



OSTIS-2011

(Open Semantic Technologies for Intelligent Systems)

УДК 372.881.111.1

АНГЛОЯЗЫЧНЫЕ ОНТОЛОГИИ В ПРОФИЛЬНОМ ЛИНГВИСТИЧЕСКОМ ОБРАЗОВАНИИ

О.М. Корчажкина (*olgakomax@gmail.com*)

ГОУ Центр образования № 1678 «Восточное Дегунино», г. Москва,
Российская Федерация

Обсуждаются перспективы использования лингвистических онтологий WordNet 2.1 и WordNet 3.0, разработанных на базе национального корпуса британского варианта английского языка, для обучения лексике английского языка и проведения лингвистических исследований в классах филологического профиля старшей школы.

Ключевые слова: лексика, лингвистическая онтология, синсет, языковой корпус.

Введение

Основой современной концепции обучения иностранным языкам (ИЯ) в средней школе является коммуникативный подход, что предполагает формирование лексической компетентности учащихся на базе не только формальной, но и в равной степени функциональной стороны языка.

Набор лексических единиц для активного и пассивного усвоения на всех уровнях обучения ИЯ формируется по принципу баланса между формой и функцией, когда лексические единицы рассматриваются как компоненты языковой системы, имеющие свои функциональные особенности. Однако на практике нередко отмечается крен в сторону одного из аспектов в ущерб другому. Если обучение лексике происходит преимущественно с опорой на языковой аспект, то это может вызвать у учащихся определённые трудности в речевом общении, когда употребление лексических единиц не отвечает замыслу высказывания и может неоднозначно восприниматься собеседником, особенно носителем ИЯ. Акцент на речевой аспект при обучении лексике ИЯ, напротив, делает коммуникацию в целом успешной, однако может привести к упрощению высказывания за счёт неоправданного «клиширования» речи, особенно её устной формы. Искусственное разведение двух необходимых составляющих при освоении иноязычной лексики снижает общий уровень владения речью. Это происходит тогда, когда лексика рассматривается и поэтому осваивается как набор слабо связанных между собой единиц вне языковой системы ИЯ, включающей и семантический, и прагматический аспект.

В профильных классах старшей школы расширение иноязычного лексического поля учащихся должно происходить в направлении не только количественных, но и качественных показателей, а именно: за счёт освоения глубинных иерархических слоёв в представлении лексических единиц, что предполагает выявление семантических, логических и контекстуальных связей¹, а также уровней прагматического взаимодействия лексем ИЯ.

Эти связи, заложенные в глубинные структуры языка и являющиеся изначально категориями имплицитными, выражаются в таких системных отношениях как синонимия/антонимия, гиперонимия/гипонимия, логические, родовидовые связи, а также семантических отношениях между морфологически связанными лексемами. Известно, что ни

¹ Контекст - указательный минимум окружения лексической единицы в тексте, помогающий сделать нужный выбор её значения [Елисеева, 2009].

один словарь ИЯ, будь то толковый или двуязычный, традиционный или электронный, не содержит той обширной информации о системных связях между лексическими единицами, которая необходима учащимся на продвинутом уровне для более эффективного овладения иноязычной лексикой. Такие связи в неявном виде заложены в национальные языковые корпуса – электронные информационно-справочные системы, основанные на собрании текстов на некотором языке и представляющие данный язык на определенных этапах его существования во всём многообразии жанров, стилей, территориальных и социальных вариантов и т. п. (см., например [Национальный корпус русского языка]). Однако изъять из языковых корпусов сложные семантические отношения между лексическими единицами и оформить их традиционными способами составления словарей в виде словарных статей, обычных структурных единиц толковых словарей, задача сложная, да и непродуктивная.

Кроме того, необходимость изучения глубинных системных взаимосвязей между лексическими единицами ИЯ связана с проблемой организации и проведения лингвистических исследований в профильных классах средней школы, что является одной из форм приобщения учащихся к научно-исследовательской деятельности. Однако выявление подобных семантических, логических и контекстуальных отношений между лексемами любого языка не под силу даже продвинутому учащемуся, поскольку возможно только в результате концептуального анализа больших массивов устных и письменных текстов, создаваемых его носителями.

Поэтому при обучении лексике или организации лингвистических исследований на профильном уровне учителю очень часто приходится создавать такие упражнения и задания и так структурировать учебный материал, чтобы он был посилен для освоения школьников. Тем не менее, даже в этих случаях учащимся приходится «вручную» устанавливать сложные семантические связи между лексемами, причём преимущественно в пределах одного текста, что не всегда приводит к желаемому результату и чаще всего заканчивается полной потерей интереса к исследовательской деятельности. Однако подобная рутинная работа может быть автоматизирована за счёт привлечения новых информационных технологий, что повысит интерес учащихся к проведению лингвистических исследований и сделает этот процесс более эффективным.

1. Онтологии как системы представления знаний

С разработкой поисковых и структурных инструментов новой версии сети, семантического веба, Web 3.0, имеющей целью решение проблемы семантической интеграции информационных ресурсов и повышения эффективности информационного поиска, стала возможной новая система представления знаний, выступающая в качестве ведущей парадигмы структурирования информационного контента – с помощью онтологий. Термин «онтология» имеет множество определений (см., например, [Гаврилова и др., 2008], [Добров и др., 2009], [Иванников и др., 2008]), однако в контексте настоящего рассмотрения мы будем придерживаться следующего толкования. *Онтология – это иерархическая структурная схема некоей семантической системы, которая устанавливает концептуальные² связи между её компонентами, представляющими собой термины и понятия в некоторой предметной области.*

Онтологии строятся в виде разветвлённого словаря концептов данной предметной области, с помощью которых осуществляется представление и обмен знаниями по интересующей проблеме [Иванников и др., 2008], поэтому становится возможным их использование как при изучении данной предметной области, так и при освоении языка, выбранного для составления онтологии. При обучении английскому языку неоценимую помощь могут оказать лексические онтологии, разработанные англоязычными IT-специалистами.

Одной из таких онтологий является англоязычный ресурс WordNet, разработанный в Принстонском университете США на основе национального корпуса британского варианта английского языка (the British National Corpus [British National Corpus]). База данных WordNet представляет собой семантически структурированный лингвистический гиперресурс,

² Концептуализация – абстрактная модель явлений (процессов) в мире, составленная посредством определения понятий, существенных для описания данных явлений (процессов) [Иванников и др., 2008].

охватывающий около 155 тысяч слов и словосочетаний, объединённых в 117 тысяч понятий, причём общее число пар «лексема-значение» насчитывает 200 тысяч [Добров и др., 2009]. В основу ресурса положены отношения синонимии между лексемами, принадлежащими четырём основным частям речи: существительным, прилагательным, глаголам и наречиям. Онтология основана на принципе контекстуальной синонимии: совокупности когнитивных синонимов выступают в виде лексикализованных концептов, то есть понятий, формирующих англоязычную картину мира [Добров и др., 2009]. В настоящее время существуют две рабочие версии онтологии – WordNet 3.0 и WordNet 2.1.

В качестве примера приведём представленность концепта *database* в трёх ресурсах, основанных на лингвистической онтологии WordNet 3.0.

1. Онтологический словарь «Free English online dictionary based on WordNet 3.0» [Free English online dictionary] даёт основные, хотя и ограниченные сведения об искомом концепте. Он позволяет проследить семантические связи слов на уровне гиперонимов, гипонимов, синонимов и определить область применения концепта согласно предметной классификации.

English *noun*: **database**

1.	database (<i>communication</i>) an organized body of related information
	Broader (hypernym) info , information
	Narrower (hyponym) computer database , electronic database , electronic information service , list , listing , on-line database , subdata base

English *noun*: **computer database**

1.	computer database (<i>communication</i>) (computer science) a database that can be accessed by computers
	Synonyms electronic database , electronic information service , on-line database
	Broader (hypernym) database
	Narrower (hyponym) lexical database , object-oriented database , relational database
	Domain category computer science , computing

Как видно из представленного примера, семантическое поле концепта *database* лежит в предметной области «информация» (*info*, *information*), что является его гиперонимом (расширенным значением лексемы). Если запросить онтологию подобрать более узкое значение концепта *database* (гипоним), то получим список следующих устойчивых словосочетаний (коллокаций): *computer database*, *electronic database*, *electronic information service*, *list*, *listing*, *on-line database*, *subdata base*, среди которых представлены не только словосочетания с исходным концептом (*computer database*, *electronic database*), но и смежные, в том числе и синонимичные значения, такие как *electronic information service*, *list*, *listing*, *on-line database*, *subdata base*.

Онтологический словарь имеет разветвлённую систему гиперссылок, что обеспечивает удобство его использования. Вводя в поисковую систему словаря искомый концепт, можно получить весьма глубокое «проникновение» в систему понятий данного концепта или смежных концептов. Если, например, проследить далее семантику одного из гипонимов концепта *database* - *computer database*, переходя по соответствующей гиперссылке, то можно определить пять уровней его семантики: толкование в пределах его семантического поля, набор синонимов, гипероним, набор гипонимов и предметную классификацию концепта (Domain category).

Таим образом, онтологический словарь «Free English online dictionary based on WordNet 3.0» даёт возможность проследить простейшие связи между концептами внутри некоторой семантической области по следующим признакам: семантическое поле, предметная классификация, толкование, синонимы, гиперонимы и гипонимы.

2. Поисковая система «WordNet Search - 3.0» [WordNet Search - 3.0].

На рис. 1 можно видеть фрагмент интерфейса системы с полями для ввода лексемы и выбора способов представления её свойств.



Рисунок 1 - Фрагмент интерфейса поисковой системы «WordNet Search - 3.0»

Для рассматриваемого нами концепта *database* результаты поиска будут выглядеть следующим образом:

Key: "S:" = Show Synset (semantic) relations, "W:" = Show Word (lexical) relations

S: (n) **database** (an organized body of related information)

direct hyponym / full hyponym

direct hypernym / inherited hypernym / sister term

S: (n) information, info (a message received and understood) и т.д.

Не рассматривая подробно поисковую систему самой онтологии «WordNet Search - 3.0», отметим, что она позволяет определить более тонкие семантические слои заданного концепта, чем предыдущая система в виде онтологического словаря: семантические и лексические связи (semantic/lexical relations) исходного концепта, его прямые/полные гипонимы (direct hyponym/full hyponym), прямые/производные гиперонимы (direct/inherited hypernym), частичные голонимы (member holonym), а также согласованные/парные понятия (sister term). Переход по гиперссылкам позволяет раскрыть полный набор свойств искомого концепта.

3. Система «WordNet Vocabulary Helper» [11] обладает наглядным интерфейсом, выводящим пользователя на четыре семантических уровня, указанные в содержании онтологической статьи - общие свойства искомого концепта (overview), гипонимы (hyponyms), синонимы/гиперонимы (synonyms/hypernyms), согласованные/парные понятия (coordinate/sister terms):

Contents

- Overview of noun *database*
- Hyponyms of noun *database*
- Synonyms/Hypernyms (Ordered by Estimated Frequency) of noun *database*
- Coordinate Terms (sisters) of noun *database*

Переход по гиперссылкам раздела **Contents** позволяет получить следующую информацию об искомом концепте:

- **Overview of noun *database***

The noun *database* has 1 sense (no senses from tagged texts)

1. **database** -- (an organized body of related information)

- **Hyponyms of noun *database***

1 sense of **database**

database – (an organized body of related information)

list, listing – (a **database** containing an ordered array of items (names or topics))

electronic **database**, on-line **database**, computer **database**, electronic information service ((computer science) a **database** that can be accessed by computers)

subdata base – (a subset of data in a **database** that are used in a specific application)

- **Synonyms/Hypernyms (Ordered by Estimated Frequency) of noun *database***

1 sense of **database**

database – (an organized body of related information)

information, info – (a message received and understood)

- **Coordinate Terms (sisters) of noun *database***

1 sense of **database**

database – (an organized body of related information)

- information, info -- (a message received and understood)
- ammunition -- (information that can be used to attack or defend a claim or argument or viewpoint; *“his admission provided ammunition for his critics”*)
- factoid -- (something resembling a fact; unverified (often invented) information that is given credibility because it appeared in print)
- misinformation -- (information that is incorrect)
- material -- (information (data or ideas or observations) that can be used or reworked into a finished form; *“the archives provided rich material for a definitive biography”*)
- details, inside information -- (true confidential information; *“after the trial he gave us the real details”*)
- fact -- (a statement or assertion of verified information about something that is the case or has happened; *“he supported his argument with an impressive array of facts”*)
- format, formatting, data format, data formatting -- (the organization of information according to preset specifications (usually for computer processing))
- gen -- (informal term for information; *“give me the gen on your new line of computers”*)

database -- (an organized body of related information)

- news, intelligence, tidings, word -- (information about recent and important events; *“they awaited news of the outcome”*)
- news -- (informal information of any kind that is not previously known to someone; *“it was news to me”*)
- nuts and bolts -- (detailed practical information about how something works or how something can be accomplished)
- intelligence, intelligence information -- (secret information about an enemy (or potential enemy); *“we sent out planes to gather intelligence on their radar coverage”*)
- confirmation -- (information that confirms or verifies)
- insider information -- (important information about the plans or condition of a corporation that has not been released to the public; use for personal profit is illegal)
- secret, arcanum -- (information known only to a special group; *“the secret of Cajun cooking”*)
- secret -- (something that should remain hidden from others (especially information that is not to be passed on); *“the combination to the safe was a secret”; “he tried to keep his drinking a secret”*)
- propaganda -- (information that is spread for the purpose of promoting some cause)
- course of study, program, programme, curriculum, syllabus -- (an integrated course of academic studies; *“he was admitted to a new program at the university”*)
- news -- (information reported in a newspaper or news magazine; *“the news of my death was greatly exaggerated”*)
- evidence -- ((law) all the means by which any alleged matter of fact whose truth is investigated at judicial trial is established or disproved)
- readout, read-out -- (the information displayed or recorded on an electronic device)
- tabulation, tabular matter -- (information set out in tabular form)
- skinny -- (confidential information about a topic or person; *“he wanted the inside skinny on the new partner”*)
- stuff -- (information in some unspecified form; *“it was stuff I had heard before”; “there's good stuff in that book”*)
- report card, report -- (a written evaluation of a student's scholarship and deportment; *“his father signed his report card”*)

Очевидно, что данный ресурс предоставляет пользователю общие свойства искомого концепта, гипонимы и синонимы/гиперонимы практически в том же объёме, что и онтологический словарь *Free English online dictionary based on WordNet 3.0*. Однако наиболее интересные с точки зрения смежных предметных областей сведения об искомом концепте содержатся в разделе **Coordinate Terms (sisters) of noun database**, где переданы различные толкования гиперонима концепта *database* -- лексемы *information*, значительно расширяющие представление учащихся об исследуемом семантическом поле. В системе «WordNet Vocabulary Helper» данный раздел представлен наиболее полно.

2. Лингвистическая онтология WordNet 2.1

Онтология WordNet 2.1 [WordNet 2.1] структурирована несколько иначе. Лексика в данной онтологии организована в логические группы, названные *синсетами* (synsets, или synonymous sets). В каждом синсете лексические единицы (слова и словосочетания) сгруппированы по принципу взаимозаменяемости в пределах определённого контекста. Каждый синсет состоит из набора синонимичных слов или словосочетаний, а также групп семантических указателей (semantic pointers), отображающих отношения между этим синсетом и другими синсетами онтологии.

Синсеты, являющиеся основными структурными единицами онтологии, бывают головными и периферийными (head/satellite synsets), причём головные синсеты связаны с одним или более периферийными синсетами. Головные синсеты, в свою очередь, объединяются в кластеры прилагательных (adjective clusters), которые организованы вокруг антонимических пар или троек прилагательных. Головной синсет в кластере – это синсет, который содержит хотя бы одно прилагательное, имеющее прямой антоним, а периферийный синсет – это синсет, в котором заключён концепт, сходный по значению с концептом, содержащимся в его головном синсете.

Отношения между синсетами онтологии представлены двумя типами указателей, или связей – лексическими и семантическими. Лексические отношения устанавливаются между семантически связанными словоформами, а семантические – между разными значениями одной лексемы. Синсеты структурированы таким образом, что отношения между лексемами внутри каждого синсета описываются не только по принципу синонимии/антонимии, но и с привлечением отношений других типов: краткой/полной гипонимии (brief/full hyponyms), прямой/производной гиперонимии (direct/inherited hypernyms), меронимии/голонимии (meronymy/holonymy), принадлежности к определённой предметной области (domain category) и т.д., поэтому одна и та же лексическая единица заложена в онтологии в более чем один синсет и представлена в виде более чем одной синтаксической категории.

Создатели онтологии WordNet 2.1 сравнивают её структуру с осью, на которую нанизаны велосипедные колёса (рис. 2). Сама ось символизирует отношения антонимии, положенные в основу онтологии, а синсеты – это колёса, причём головные синсеты располагаются непосредственно на оси, а периферийные – отходят от неё по радиусам (спицам колеса).

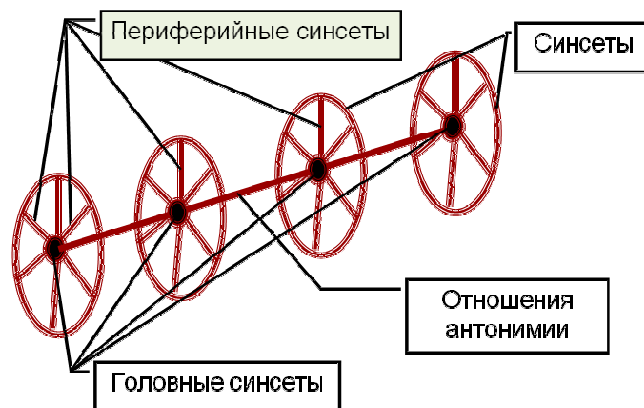


Рисунок 2 - Модель онтологии WordNet 2.1

Рассмотрим кратко работу с онтологией WordNet 2.1. Браузер онтологии представляет собой окно с тремя рядами панели инструментов (рис.3). В первом ряду находятся кнопки *File*, *History*, *Options* и *Help*. Во втором ряду расположено окно для ввода искомой лексемы (командная строка браузера), а в третьем ряду – вспомогательное окно *Senses*. После ввода исходной лексемы в командную строку браузера и нажатия кнопки *Enter* на клавиатуре компьютера в главном окне онтологии открывается список всех значений лексемы, детализированных по частям речи и сопровождаемых синонимами и примерами в виде словосочетаний или предложений (см. рис.3).

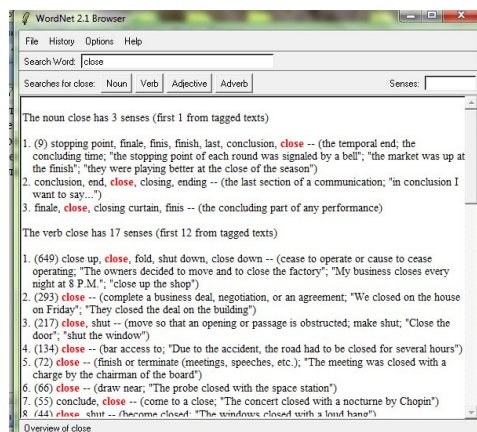


Рисунок 3 - Полный набор значений искомой лексики

Для получения более подробных сведений о семантических значениях исходной лексики в онтологии предусмотрена отдельная градация по частям речи, к которой пользователь переходит по кнопкам, появляющимся на панели браузера после введения исходной лексики: *Noun*, *Verb*, *Adjective* и *Adverb*. Выбирая указатели в раскрывающихся списках для каждой части речи, пользователь может перейти к следующим отношениям (рис. 4):

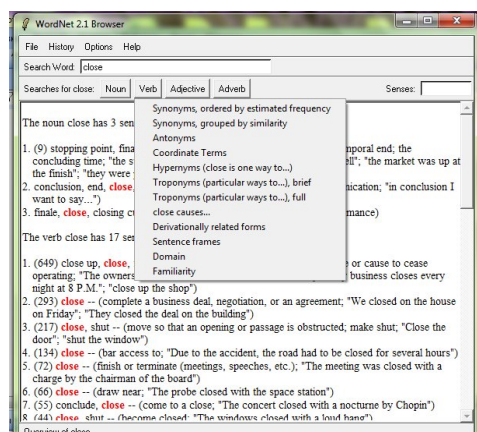


Рисунок 4 - Список семантических указателей для глаголов

1. для существительных:

Synonyms, ordered by estimated frequency / Coordinate terms / Hypernyms (*the noun* is a kind of ...) / Hyponyms (... is a kind of *the noun*), brief /full / Holonyms (*the noun* is a part of ...), inherited/regular / Meronyms (parts of *thenoun*), inherited/regular / Derivationally related forms /Domain / Familiarity (polysemy count);

2. для глаголов:

Synonyms, ordered by estimated frequency/grouped by similarity / Antonyms / Coordinate terms / Hypernyms (*the verb* is one way to ...) / Troponyms (particular ways to...), brief/full / Derivationally related forms / Sentence frames / Familiarity (polysemy count);

3. для прилагательных:

Synonyms/Related Nouns / Antonyms / Derivationally related forms / Familiarity;

4. для наречий:

Synonyms/Stem Adjectives / Antonyms / Domain / Familiarity.

Дополнительно к указанным отношениям некоторые лексемы имеют более конкретную детализацию, например: *go* entails doing ...; *fly* causes ...; *interesting* is a value of

Таким образом, переход по указателям даёт возможность выявить все лексемы, связанные с исходной лексемой лексико-семантическими отношениями, содержащимися в онтологии.

Интересна функция, которая называется *Find key words by substring* (найди ключевые слова с помощью метода извлечения данных substring). Она позволяет получить полный список словосочетаний, в которых встречается искомая лексема (рис. 5). Переход к ней осуществляется путём выбора соответствующей опции в раскрывающемся списке при нажатии кнопки *File*.

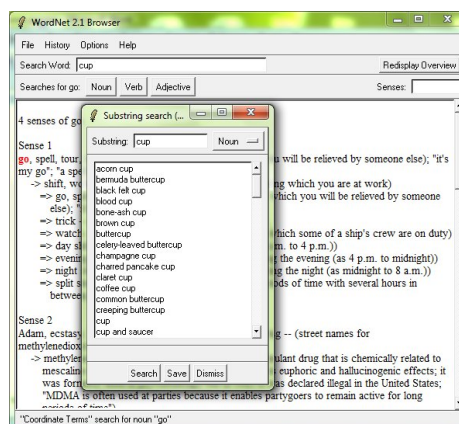


Рисунок 5 - Окно со списком устойчивых словосочетаний

В качестве примера приведём отношения, связывающие исходную лексему **permanent** с другими лексемами онтологии WordNet 2.1. Информация в главном окне выглядит следующим образом:

The noun **permanent** has 1 sense (no senses from tagged texts)

1. permanent wave, permanent, perm (a series of waves in the hair made by applying heat and chemicals)

The adj **permanent** has 2 senses (first 1 from tagged texts)

1. (13) permanent, lasting (continuing or enduring without marked change in status or condition or place; "permanent secretary to the president"; "permanent address"; "literature of permanent value")

2. permanent (not capable of being reversed or returned to the original condition; "permanent brain damage")

При этом на панели онтологии появляются две кнопки *noun* и *adjective*, предлагающие перейти к следующим отношениям: для существительного – Synonyms, ordered by estimated frequency / Coordinate terms / Hypernyms (**permanent** is a kind of ...) / Familiarity, а для прилагательного – Synonyms/Related Nouns / Antonyms / **permanent** is a value of ... / Derivationally related forms / Familiarity.

Например, отношения *Synonyms, ordered by estimated frequency* (a), *Hypernyms* (**permanent** is a kind of ...) (b) и *Familiarity* (c) следующим образом представлены в онтологии для лексемы **permanent** как существительного:

(a) permanent wave, permanent, perm -- (a series of waves in the hair made by applying heat and chemicals)

=> wave -- (a hairdo that creates undulations in the hair)

(b) permanent wave, permanent, perm -- (a series of waves in the hair made by applying heat and chemicals)

=> wave -- (a hairdo that creates undulations in the hair)

=> hairdo, hairstyle, hair style, coiffure, coif -- (the arrangement of the hair (especially a woman's hair))

=> hair -- (a covering for the body (or parts of it) consisting of a dense growth of threadlike structures (as on the human head); helps to prevent heat loss; "he combed his hair"; "each hair consists of layers of dead keratinized cells")

=> body covering -- (any covering for the body or a body part)

=> covering, natural covering, cover -- (a natural object that covers or envelops; "under a covering of dust"; "the fox was flushed from its cover")

- => natural object -- (an object occurring naturally; not made by man)
- => whole, unit -- (an assemblage of parts that is regarded as a single entity; "how big is that part compared to the whole?"; "the team is a unit")
- => object, physical object -- (a tangible and visible entity; an entity that can cast a shadow; "it was full of rackets, balls and other objects")
- => physical entity -- (an entity that has physical existence)
- => entity -- (that which is perceived or known or inferred to have its own distinct existence (living or nonliving))

(c) permanent used as a noun is very rare (polysemy count = 1)

А отношения *Derivationally related forms* (a) и *Familiarity* (b) для лексемы **permanent** как прилагательного выглядят так:

(a) permanent (vs. impermanent), lasting -- (continuing or enduring without marked change in status or condition or place; "permanent secretary to the president"; "permanent address"; "literature of permanent value")

RELATED TO->(noun) permanency#1

=> permanence, permanency -- (the property of being able to exist for an indefinite duration)

(b) permanent used as an adjective is rare (polysemy count = 2)

Как видно из приведённого примера, онтология WordNet 2.1 обладает многообразием лексико-семантических связей, полнотой представления информации, дополнительными функциями, отсутствующими в других онтологиях того же уровня, и удобна в использовании. Она свободно распространяется в виде отторгаемого ресурса объёмом менее 20 Мб и имеет подробную инструкцию. Всё это делает её одной из самых популярных онтологий в среде учёных-лингвистов, и, возможно, со временем к ней всё чаще будут обращаться учителя английского языка и их ученики.

Заключение

К сожалению, онтологии WordNet не предусматривают возможность лексико-семантического анализа произвольных англоязычных текстов, поскольку создавались совершенно для иных целей. Однако очевидно, что способ представления ими лексико-семантической информации позволяет проводить изучение и исследование огромного пласта англоязычной лексики не только на базе родовидовых отношений, но и на глубинных уровнях. Это даёт возможность учащимся усваивать концепты, составляющие основу языковой картины мира, близкой к картине мира носителей английского языка. Благодаря указанным характеристикам, которые не имеет ни один из известных словарей английского языка, онтологии WordNet 3.0 и WordNet 2.1 можно рассматривать как новые высокотехнологичные инструменты и с успехом использовать их для изучения лексики и проведения лингвистических исследований по английскому языку на продвинутом уровне.

Глоссарий терминов, используемых в иерархии онтологии WordNet 2.1

Adjective cluster (кластер прилагательных) – группа синсетов, содержащих прилагательные, которые организованы вокруг антонимических пар, или триплетов. *Кластер прилагательных* состоит из двух или более головных синсетов, представляющих концепты, связанные отношениями антонимии. Каждый головной синсет имеет хотя бы один периферийный синсет.

Antonyms, direct/indirect (прямые/непрямые антонимы): прямые антонимы – пара слов, между которыми существуют ассоциативные связи, являющиеся результатом частотности их употребления (в кластерах прилагательных прямые антонимы появляются только в головных синсетах); непрямой антоним – имя прилагательное в периферийном синсете, которое не имеет прямого антонима, однако через прямой антоним в головном синсете получает непрямые антонимы.

Coordinate terms (согласованные, или парные понятия) – существительные или глаголы, имеющие одни и те же гиперонимы.

Derivationally related forms (родственные формы слов) – понятия, выраженные разными синтаксическими категориями, но имеющие общий корень и связанные семантически.

Domain (область, сфера) – предметная классификация, с которой синсет связан указателями КАТЕГОРИЯ, ОБЛАСТЬ или УПОТРЕБЛЕНИЕ.

Familiarity (осведомлённость) – уровень, обозначающий частотность употребления слова.

Holonym, inherited/regular (голоним, производный/обычный) – название целого, чья часть называется меронимом. *Y* является *голонимом X*, если *X* является частью *Y*.

Hypernym – общий термин, используемый для обозначения полного класса определённых сущностей. *Y* является *гиперонимом X*, если *X* является частью *Y*.

Hyponym, brief /full (гипоним, краткий/полный) – специальный термин, используемый для обозначения члена какого-либо класса. *X* является *гипонимом Y*, если *X* является частью *Y*.

Meronym, inherited/regular (мероним, производный/обычный) – название составной части, содержания или члена чего-либо. *X* является *меронимом Y*, если *X* является частью *Y*.

Polysemy count (индекс полисемии) – количество значений слова в онтологии WordNet 2.1, принадлежащего данной синтаксической категории.

Sentence frame – структура предложения.

Synonyms, ordered by estimated frequency/grouped by similarity (синонимы, расставленные в соответствии с уровнем частотности или сгруппированные по признаку подобия) – два слова, которые могут быть взаимозаменяемыми в определённом контексте.

Synonyms/Related Nouns (для прилагательных – синонимы в виде родственных существительных) – существительные, связанные с прилагательными логической или причинной связью, или имеющие с ними общие свойства.

Synonyms/Stem Adjectives (для наречий – синонимы, прилагательные с той же основой) – прилагательные, имеющие тот же корень, что и наречие.

Synset, head/satellite (синсет, головной/периферийный): головной синсет – набор синонимов в кластере прилагательных, содержащий по крайней мере одно слово, имеющее прямой антоним; периферийный синсет – набор синонимов в кластере прилагательных, выражающих концепт, близкий по значению концепту, представленному его головным синсетом.

Syntactic category (синтаксическая категория) – часть речи, например, прилагательное, существительное, глагол и т.д.

Troponym, brief/full (тропоним, краткий/полный) – глагол, описывающий особый способ совершения действия, выраженного другим глаголом. *X* является *тропонимом Y* если глагол *X* описывает способ действия, который в общем виде выражается глаголом *Y* (например, **to march** – это *тропоним to walk*).

Библиографический список

[Гаврилова и др., 2008] Гаврилова Т.А., Муромцев Д.И. Интеллектуальные технологии в менеджменте: инструменты и системы. Учеб. пособие. 2-е изд. / Т.А. Гаврилова, Д.И.Муромцев; Высшая школа менеджмента СПбГУ. – СПб.: Изд-во «Высшая школа менеджмента; Издат. дом С.-Петерб. гос. ун-та, 2008.

[Добров и др., 2009] Добров Б.В. Онтологии и тезаурусы: модели, инструменты, приложения: учебное пособие / Б.В. Добров, В.В. Иванов, Н.В. Лукашевич, В. Д. Соловьёв – М.: Интернет-Университет информационных Технологий; БИНОМ. Лаборатория знаний, 2009. – (серия «Основы информационных технологий»).

[Елисеева, 2009] Елисеева В.В. Лексикология английского языка [Электронный ресурс], 2009. – Режим доступа: http://www.classes.ru/grammar/24.leksikologiya_angliyskogo_yazyka/3.1.htm.

[Иванников и др., 2008] Иванников А.Д. и др. Получение знаний для формирования информационных образовательных ресурсов / А.Д.Иванников, В.П.Кулагин, В.А.Мордвинов, Л.В.Найханова, Б.Б.Овезов, А.Н.Тихонов, В.Я.Цветков – М.: ФГУ ГНИИ ИТТ «Информика», 2008.

[Национальный корпус русского языка] Национальный корпус русского языка [Электронный ресурс]. – Режим доступа: http://www.ruscorpora.ru/corpora_intro.html.

[British National Corpus] British National Corpus, The [Электронный ресурс]. – Режим доступа: <http://www.natcorp.ox.ac.uk>.