Базовой структурной единицей Принстонского WordNet является синонимический ряд (синсет), объединяющий слова со схожим значением. Предполагается, что каждый синсет репрезентирует в словаре некоторое лексикализованное понятие данного языка. Для удобства использования словаря человеком каждый синсет дополнен дефиницией и примерами употребления слов в контексте. Синсеты в WordNet связаны между собой такими семантическими отношениями, как гипонимия (родовидовое), меронимия (часть-целое), лексический вывод (каузация, пресуппозиция) и др.; среди них особую роль играет гипонимия: она позволяет организовывать синсеты в иерархические структуры (деревья). Лексика каждой части речи представлена в виде набора деревьев (леса). Для разных частей речи родовидовые отношения могут иметь дополнительные характеристики и различаться областью охвата, например, только некоторые группы прилагательных в WordNet связаны гипонимическими отношениями.

В оригинальной версии словаря не содержится информации о тематической организации лексики, так как, по мнению Дж. Миллера, она представляет собой иной тип структурирования лексики, отличный от собственно семантического, представленного в WordNet.

Основные принципы WordNet не претерпели значительных изменений со времени его создания, однако работы по его усовершенствованию продолжаются, например, в последней версии WordNet 1.7.1 введена информация о частотности значений. Широкое использование WordNet обусловлено в значительной степени тем, что он распространяется свободно.

### Проект EuroWordNet

Новый этап в эволюции wordnet-словарей связан с проектом EuroWordNet<sup>1</sup> (1996-1999 гг.), в рамках которого не только было создано несколько тезаурусов для европейских языков (голландского, испанского, итальянского, немецкого, французского, чешского и эстонского), но и впервые была реализована идея об объединении отдельных wordnet-представлений в общую систему. Все компоненты EuroWordNet были построены по единой модели, что, однако, не предполагало прямого перевода английского варианта WordNet 1.5., перед разработчиками стояла задача – отразить все особенности лексических систем национальных языков. Совместимость компонентов EuroWordNet обеспечивалась единством принципов и заданным набором общих понятий (Basic Concepts), на которых определялась система межъязыковых отсылок (Inter-Lingual-Index), дающих возможность переходить от лексикализованных значений одного языка к сходным, но не обязательно тождественным значениям в другом языке. Данный индекс позволяет использовать EuroWordNet не только для информационного поиска в рамках одного языка, но и для многоязычного поиска.

В рамках проекта EuroWordNet первоначальная структура словаря претерпела серьезные изменения. Был расширен набор семантических отношений за счет парадигматических отношений, связывающих слова разных частей речи (например, XPOS\_NEAR\_SYNONYMY: *dead – death*; XPOS\_HYPERONYMY: *to love – emotion*; XPOS\_ANTONYMY: *to live – dead*) и синтагматических отношений между глаголами и актантами-существительными (например, ROLE\_INSTRUMENT: *to write – pencil*). Увеличение числа связей сопровождалось на практике ограничением классов знаменательной лексики: в тезаурусе представлены только существительные и глаголы.

Был сформирован новый подход к построению wordnet-словарей: с опорой на использование лексикографических источников (толковых, переводных и синонимических словарей) и результатов обработки корпусов современных текстов.

Успешное завершение проекта EuroWordNet послужило толчком к созданию большего числа wordnet-представлений для языков разных типов (например, венгерского, турецкого, арабского, тамильского, китайского и пр.), а также многоязычных ресурсов типа EuroWordNet (например, проект BalkaNet нацелен на объединение греческого, румынского, бол-

---

<sup>1</sup> Vossen P. EuroWordNet: A Multilingual Database with Lexical Semantic Networks. Dordrecht, 1998.

гарского, сербского, турецкого и чешского wordnet-словарей). В 2001 г. была создана Всемирная Ассоциация WordNet (Global WordNet Association), целью которой является объединение уже существующих и только развивающихся национальных ресурсов этого типа, усовершенствование системы межъязыковых индексов и разработка общих стандартов, позволяющих использовать модель WordNet для языков разных типов.

### ***RussNet: компьютерный тезаурус русского языка типа WordNet***

С 1999 г. на кафедре математической лингвистики исследовательская группа под руководством И. В. Азаровой ведет работы по проекту RussNet – созданию русской версии компьютерного словаря типа WordNet. Целью проекта является построение лексико-семантического ресурса

- для отражения организации лексической системы русского языка в целом (в противоположность терминологическим или частным словарям);
- для представления ядра общеупотребительной лексики русского языка;
- для фиксации всех семантических, семантико-грамматических и семантико-деривационных отношений, существенных для лексикона русского языка.

Для реализации поставленной цели мы провели ряд собственных теоретических и практических исследований<sup>1</sup>. На первом этапе было выявлено, какие традиционные источники информации можно использовать при создании RussNet, какова стратегия наполнения отдельных частей RussNet, были рассмотрены различные лексико-семантические поля.

Стандартная процедура построения wordnet предполагает обращение к частотным словарям языка для выбора употребительной лексики, с которой обычно начинается построение тезауруса. Наполнение структур RussNet осуществляется следующим образом: сначала исследуется ядро лексики русского языка – наиболее частотные слова с наиболее общим значением, затем полученные иерархические структуры расширяются за счет менее употребительной лексики.

Исследование семантических связей лексикализованных понятий осуществляется в рамках лексико-семантических групп, набор и состав которых определяется составом ядерной части тезауруса, при этом используются методы дефиниционного, контекстного и деривационного анализа.

В результате проведенных исследований лексико-семантических полей и их структуры нами было выявлено около 15 000 лексико-семантических вариантов, отредактировано и представлено в xml-формате – 5 000 синсетов с общим количеством связей около 8 000. Особую сложность при редактировании составляет согласование отдельных фрагментов RussNet-описания, для решения этой задачи было разработано программное обеспечение, позволяющее находить пересечения синсетов, несоответствия в дефинициях синсетов, соотносить разные лексико-семантические варианты одного слова. Для ввода отредактированного материала в компьютерную базу данных мы используем систему VisDic<sup>2</sup>, разработанную в чешской группе BalkaNet.

### **Структура RussNet**

Следуя принципам построения словарей WordNet и EuroWordNet, сформировавшихся в современной компьютерной лингвистике, RussNet имеет следующие структурные характеристики:

---

<sup>1</sup> Подробнее о результатах исследований см.: Материалы к компьютерному тезаурусу лексики русского языка / Сост. И. В. Азарова, О. А. Митрофанова. СПб., 2002. 232 с.; *Azarova I., Mitrofanova O., Sinopalnikova A., Yavor-skaya M., Oparin I.* RussNet: Building a Lexical Database for the Russian Language // Workshop Proceedings: Workshop on WordNet Structures and Standardisation, and how these affect Wordnet Application and Evaluation. 28th May 2002. Las Palmas de Gran Canaria, 2002. P. 60–64.

<sup>2</sup> *Pavelek T., Pala K.* VisDic – A New Tool for WordNet Editing // First International WordNet Conference. Mysore, 2002. P. 192-195.

- состоит из четырех взаимосвязанных файлов для основных частей речи: существительных, глаголов, прилагательных и наречий;
- основными единицами структуры RussNet, на которых задаются семантические отношения, являются синонимические ряды (синсеты);
- синсеты связаны между собой различными семантическими отношениями: парадигматическими и синтагматическими.

Дополнительно в структуру RussNet нами внесены следующие элементы, позволяющие фиксировать специфическую для русского языка информацию:

- расширен набор собственно семантических отношений за счет введения инхоативных отношений и выделения конверсии в самостоятельное отношение;
- вводится набор семантико-деривационных и семантико-грамматических отношений;
- основными единицами структуры RussNet, на которых задаются семантико-деривационные и семантико-грамматические отношения, являются составляющие синсетов – лексико-семантические варианты слов;
- введено описание валентностной структуры глаголов, включающее как семантические, так и поверхностно-грамматические атрибуты;
- прилагательные организованы в иерархические структуры на основе их сочетаемости с существительными.

Между RussNet и его прототипами, кроме структурных, существуют еще некоторые содержательные различия, касающиеся трактовки основных семантических отношений, что описано ниже.

### **Синонимия**

Синонимические отношения, связывающие слова по общности значений, являются базовыми для всех словарей типа WordNet. В рамках проектов WordNet и EuroWordNet синонимия определяется через понятие взаимозаменяемости. «Два слова (выражения) считаются синонимами, если существует хотя бы один контекст С, в котором замена одного слова другим не приводит к изменению истинностного значения» (Miller et al., 1990). Однако, на практике использовать этот критерий оказывается довольно сложно: взаимозаменяемость в контексте не всегда связана с общностью значений. Для слов любой части речи можно подобрать такой контекст, в котором будут взаимозаменяемы слова, семантически не связанные друг с другом. Например, метаязыковое употребление слов: «*Идти – это глагол. Молчать – это глагол*» или контекстуальная синонимия: «*во всем была сладкая и горькая грусть*» (И. Бунин). Вместе с тем некоторые синонимы не являются взаимозаменяемыми в контексте из-за особенностей синтаксической (*начать издавать, приступить к изданию, приняться за издание*) или же лексической сочетаемости (*актер играл; певец, танцор исполнил; актер представлял [устар.]*)<sup>1</sup>.

В силу этих причин мы используем критерий взаимозаменяемости как дополнительный по отношению к критерию семантической близости, который проверяется при дефиниционном анализе: установлении идентичности словарных определений или взаимной отсылке в синонимических определениях. Следовательно, в рамках RussNet отношение синонимии устанавливается между лексико-семантическими вариантами слов, которые

- принадлежат одной части речи,
- имеют сходные значения,
- могут быть взаимозаменяемы в контексте.

Использование данных словарных дефиниций обеспечивает относительную объективность нашего анализа, однако может затруднить разграничение синонимических и гипонимических отношений, поскольку не всегда ясно, какой тип определения представлен в

<sup>1</sup> Апресян Ю. Д. Лексическая семантика. Синонимические средства языка. Т. 1. М., 1995.

словарной статье. Поскольку синонимия является симметричным отношением: если А является синонимом В, то В является синонимом А; для верификации отношения синонимии мы используем двусторонние имплицативные тесты, т. е. проверяем истинность двух причинно-следственных конструкций, построенных на основе реальных контекстов употребления слов:

*Слово<sub>А</sub> в контексте С, следовательно, слово<sub>В</sub> в контексте С (истина).*

*Слово<sub>В</sub> в контексте С, следовательно, слово<sub>А</sub> в контексте С (истина).*

Например, пусть слово<sub>А</sub> = *опасаться*, слово<sub>В</sub> = *остерегаться*, контекст С = *Слуги \_\_\_\_\_ подслушивать мои разговоры* (О. Авраменко), тогда *Слуги опасаются подслушивать мои разговоры, следовательно, слуги остерегаются подслушивать мои разговоры (истина). Слуги остерегаются подслушивать мои разговоры, следовательно, слуги опасаются подслушивать мои разговоры (истина).* Таким образом, в данном значении слова *опасаться* и *остерегаться* являются синонимами.

Такой подход к определению синонимии приводит к тому, что в RussNet в рамках одного синсета объединяются следующие разновидности синонимов:

- абсолютные синонимы – слова, тождественные по значению и взаимозаменяемые в любом контексте, например, *гиппотам – бегемот, языкознание – языковедение – лингвистика*;
- фонетические и морфологические дублиеты – слова, отличающиеся написанием, разница в значениях которых зачастую не поддается определению, например, *камешек – камушек, снежный – снеговой, волочь – волочить*;
- стилистические синонимы – слова со сходными значениями, но разной стилистической окраской, которая не позволяет заменять их в одном контексте, например, *глаза – очи – зенки*; в рамках синсета такие слова сопровождаются соответствующими стилистическими пометами (*разг., книжн., устар.* и т.п.);
- деривационные синонимы (см. ниже).

Для придания более четкой структуры синсету мы рассматриваем один из его членов как доминанту, наименее закрепленный синтагматически синоним-представитель синсета с наиболее общим значением и нейтральной стилистической окраской. При выделении доминанты мы опираемся, прежде всего, на статистические критерии: представитель синсета в силу своих нейтральных свойств употребляется в текстах корпуса значительно чаще, чем остальные члены синсета. Таким образом, в RussNet члены синсета оказываются упорядочены в соответствии с частотой их появления в текстах: от нейтральных абсолютных синонимов и дублиетов – к эмоционально и стилистически окрашенным.

### *Гипонимия*

Гипонимические отношения являются основным средством формирования структуры тезауруса. В рамках RussNet гипонимия определяется как семантическое отношение, связывающее

- синсеты, относящиеся к одной части речи;
- синсеты, между содержанием понятий которых существует отношение семантического включения, т. е. значение гиперонима полностью включено в значение гипонима, например, значение синсета с доминантой *бояться<sub>2</sub>* входит в качестве основного семантического компонента в значение синсета {*опасаться, остерегаться*};
- гиперонимом является непосредственный родовой термин.<sup>1</sup>

Мы используем при выявлении гипонимических отношений тот факт, что в текстах гиперонимы употребляются значительно чаще по сравнению со своими гипонимами, поскольку, родовые термины, обладая более общим значением и более широкой сочетаемостью, регулярно выступают в роли заместителей видовых. Кроме того, существуют типичные конструкции для родовидовых отношений типа: «А – это [разновидность] В», которые

---

<sup>1</sup> В принципе, у синсета может быть несколько потенциальных гиперонимов, например, синсет *нести<sub>1</sub>* является видовым по отношению к синсетам *идти* и *держаться<sub>2</sub>*.

регулярно встречаются в текстах и часто используются при построении определений типа «genus proximum + differentia specifica» в толковых словарях.

Поскольку гипонимия является несимметричным отношением: если А является гиперонимом В, то В является гипонимом А, то для верификации гипонимических отношений используются тесты, аналогичные тестам для верификации синонимии, с учетом несимметричности отношения:

*Член синсета<sub>А</sub> в контексте С, следовательно, член синсета<sub>В</sub> в контексте С (истина).*

*Член синсета<sub>В</sub> в контексте С, следовательно, член синсета<sub>А</sub> в контексте С (ложь).*

Например, пусть **синсет<sub>А</sub>** = {роза}, **синсет<sub>В</sub>** = {цветок}, **контекст С** = Мне подарили \_\_\_\_, тогда *Мне подарили розы, следовательно, мне подарили цветы (истина). Мне подарили цветы, следовательно, мне подарили розы (ложь).* Таким образом, синсет {роза} является гипонимом для синсета {цветок}.

### **Антонимия**

Отношение антонимии связывает синсеты, которые противопоставляются по наиболее существенному компоненту значения<sup>1</sup>. В рамках концепции Дж. Миллера считается, что логический анализ не дает возможности выявить антонимы, и предлагается ориентироваться на результаты ассоциативного противопоставления слов. В RussNet мы совмещаем психолингвистический и логический подходы к установлению антонимических отношений. Для этой цели проводится (1) анализ материалов ассоциативных словарей; (2) лингвистический и логический анализ словарных определений, построенных на основе антонимических отсылок; (3) выявление конструкций противопоставления как вспомогательных контекстуальных маркеров антонимов. Поскольку противопоставлены могут быть любые пары нетождественных значений (*Петр не спит, а ест*), то антонимы выделяются как противоположности в рамках некоторой общности – одной лексико-семантической группы, общего гиперонима. Возможны различные виды антонимических противопоставлений с нейтрализацией оппозиции (*горячий – теплый – холодный*) или без нее (*удачный – неудачный*).

Выделение доминанты синонимического ряда, помогает нам решить проблему представления непосредственной и опосредованной антонимии. Доминанты двух синсетов связываются отношением антонимии напрямую, остальные члены синсета противопоставляются косвенным образом, опосредованно – через доминанту.

Поскольку антонимия является симметричным отношением, т. е. если А является антонимом В, то В является антонимом А, для верификации антонимии мы используем двусторонние имплицативные тесты с отрицанием.

*Член синсета<sub>А</sub> в контексте С, следовательно, не член синсета<sub>В</sub> в контексте С (истина).*

*Член синсета<sub>В</sub> в контексте С, следовательно, не член синсета<sub>А</sub> в контексте С (истина).*

Например, пусть **синсет<sub>А</sub>** = {подтверждать, доказывать}, **синсет<sub>В</sub>** = {опровергать}, **контекст С** = Он \_\_\_\_ мои опасения (Ю. Латынина), тогда *Он подтвердил мои опасения, следовательно, он не опроверг мои опасения (истина); Он опроверг мои опасения, следовательно, он не подтвердил мои опасения (истина).* Таким образом, мы приходим к заключению, что синсеты {подтверждать, доказывать} и {опровергать} являются антонимами.

Таким образом, мы устанавливаем антонимические отношения между синсетами, которые

- принадлежат одной части речи,
- имеют общую часть значения – принадлежат одной лексико-семантической группе, имеют общий гипероним,
- противопоставляются по существенному компоненту значения,
- взаимозаменяемы в контексте под отрицанием или в противительной конструкции.

---

<sup>1</sup> По аналогии с EuroWordNet мы трактуем антонимию как семантическое отношение между синсетами, а не лексико-семантическими вариантами слов.

В словаре RussNet отражены и другие собственно семантические отношения, устанавливаемые между синсетам: меронимия, конверсия, лексический вывод (каузация и пресуппозиция), ролевые отношения<sup>1</sup>. Процедуры их выявления во многом аналогичны применяемым при установлении отношений синонимии, антонимии и гиперонимии: привлекаются источники лингвистической информации<sup>2</sup>, проводятся тесты, построенные по единому для RussNet принципу.

### **Семантико-грамматические и семантико-деривационные отношения**

Для лексики русского языка, особенно существительных и глаголов, характерна многообразия и высокая продуктивность словообразовательных моделей (доля мотивированной лексики составляет в русском языке около 85 %). В большинстве случаев набора традиционных для wordnet-словарей отношений оказывается недостаточно для того, чтобы описать связи между словом и его производными, поэтому в рамках RussNet мы вводим ряд особых семантических отношений, предполагающих обязательное формальное выражение на грамматическом или деривационном уровне. Специфика данных отношений заключается в том, что они устанавливаются между членами синсетов – лексико-семантическими вариантами слов.

**Деривационная синонимия** – отношение, связывающее нейтральное слово и его экспрессивные, эмоционально окрашенные дериваты, например, *старик*<sub>1</sub> – *старикан*, *старик* – *старичок*. Такие слова обладают тождественным референциальным значением, различия между ними касаются только отношения говорящего к референту. Слова с такой коннотацией, как правило, не взаимозаменяемы в контексте. Экспрессивные синонимы включаются в тот же синсет, что и нейтральное слово (например, {*старик*, *старикан* (*нейор.*), *старичок*<sub>1</sub> (*мелиор.*)}), при этом дериваты сопровождаются пометами, характеризующими их коннотативное значение: *нейор.* – отрицательная коннотация; *мелиор.* – положительная коннотация.

**Деривационная гипонимия** отличается от деривационной синонимии тем, что деривационный аффикс придает производному слову дополнительный смысловой оттенок, не сводимый к эмоционально-экспрессивному или стилистическому отношению: *белый* – *беловатый* (суффикс *-оват-* указывает на слабую интенсивность признака), *старик* – *старичок*<sub>2</sub> (в данном случае *-ок* указывает на изменение референции – «маленький старичок»).

В зависимости от того, какой дополнительный компонент присутствует в значении производного слова, мы выделяем следующие подтипы деривационной гипонимии:

- диминутивная (*цветок* – *цветочек*, *книга* – *книжица*)
  - минимальная (*маленький* – *малюсенький*)
- аугментативная (*рука* – *ручища*, *дом* – *домина*, *высокий* – *превысокий*)
  - максимальная (*последний* – *распоследний*)

**Деривационные ролевые отношения** охватывают ряд семантико-деривационных связей типа «глагол – актанта»: «глагол – объект», «глагол – деятель», «глагол – инструмент» и др., например, *сеять* – *сеянец*, *сеятель*, *сеялка*. Необходимость введения данных отноше-

---

<sup>1</sup> Подробнее о семантических отношениях в RussNet см. *Азарова И. В., Митрофанова О. А., Синопальникова А. А., Ушакова А. А., Яворская М. В.* Разработка компьютерного тезауруса русского языка типа WordNet // Доклады научной конференции «Корпусная лингвистика и лингвистические базы данных» / Под ред. А. С. Герда. СПб., 2002. С. 6-18.

<sup>2</sup> Караулов Ю. Н., Уфимцева А. А. и др. Русский ассоциативный словарь. Т. 1-6. М., 1994, 1996, 1998; Новый объяснительный словарь русских синонимов русского языка / Под ред. Ю. Д. Апресяна. Вып. 1. М., 1997. Вып. 2. М., 2000; Ожегов С. И. Словарь русского языка. М., 1984; Ожегов С. И., Шведова Н. Ю. Толковый словарь русского языка. М., 1992; Словарь ассоциативных норм русского языка / Под ред. А. А. Леонтьева. М., 1977; Словарь русского языка / Под ред. А. П. Евгеньевой. Т. 1-4. М., 1985-88; Словарь синонимов русского языка / Под ред. А. П. Евгеньевой. Т. 1-2. Л., 1971; Словарь современного русского литературного языка. Т. 1-17. М.-Л., 1991. Словарь сочетаемости слов русского языка / Под ред. П. Н. Денисова, В. В. Морковкина. М., 2002; Толковый словарь русских глаголов: Идеографическое описание. Английские эквиваленты. Синонимы. Антонимы / Под ред. Л. Г. Бабенко. М., 1999; Частотный словарь русского языка / Под ред. Л. Н. Засориной. М., 1977; Шаров С. А. Список наиболее частотных слов русского языка. [www.artint.ru/projects/frqlist](http://www.artint.ru/projects/frqlist)

ний обусловлена тем, что дериваты наследуют не только семантические, но и синтаксические свойства, например, рамки управления (*бороться за правое дело – борьба за правое дело – борец за правое дело*).

В RussNet мы трактуем противопоставление глаголов по виду как **семантико-грамматическое отношение**, которое связано со сложной структурой оппозиций значений (достигший предела – стремящийся к пределу, не имеющий предела и пр.), обладает чертами грамматической категории, связано с определенными деривационными процессами. В RussNet устанавливаются отношения перфективации (*делать – сделать, ссориться – поссориться*) и имперфективации (*помочь – помогать, дать – давать*), которые присоединяют члены видовой пары к структуре семантических связей словаря.

### **Построение дефиниций**

После окончательного формирования синсетов и установления их семантических связей мы строим дефиниции, призванные наиболее полно отразить значение синсетов. При этом мы используем материалы лексикографических источников как дополнительное средство. Большая часть определений опирается на гипонимические отношения, на основе которых строятся родовидовые определения. Основной задачей является отображение в дефинициях синсетов-согипонимов дифференциальных признаков и соотнесенность с общим родовым понятием (*шутить<sub>1</sub>* – «говорить с кем-л. несерьезно, чтобы развлечься», *отвечать* – «говорить что-л. кому-л., отвечая на вопрос»). Другими важными отношениями, используемыми при конструировании дефиниций являются «часть – целое» (*голень* – «часть ноги от колена до стопы») и отношения лексического вывода, в частности, каузативное (*злить* – «вызывать злость» или «заставлять злиться»).

### **Заключение**

Оценивая современное состояние RussNet, можно сказать, что объем представленной лексики покрывает ядро словарного состава современного русского языка, отражая наиболее значимые для лексикона структурные отношения. Созданный лингвистический ресурс объединяет в себе черты разноплановых лингвистических описаний (традиционных лексикографических источников, с одной стороны, и моделей языкового сознания, с другой стороны), представление данных в компьютерной форме дает возможность использовать RussNet в качестве компонента разнообразных информационных систем.