

Корпусное исследование лексико-семантических отношений между 6 русскими словами, обозначающими капитальные объекты (контексты с однородностью)

Corpus-based evidence for lexico-semantic relations for 6 russian words designing permanent structures

Соколова Е. Г. (minegot@rambler.ru)

РГГУ, Москва

В статье описываются результаты корпусного исследования лексико-семантических отношений между шестью словами русского языка: дом, здание, строение, постройка, корпус и сооружение. Предлагается методика исследования, основанная на классификации контекстов, в которых эти слова совместно встречаются в корпусе текстов. Перечисляются синтаксические конструкции, представляющие контексты с однородными объектами, и ассоциируемые с ними лексико-семантические отношения.

1. Введение

Задача, решаемая в данной статье связана с исследованиями по автоматической генерации описаний изображений на ЕЯ (1, 2, 3, 4), в которых авторы системы столкнулись с проблемой выбора номинаций для упоминаемых в тексте домов и других капитальных объектов и их описаний в онтологии, используемой генератором при порождении текстов.

Систематизация понятий капитальных объектов для генерации описаний изображений обслуживает несколько процессов. Назовем три: 1) первичная номинация объектов, основанная на зрительно воспринимаемых свойствах объекта или зрительном образе, 2) упоминание в тексте понятия, отсылающего к уже введенному в дискурс понятию (вторичная номинация) и 3) наследование свойств от более общих к более конкретным понятиям в онтологии. Одно и то же слово может использоваться как первичная и как вторичная номинация. Например, [1]¹

¹ В статье сквозная нумерация примеров. Нумеруются только примеры, которые далее рассматриваются в р.б. Примеры, иллюстрирующие не рассматриваемые и терминологические контексты, не нумеруются. Номер ставится перед примером в квадратных скобках. Длинные примеры сокращены без ущерба для рассматриваемого явления, удаленные куски заменены многоточием.

«Повернув за угол этого здания, любопытствующий посетитель или вновь прибывший арестант видит в дальних углах двора два одноэтажных строения, гораздо меньших, чем главный корпус.» — первичная номинация объектов словами *строение* и *корпус*. В следующем контексте *дом* — первичная номинация, *строение* — вторичная: *«Ты назначишь ему свидание в квартире, которая расположена в этом доме. — Гречишников указал на сталинское строение, увенчанное стеклянным газовым ангелом, где между словами *дом* и *строение* существует родовидовое отношение или отношение синонимии.*

Исследуются связи пяти слов, описываемых в словарной статье ДОМ 1 Нового Объяснительного Словаря Русского Языка (НОССРЯ) (5, с. 82–84) как синонимы, каковыми они могут являться в определенных своих значениях, и которые удовлетворяют следующему смыслу²: «наземное сооружение, имеющее внутри помещения, которые занимают большую часть его объема», сформулированному для них в этой словарной статье. Подход НОССРЯ близок подходу WordNet (6), в котором похожие слова объ-

² В статье используется термин «смысл» в отличие от термина «толкование», используемого в словарях типа НОСС (см. ниже ссылку 4), поскольку обсуждается именно употребление слов в текстах.

единяются в «синсеты» по общему значению/смыслу. В WordNet найдено удачное решение — созданы синсеты и формализованы лексико-семантические отношения (ЛСО) между ними, что превращает традиционный словарь в электронный ресурс. Прочая информация, в частности, толкования, остаются традиционными. В отличие от WordNet, представляющего собой интернет-ресурс и полную лексико-семантическую базу английского языка, которая свободно доступна в Интернете для использования людьми и автоматизированными системами, НОССРЯ имеет вид традиционного филологического исследования, адресованного специалистам, и представляет собой фактически сборник исследований, более или менее субъективных в зависимости от слова (служебное / полнозначное) и автора статьи. Наша попытка опереться на НОССРЯ для включения в онтологию понятий капитальных объектов показывает, что статья ДОМ 1 не является достаточной базой для моделирования употребления слов в тексте. Традиционный словарь дает мало информации, выделяя только точные синонимы. Например, в (7) две пары из десяти комбинаторно возможных для вышеперечисленных пяти слов, считаются синонимами — *строение* и *постройка*, *здание* и *дом*.

Цель исследования состоит в том, чтобы найти объективные основания для различения ЛСО между двумя словами с позиции их употребления в тексте. Материалом исследования являются контексты, найденные в Национальном Корпусе Русского Языка (НКРЯ) (8) для пар слов из рассматриваемой группы, состоящей из пяти вышеперечисленных слов и слова *сооружение*, которое по его месту в определении смысла синсета, данного в НОССРЯ, является гиперонимом этих слов, следовательно, его смысл пересекается со смыслом данного синсета. По нескольким поисковым запросам из НКРЯ извлекаются и исследуются контексты совместного употребления пар слов. Возможны два типа контекстов: а) двумя словами обозначаются разные, но однородные (в силу принадлежности к одной семантической группе) объекты — контекст однородности; б) двумя словами обозначается один и тот же объект — контекст тождества. Размер статьи не позволяет рассмотреть оба типа контекстов, поэтому в данной статье мы ограничиваемся контекстом однородности. Кроме того, не рассматриваются контексты, близкие к тождественным, в которых понятия, обозначенные парой слов, находятся в различных ракурсах (см. р. 3).

2. Метод (часть 1)

Рассматриваются все возможные кортежи пар слов из заданного набора, всего 30 кортежей. Для определения наличия между двумя словами родови-

догового отношения и синонимии обычно используются психолингвистические тесты: а) тест на заменимость слова в тексте его синонимом — для синонимии, б) диагностические контексты: «*Это X, следовательно, это Y*» и «*Это Y, следовательно это X*» для обоих отношений. Если верны оба утверждения, то диагностируется отношение синонимии, если одно верно, а другое нет — родовидовое отношение, например, слово *птица* — гипероним слова *соловей*, потому что (верное утверждение отмечается «t», неверное — «f»): t(*Это соловей, следовательно, это птица*) и f(*Это птица, следовательно, это соловей*). *Судно* — синоним *корабль*, потому что: t(*Это корабль, следовательно, это судно*) и t(*Это судно, следовательно, это корабль*). Однако опереться на диагностический контекст в нашем случае не удастся по двум причинам: а) значения трех из шести слов ассоциируются с объектом (и с некоторым зрительным образом³): *дом*, *здание*, *корпус*, а другие воспринимаются как результат созидательного процесса, т. е. ассоциируются с процессом, например, «сооружение — это «то, что сооружено» (9). Для слов с объектной ассоциацией слова с процессной ассоциацией окажутся гиперонимами: t(*Это дом / здание / корпус, следовательно, это строение / постройка / сооружение*) по известному нам способу возникновения объектов — они *построены*, *сооружены*. Но знание не тождественно ЛСО языка. Гипотеза, проверяемая в данной статье, состоит в том, что ЛСО должны проявляться в текстах. Диагностические контексты в реальных текстах не встречаются, возможности проверить взаимозаменяемость слов нет. В статье рассматриваются контексты, в которых исследуемые слова встречаются совместно, найденные в НКРЯ по трем следующим шаблонам:

[X (1 1) Y] — выделяет контексты, в которых слова рассматриваемой пары следуют непосредственно друг за другом (знаки препинания в НКРЯ не считаются словом), например, [дом (1 1) здание], по которому найдется, в частности, контекст: [2]Работа архитекторов видна всем: вот они, их дома, здания. [Михаил Песин. Картинки с выставки (2002) // «Биржа плюс свой дом» (Н. Новгород), 2002.05.20].

[X (1 1) это (1–6) Y] и [X (1 1) быть (1–6) Y] — шаблоны, предназначенные для выделения конструкции типа «определение», например, «сосна — это дерево». В связи с тем, что между словом «это/быть» и вторым словом пары допускается вставка до 6 слов, эти шаблоны выделяют дополнительные контексты — с указательным местоимением *это*, с глаголом *быть* в прошедшем времени, и др.

В табл.1 приведены данные о количестве найденных в НКРЯ контекстов для всех пар слов. Ко-

³ Можно сказать: «Нарисуй дом, здание», плохо «нарисуй корпус», и совсем невозможно «нарисуй сооружение, строение, постройку».

лонка «?» содержит сведения о количестве не интерпретируемых однозначно примеров. Колонка «Не рассм.» содержит контексты, в которых слова имеют значения, отличные от смысла синсета. Это примеры со словами *строение*, *дом* и *корпус* как элементами адреса; слова *строение*, *постройка*

и *сооружение* как процесс, например, *сооружение домов*, *корпус постройки 19 в.*; *корпус* как оболочка, например, *И тут Христо заметил в темноте — на минуту совсем ясно — огромные, как облака, паруса и высокий, как дом, корпус*. [Б. С. Житков. Элчан-Кайя (1926)], *сооружение* как устройство.

Табл. 1

		Всего	?	Не рассм.
1	дом здание	14	1	
2	дом строение	4	1	
3	дом постройка	11		10 дома постройки 70ых годов
4	дом корпус	11		9 — адрес (см.)
5	дом сооружение	1		
6	здание дом	39	2	30 здание дома престарелых
7	здание строение	84		
8	здание постройка	15	1	13 здание постройки прошлого века
9	здание корпус	9		
10	здание сооружение	107		
11	строение дом	22		12 (устар.) строение домов
12	строение здание	6		
13	строение постройка	-		
14	строение корпус	1		
15	строение сооружение	65		
16	постройка дом	101		94 постройка дома, старой постройки дом,...
17	постройка здание	19		16 постройка зданий, старой постройки здание
18	постройка строение	-		
19	постройка корпус	-		
20	постройка сооружение	2		
21	корпус дом	19		16 корпус дома
22	корпус здание	7		1 корпус здания
23	корпус строение	2		— (устар.)
24	корпус постройка	1		1 корпус постройки 18 века
25	корпус сооружение	-		
26	сооружение дом	11		8 сооружение дома
27	сооружение здание	3	2	
28	сооружение строение	3	1	
29	сооружение постройка	3		
30	сооружение корпус	-		

Не рассматриваются контексты, в которых слово сохраняет ассоциацию с процессом, например, а) *В то же время Петербург стал украшаться и новыми, наконец достойными и гармонирующими с характером Петербурга зданиями — постройками молодых архитекторов Фомина, Шуко, Таманова и др.* [М. В. Добужинский. Облик Петербурга (1938) // «Звезда», 2003]. б) *К городу примыкали громадные постройки домов и судостроительных верфей.* [А. С. Новиков-Прибой. Цусима (1932–1935)]. Пример а) можно перефразировать: «зданиями постройки молодых архитекторов», где слово *постройка* будет иметь только процессное значение, в б) реализуется устаревшее значение слова *постройка* — «место, где строят» (9).

Пустые строки табл. 1 интерпретируются в р. 4. В строках 7, 10 и 15 особенно много примеров. Это административные штампы. Они рассматриваются в р. 5. Оставшиеся контексты интерпретируются по типам и ЛСО для каждой пары слов в р. 6.

3. Метод (часть 2)

Рассматриваются попарно вышперечисленные слова, и проверяется наличие родовидового отношения и синонимии между членами этих пар. Обозначим отношения стрелками:

$X \rightarrow Y$ - родовидовое отношение: Y — гипоним,
 X — гипероним;
 $X \leftrightarrow Y$ - синонимия: Y и X — синонимы.

Пробуем найти манифестацию этих отношений в контекстах из НКРЯ. Заметим, что слова выделены на основании смысла синсета и точными синонимами не являются. Наоборот, в НОССРЯ автор словарной статьи ДОМ 1 пытается «развести» их значения, указывая различительные признаки, которыми обычно обладают обозначаемые ими объекты: размер, материал и др.⁴. Различительные признаки рассматриваются как часть лексического значения слова. Они сообщают гипониму конкретность по отношению к гиперониму в случае, если между смыслами слов обнаруживается родовидовое отношение, например, *дом* имеет некоторые конкретные свойства, которые выделяют его как вид наземных сооружений и отличают от других видов наземных сооружений (*здание, строение* и др.), вступающих с ним в отношение квазисинонии. Получается, что оба отношения, родовидовое и квазисинонии, базируются на понятии различительного признака. Отношение квазисинонии обозначим разорванной стрелкой: $X \leftarrow \{ _ | _ \} \rightarrow Y$, где в фигурных скобках помещаются различительные признаки слов X и Y , соответственно. Если одновременно есть свидетельство о квазисинонии и наличии родовидового отношения, будем обозначать эту ситуацию такой же формулой, но с удвоенной стрелкой, указывающей на гипоним: $X \leftarrow \{ _ | _ \} = > Y$. Родовидовое отношение тоже может указывать различительный признак гипонима, например, «←+жилой|», или гиперонима, например, «←|+рез.процесса».

В данном исследовании оказалось полезным использовать комплексный признак, привязанный как к коммуникативной ситуации, так и к обозначаемой реальности, назовем его ракурсом. Приблизительно, ракурсы составляют две группы: застройка — когда речь идет об объектах вообще, объектах в регулирующих документах, объектах населенных

пунктов определенного типа, и визуальный ракурс — когда объект представляется (как бы) в зрительном восприятии говорящего:

- 1 «застройка» — совокупность капитальных объектов, составляющих населенный пункт, например, *московские здания*, включая объект как элемента застройки, например, *здание, строение* и *сооружение* в документах, в частности, этот ракурс свойственен официально-деловой сфере функционирования текстов.
- 2 «мини-жил» — совокупность капитальных объектов в «распределенном капитальном объекте» (РКО), предназначенном для постоянного (в том числе сезонного) проживания — *усадеб, дворе, даче*. Под РКО понимается объект, занимающий относительно небольшой и часто огороженный участок земли, объединенный общим собственником или пользователем, состоящий обычно из нескольких капитальных объектов;
- 3 «мини-спец» — совокупность капитальных объектов в РКО специального назначения — *монастыре, больнице, заводе* и др. Ракурсы 2 и 3 называем мини-ракурсами.
- 4 ракурс зрительного восприятия («ЗВ»), ракурс наблюдателя — застройка, не огороженная, например, *застроенная поляна, берег, совхозные угодья, дачный поселок*, зрительно воспринимаемые рассказчиком;

В данной статье рассматриваются контексты, в которых оба слова находятся в одном и том же ракурсе. Кроме того, что каждое слово пары обладает определенным ракурсом, оба слова погружены в определенный синтаксический контекст.

В данной статье мы рассматриваем контексты однородности, в которых соучаствуют слова одного синсета. Он создается при перечислении объектов в составе сочиненной цепочки или при наличии в предложении указателей, создающих единый ситуационный контекст для обоих слов, например, общее обстоятельство для двух сочиненных предложений, соединение сравнительным предикатом и др. Выделяется три типа контекстов однородности: $O_{\text{сочинение}}$, $O_{\text{сочинение}_c}$ обобщением и $O_{\text{—}}$ — остальные контексты однородности.

$O_{\text{сочинение}}$ — цепочка однородных членов, в которой однородность связана либо с принадлежностью членов одному синсету, либо наличием у них общего гиперонима. В первом случае однородные члены являются квазисинонимами с различительными признаками, например, [дома, здания] → дом←{+жилой|-жилой}→здание; во втором имеют общего гиперонима, например, в терминосистеме «недвижимость» (см. ниже). Результат можно записать:

[X , Y] и (X и Y принадлежат одному синсету)
 → $X \leftarrow \{ a | b \} \rightarrow Y$, где a и b — различительные признаки.

⁴ Можно понять, что толкование и различительные признаки — это элементы метаязыка описания слова в словаре типа НОСС (Июмдин Б. Л. Как составляются словари. Конспект семинаров // Лингвистика для всех. Летние лингвистические школы 2005 и 2006. М: «МЦНМО», 2008. С. 85). Встречается два типа различительных признаков: маркированность, например, «+ — наличие архитектурного замысла» в статье ДОМ 1 НОССРЯ, и положение на шкале, по которой противопоставляются слова — члены синонимического ряда, например, для слов *тоска, уныние* и *грусть* по шкале «интенсивность, глубина и длительность чувства в вышеуказанном источнике с. 85. Мы тоже используем термин «различительный признак», причем в нашем случае это — маркированность. В статье «Синонимия» энциклопедии Кругосвет (без подписи) подобный термин не используется, просто речь идет о смысловых, стилистических и др. различиях http://www.krugosvet.ru/enc/gumanitarnye_nauki/lingvistika/SINONIMIYA.html.

$[X, Y]$ и $\neg(X$ и Y принадлежат одному синсету)
 \rightarrow найдется Z ($Z \rightarrow X$ и $Z \rightarrow Y$).

О_сочинение_с_обобщением — сочиненная цепочка, вводимая обобщающим словом, причем обобщающее слово и некоторые члены цепочки принадлежат одному синсету, например, «*возведение каких-либо строений (домов, поселков, предприятий и т. д.)*». Члены цепочки отделяются от обобщающего слова двоеточием, заключаются в скобки, иногда с уточнением, например, *в том числе*:

$[Z(X, Y)]$ и $(X$ и Z принадлежат синсету) $\rightarrow Z \rightarrow X$.

Разновидностью этого типа являются сложно-сочиненные предложения с квантором всеобщности в первом предложении

$[S1(\text{все } Z), S2(X)]$ и $(X$ и Z принадлежат одному синсету) $\rightarrow Z \rightarrow X$, например, [32].

О_ — слова одного синсета обозначают однородных участников одной ситуации. Тип отношения между ними зависит от контекста. Это может быть синонимия: употребление для номинации слов-синонимов имеющее стилистическую мотивацию — исключение повторения слова:

[3] *Один из них, 50-летний безработный, позвонил в милицию из квартиры на Сумской улице и сообщил, что соседнее с его домом здание ОВД «Чертаново-Северное» заминировано.* [Дмитрий Павлов. Ведомости (1996) // «Коммерсантъ-Daily», 1996.01.20].

Оба объекта — дом безработного и здание ОВД, могут быть названы и одним, и другим словом, например, «*один из домов (одно из зданий) на Сумской улице заминирован(/о)*»⁵. Это может быть отношение часть, см. р. 6.4 *дом* — *корпус*, где содержание отношения создается определениями слова *корпус*. Можно указать на следующие синтаксические типы контекстов:

$[Z \text{ Pr} \dots X Y]$, где Z — согласованное определение к Y , (Pr зависит от Z),

$(X$ и Y принадлежат синсету) $\rightarrow X \leftrightarrow Y$, например, [3].

$[Z \text{ как} \dots X Y]$, где Z — согласованное определение к Y ,

$(X$ и Y принадлежат синсету) $\rightarrow X \leftrightarrow Y$:

⁵ Замена слов в данном примере невозможна, потому что *здание* не сочетается с притяжательным местоимением, обозначающим собственника.

[4] *Давно не крашенный, серый, как большинство старых деревянных строений, дом тем не менее носил следы хорошей хозяйской руки.* [И. А. Ефремов. Лезвие бритвы (1959–1963)]

4. Пустые строки табл. 1 (13 и 18, 19 и 24, 25 и 30)

Из табл. 1 видно, что пары слов «*сооружение* и *корпус*», «*постройка* и *корпус*», а также «*постройка* и *строение*» не встречаются в одном контексте.

Слова *строение* и *постройка* по тесту с диагностическим контекстом являются точными синонимами и включены в (7) и при этом не встречаются в одном контексте, который мог бы диагностировать родовидовое отношение или квазисинонимиию. Следовательно, можно рассматривать отсутствие таких контекстов как подтверждение синонимии этих слов. Синонимии не противоречит факт, что в современном языке сферы употребления этих слов различны: *постройка* в отличие от *строение* крайне редко употребляется в текстах официально-деловой сферы.

Для пар «*корпус-постройка*», «*корпус-сооружение*» с разной природой значения диагностический контекст не показателен. В контекстах из НКРЯ для этих пар слов ЛСО не диагностируются, следовательно, их нет. Это утверждение не противоречит знанию того, что *корпус*, по сути, является *зданием* или *домом* (только называется *корпусом*) и как таковой ассоциативно может быть связан с понятиями, находящимися с ними в лексико-семантических отношениях. Учитывая ограниченность эксперимента также нельзя утверждать, что такие контексты в принципе невозможны, но в контекстах однородности эксперимента эта возможность не проявляется, и эти ассоциации не работают.

5. Здание, строение, сооружение (строки 7, 10, 15, 28) как термины

В строках 7, 10, и 15 табл. 1 оказалось особенно много примеров, причем подавляющее большинство встречаются в одном сочетании: «*здание, строение, / и /или сооружение*». Это сочетание является штампом официально-деловой сферы — документации по недвижимости. Строка 7 выделяет начальную пару слов из этой тройки (*здание строение*) — 84 контекста, строка 15 — конечную пару (*строение сооружение*) — 65 контекстов. *Дом* встречается в тройке в найденных контекстах вместо слова *здание* всего два раза, причем в «маргинальных» ситуациях — *заложены* или *самовольной постройки*. Кроме того, в этих двух примерах слово *строение* предваряется определением «*другой, иной*»: а) *Самовольной по-*

стройкой является жилой дом, другое строение, сооружение или иное недвижимое имущество, созданное на земельном участке, не отведенном для этих целей [Гражданский кодекс Российской Федерации. Часть первая (1994)]; б) ... решением об изъятии земельного участка, на котором находятся заложенные дом, иные строения, сооружения или насаждения, убытки, ..., возмещаются залогодержателю в полном [Закон Российской Федерации «О залоге» (1992)]. Это свидетельствует о том, что в этих примерах реализуется отношение синонимии в паре «дом–строение» как терминов или «общеязыковой» синонимии слов дом и здание в несвойственной для дома сфере употребления в маргинальных для здания контекстах. Второе предпочтительнее.

Строка 10 содержит административный штамп «здания, сооружения», который используется часто в сочетании с «земельным участком» и представляет сжатый вариант трехчленного штампа — 107 контекстов; вариант двучлена со словом строение встречается только два раза. Термины встречаются только в одном типе контекстов — **О_сочинение**.

В табл. 2 представлены ЛСО официально-деловой сферы «недвижимость» для нашей группы слов.

Табл. 2

		застройка	
1	объект_недвижимостиТ	→	зданиеТ, строениеТ, сооружениеТ
2	зданиеТ, строениеТ	←{ _ -синсет*, -процесс }→	сооружениеТ
3	зданиеТ	←{ ? ? } = >>	строениеТ

* СооружениеТ не принадлежит синсету и не имеет процессного значения, определенного для него в (9). Главная составляющая значения термина в том, что он представляет «объект недвижимости», т. е. является его гипонимом. Дом как объект недвижимости встречается крайне редко в основном для обозначения «неполноценного» объекта — самовольно построенного или заложенного.

Комментарии. В табл. 2 слова-термины помечены большой буквой «Т» в конце слова. Значения терминов сужены по сравнению с теми же словами — не терминами. Сооружение не имеет процессного значения, определенного для него в (9) — (-процесс). Главная составляющая его значения в том, что оно обозначает «объект недвижимости», т. е. является гипонимом этого понятия, и обозначает объекты инфраструктуры населенных пунктов, например, очистные, спортивные сооружения, дороги и т. п., отличные от зданий и строений. Строение является квазисинонимом термина здание с невыявленным различительным признаком и его же гипонимом.

6. Контексты с нетерминологическими значениями

6.1. Пара слов «дом и здание» (строки 1 и 6)

Обе части диагностического контекста для данной пары слов в целом истинны: (т)Это дом, следовательно, это здание. и (т)Это здание, следовательно, это дом. Т.е. слова могут считаться синонимами, что подтверждается и их наличием в (7).

О_ (6 контекстов) также свидетельствуют в пользу синонимии (в ракурсе dev, dev-el):

- [5] В городе много домов было разрушено, чувствовалась острая нехватка зданий. [Любовь Кузнецова. «...Собираю разрозненные брёвнышки народа своего...» (2003) // «Вестник США», 2003.09.03]
- [6] Цены напрямую зависят от технического состояния объектов, и самые лучшие в этом смысле здания будут стоить не ниже сталинского дома. [Квартиры б/у: не хуже новостроек (2003) // «Мир & Дом. City», 2003.09.15]
- [7] И только потому, что школьные дома были поместительнее других зданий, принадлежавших сельским коммунаам, школы естественно служили местом собраний [Ф. К. Сологуб. Королева Ортруда (1909)].

Пример в ракурсе «ЗВ»: [4], а также:

- [8] Они миновали уж здание; вон за зданием — дом; и вон — окна... [Андрей Белый. Петербург (1913–1914)].

О_сочинение — наиболее частый (11 контекстов) для этой пары: см. [3], а также:

- [9] Все его сооружения — жилые дома, здания городских учреждений, заводские цеха, проезжие и пешеходные дороги — покоились на сваях. [Евгений Рубин. Пан или пропал. Жизнеописание (1999–2000)]
- [10] Я с ними рассматриваю рисунки древних зданий, домов, утвари, — сам черчу, объясняю, как, бывало, тебе: что сам знаю, всем делюсь. [И. А. Гончаров. Обрыв (1869)]
- [11] Но декорация совершенно переменялась, когда мы выехали на берег Сены; тут представились нам красивые здания, дома в шесть этажей, богатые лавки. [Н. М. Карамзин. Письма русского путешественника (1793)]

В рассматриваемых сочиненных цепочках наблюдается следующие закономерности:

- а) в контекстах до середины 20 в. встречается только кортеж «здание дом» [10, 11], в современном языке наоборот — «дом здание» [9].
- б) в современном языке⁶ ясен различительный признак, поскольку он проявляется в половине контекстов, например, в [9]: дом — жилой, здание — учреждение. До середины 20 в. этой дифференциации, по-видимому, не было, в частности, функция «+учреждение» проявляется только в одном примере и ассоциирована со словом дом [7].

О_сочинение_с_обобщением. Контексты находятся для обоих кортежей, но в данной статье рассматривается только <здание — дом>. В следующих примерах ракурс — элемент застройки, его оснащение, реконструкция и т. п. Здание является гиперонимом дома:

[12] Договор строительного подряда заключается на строительство или реконструкцию предприятия, здания (в том числе жилого дома), сооружения или иного объекта, ... [Гражданский кодекс Российской Федерации. Часть вторая (1995)]

[13] Трудно себе представить современное здание (будь то жилой дом, торговый комплекс или бизнесцентр), которое обходилось бы без лифтов. [Вверх по рельсам, ведущим вниз (2003) // «Строительство», 2003.05.26]

6.2. Пара слов «дом и строение» (строки 2 и 11)

О_сочинение — 3 примера в ракурсе «ЗВ», где дом имеет различительный признак +жилой, который проявляется в виде указания на хозяина дома или проживающих [15, 16]:

[14] Спасать дома́, строения́ бесполезно — огонь распространяется с колоссальной скоростью. [Д. А. Гранин. Месяц вверх ногами (1966)]

[15] Место сборищ народных было обыкновенно при кирках, возле которых находится всегда несколько строений, дома пастора и кистера, шинок и несколько домов, в которых помещались ремесленники. [Ф. В. Булгарин. Воспоминания (1846–1849)]

[16] В следующую ночь дом Белосельских был тоже окружен мушкетерами и пожарными, и в надворных строениях была задержана разбойничья шайка, переселившаяся из дома Гурьева. [В. А. Гиляровский. Москва и москвичи (1934)]

О_сочинение_с_обобщением. В ракурсе застройки строение является гиперонимом дома.

[17] ...жилищный фонд образуют находящиеся на территории Российской Федерации жилые дома́, а также жилые помещения в других строениях (т. е. дома́, строения, строительство которых закончено с осуществлением регистрации прав на построенный объект недвижимости). [Недостроенное жилье и земельный налог (2004) // «Бухгалтерский учёт», 2004.12.06]

[18] ... возведение каких-либо строений (домов, поселков, предприятий и т. д. [В. Ларичев и др. Характеристика коррупционных проявлений в сфере природопользования (2 часть) (2004) // «Адвокат», 2004.12.01]

6.3. Пара слов «дом и постройка» (строки 3 и 16)

О_сочинение (8 контекстов). Эта пара встречается исключительно в ракурсах «мини-жил» и «ЗВ», причем постройки — во множественном числе. Оба слова часто маркируются по функции, например, надворные, хозяйственные, совхозные, главный, запасной... При этом выявляется «прикладной к дому» характер построек [22, 23], а в [24] отсутствие построек воспринимается как аномалия. Квазисинонимия:

[19] Не умеют быстро приспособливать к обороне местные предметы (дома, постройки и т. д. [Сборник боевых документов Великой Отечественной войны. Выпуск 11 (1941–1945)]

[20] Все эти дни Андрей Сургеев прожил как бы человеком-невидимкою, ..., анадсовхознымипостройками, дома́ми и клубом, над машинным двором, над совхозной землёй, ..., над потерявшими листву деревьями — не солнце и луна, не облака, набухшие влагой, а чавканье и чмоканье сапожищ Шишлина. [Анатолий Азольский. Лопушок // «Новый Мир», № 8, 1998]

[21] Во-вторых, нежелательно, чтобы пруд оказался в тени какой-либо постройки, дома или забора — его обитателям необходим солнечный свет в течение 4–5 часов в день. [Водный сад (2003) // «Сад своими руками», 2003.03.15]

⁶ Под «современным языком» иногда понимается русский язык со времени А. С. Пушкина. Для наших слов это — русский язык начиная со второй половины 20 в., когда изменился характер застройки, а с ним и средства языка, ее обслуживающие.

[22] *Когда-то дом был главой целого семейства постройке.* [Василий Белов. Плотничьи рассказы (1968)]

[23] *Вместе с другими хозяйственными постройками дом образует маленький готический поселок.* [Николай Гаврюшин. Суханово: соприкосновение двух миров (2002) // «Ландшафтный дизайн», 2002.03.15]

[24] *Седьмой двор выделялся издали какой-то запущенной бедностью — не только потому, что дом был низок и мал и надворные постройки почти отсутствовали,* [Сергей Бабаян. Ротмистр Неженцев (1995–1996)]

О_сочинение_с_обобщением. В ракурсах «мини-жил» и «ЗВ» застройка обозначается словом «постройки» и ассоциируется со стихийностью (*возникают, разбросанные*) и своеобразием (*стиль*). Дом — гипоним постройки::

[25] *Одна за другой в «Муромцево» возникают выполненные в едином стиле постройки: дом главный и дом запасной, дом управляющего, летний театр, купальня, охотничий домик.* [Мария Ожерельева. Современники называли ее «царской» (2002) // «Ландшафтный дизайн», 2002.01.15]

[26] *Все постройки были солидные, кирпичные, дома крыты черепицей, архитектура частью русская.* [В. А. Обручев. От Кяхты до Кульджи. Путешествие в Центральную Азию и Китай (1940)]

6.4. Пара слов «дом и корпус» (строки 4 и 21)

О_. Пара встречается в ракурсах «мини-жил» и «ЗВ». Слова обозначают объекты, которые представлены в дискурсе как разные, причем корпус обозначает *часть дома*, что определяется выделенными жирным шрифтом определениями к слову корпус:

[27] *Все перечисленные системы, агрегаты и механизмы удобно разместить в пристройке или **отдельном от** дома корпусе вместе с гаражами, мастерской и другими необходимыми в усадьбе подсобными помещениями.* [Михаил Харит. Автономный дом (2002) // «Домовой», 2002.05.04]

[28] *Возле дома... горел костер Дом был большой, **в несколько корпусов.*** [В. П. Катаев. Отче наш (1946)]

[29] *Основная часть дома была на фундаменте, **исключение составлял боковой корпус, где поме-***

щались кухня, туалет, склад дров и запасная комната, где жил слуга-сарт. [А. А. Татищев. Земли и люди: В гуще переселенческого движения (1906–1921) (1928)]

6.5. Пара слов «дом и сооружение» (строки 5 и 26)

О_сочинение_с_обобщением. Есть только один пример, диагностирующий родовидовое отношение, которое создается контекстом и «процессной» ассоциацией значения слова *сооружение* (см. р. 2), в следующих примерах слова, поддерживающие эту ассоциацию, выделены жирным шрифтом:

[30] *Малой динамической сложностью (неопределенностью) обладают такие **системы, как строительные конструкции, сооружения:** дома, мосты, мебель, автомобили, самолеты, ...* [И. М. Гуревич. Законы информатики — основа исследований и проектирования сложных систем (2003) // «Информационные технологии», 2003.11.24]

[31] *В ролевой игре по курсу «Основы производственной психологии и педагогики» студенты **возводят сооружение** (дом, школу, гостиницу).* [Вуз будущего // «Студенческий меридиан», 1985]

[32] *Все его сооружения — жилые дома, здания городских учреждений, заводские цеха, проезжие и пешеходные дороги — **покоились на сваях.*** [Евгений Рубин. Пан или пропал. Жизнеописание (1999–2000)]

6.6. Пара слов «здание и строение» (строки 7 и 12)

О_сочинение. Всего 2 контекста, не считая штампа «здания, строения, сооружения». Следующий пример из текста закона отнесем к терминам (см. р. 5): «...**сооружения связи — объекты инженерной инфраструктуры, в том числе здания, строения, созданные или приспособленные для размещения средств связи и кабелей электросвязи;** [Федеральный закон «О связи» (2003)]», второй пример не принадлежит к современному языку, как и единственный пример, свидетельствующий о синонимии слов здание — строение:

О_ : *До тех пор великие князья московские жили не иначе, как в деревянных домах, да и вообще на всем русском севере каменными зданиями были только церкви: жилые строения были исключительно деревянные.* [Н. И. Костомаров. Русская история в жизнеописаниях ее главнейших деятелей. Выпуск второй: XV–XVI столетия (1862–1875)]. Следовательно

но, в современном языке эти слова вне официально-деловой сферы не ассоциированы друг с другом в контекстах однородности.

6.7. Пара слов «здание и постройка» (строки 8 и 17)

Встречаются в трех видах однородных контекстов, но в иных, чем *дом-постройки* ракурсах. Различительные признаки *построек* — «несамостоятельность», «древность» — поддерживающие контексты выделены жирным шрифтом.

О_сочинение:

[33] Для разрушения особо прочных зданий, *построек и инженерных сооружений* создаются корпусные или дивизионные группы разрушения из артиллерии особой мощности и гвардейских миномётных частей (М-31). [Сборник боевых документов Великой Отечественной войны. Выпуск 1 (1941–1945)]

О_сочинение_с_обобщением:

[34] Внутри шестиугольной крепости обозначены три *постройки*: здание с двумя башнями, «кордегардия» («*Corp de Guard*») и «дом коменданта» («*Comend Niuis*»). [Т. А. Базарова. План петровского Петербурга. Источниковедческое исследование (2003)].

О_ : — квазисинонимия

[35] При занятии противником *прилегающих к занятому зданию построек* организует их очистку в том же порядке, как указано выше. [Сборник боевых документов Великой Отечественной войны. Выпуск 1 (1941–1945)]

[36] Судя по тому, что здания этого времени сохранялись хуже, чем *постройки домонгольской поры*, многое в умениях было утрачено — как правило, уже через 100–150 лет каменные здания приходили в негодность. [С. А. Еремеева. Лекции по русскому искусству (2000)]

6.8. Пара слов «здание и корпус» (строки 9 и 22)

Встречаются с РКО специального назначения: кадетским корпусом, духовной академией, больницей, комплексом администрации и др. *Корпусом* называется *главный* объект и все другие объекты, связанные со специализированной функцией всего заведения. Здания — тоже функционально отмече-

ны, но не специальными функциями, а обычными (культурными) функциями. Различительный признак «+главн.» реализуется разными прилагательными: *главный корпус*, но *основное здание*.

О_сочинение:

[37] «Храмы, башни, стены монастыря, здания, корпуса, часовни — все в полуразвалившемся состоянии» [архимандрит Макарий (Веретенников). Первый наместник возрожденной Лавры // «Альфа и Омега», 2001]

[38] Пансионат «Крымское приморье» — три *жилых корпуса*; *процедурный корпус*; здание *клуба-столовой* (с колоннами в добротном сталинском стиле), ... [Ольга Ляпунова. Отчет-лория о путешествии в Крым (2002)].

[39] Отправленный с рабочим визитом в Грозный новый вице-мэр Москвы Михаил Мень заявил, что параллельно восстановлению *основного здания* будет возводиться и *дополнительный корпус* поврежденной чеченской администрации. [Нина Зверева. Воспитательная спецоперация Владимира Путина (2003)].

6.9. Пара слов «здание и сооружение» (строки 9 и 22)

О_сочинение_с_обобщением (аналогично р. 6.5):

[40] ...*сооружения* связи — объекты инженерной инфраструктуры, в том числе здания, строения, созданные или приспособленные для размещения средств связи и кабелей электросвязи; [Федеральный закон «О связи» (2003)].

6.10. Пара слов «строение и корпус» (строки 14 и 23)

На эту пару слов встретилось всего три примера. Два — на кортеж *корпус* — *строение* относятся к 19 в. и производят впечатление устаревших: а) Он заложил еще *новый сад* и *новый корпус, строение* для дворовых. [Л. Н. Толстой. Война и мир. Том третий (1867–1869)]; б) Та половина была на дворе *длинная пристройка к главному корпусу строения*. [Ф. Ф. Вигель. Записки (1850–1860)], так как в слово *строение* без определений в современном языке (см. сноску 7) не ассоциируется с жилым объектом в а) и не может обозначать главный объект мини-застройки в б). Пример на кортеж *строение* — *корпус* [1] в р. 1 относится к типу О_ со сравнительным предикатом. Мог бы свидетельствовать о синони-

мии, но скорее это квазисинонимия с различительным признаком корпуса «+главный», который в тексте и фигурирует.

6.11. Пара слов «постройка и сооружение» (строки 20 и 29)

Только сочинительный контекст и всего два примера. Выделенные жирным шрифтом контексты указывают на маргинальность объектов и привязанность к месту.

О_сочинение:

[41] *возводить на полученных в установленном порядке земельных участках **постоянные или временные постройки, сооружения и дороги, необходимые для осуществления хозяйственной деятельности, связанной с пользованием животным миром;*** [Федеральный закон «О животном мире» (1995)]

[42] *Здания, сооружения, постройки давно **физически устарели**, большинство приемных пунктов рыбы еще колхозной постройки, капитальные и текущие ремонты их не производились десятки лет.* [Строить быстро экономично, качественно (1976) // «Маяк севера», 1976.09.04].

7. Выводы

Предложен Метод корпусного исследования ЛСО между двумя словами и изложены результаты исследования ЛСО внутри пар слов из одной семантической группы: *дом, здание, строение, постройка, корпус, сооружение* этим методом на материале НКРЯ (8).

Метод состоит в а) отборе из корпуса контекстов совместного употребления пары изучаемых слов; б) классификации контекстов на однородные (слова обозначают разные объекты) и тождественные (слова обозначают один и тот же объект); в) классификации синтаксических конструкций рассматриваемого контекста на предмет возможности определенных синтаксических конструкций диагностировать определенные ЛСО. В отличие от психолингвистических диагностических контекстов⁷ такие реальные контексты диагностируют ЛСО с учетом комплексных условий конкретного речевого акта, получивших в статье название «ракурсы».

⁷ Фактически психолингвистические контексты рассматривают ЛСО только в одном из возможных ракурсов, остается вопрос, можно ли рассматривать этот ракурс как обобщение всех других.

Результаты получены с учетом следующих ограничений: а) метод корпусного исследования ЛСО, примененный к парам слов; б) рассматривались контексты из НКРЯ не ранее второй половины 20 в., поскольку до этого времени застройка и средства языка ее отражающие, были иными; в) за недостатком места представлены результаты исследования только однородных контекстов.

Рассматривались три типа отношений: синонимия, квазисинонимия, родовидовое. Показано, что в основе двух последних лежит понятие различительного признака. Для рассматриваемой группы слов с указанными ограничениями оказалось важным выделить следующие различительные признаки: «+жилой», «+учреждение», «+специальный», «+хозяин», «+функциональн.» (например, *хозяйственный, надворный,...*), «+главн.», а также «+несамостоятельность», «+древность», не исключаются и другие. Физические и конструкционные свойства, например, наличие архитектурного замысла, материал, размер и др. не существенны при установлении ЛСО между словами этой группы.

В ходе исследования понадобились следующие четыре ракурса:

1 «застройка», **2** «мини-жил», **3** «мини-спец», **4** «ЗВ» (см. р. 3), которые оказались важными для данной группы слов. Далее приводится сводная таблица типов ЛСО, выявленных для данной группы слов, с комментариями для каждой пары слов.

Дом — здание. Эти слова являются основным средством современного русского языка для обозначения капитальных объектов в рассматриваемой группе слов. Они вступают в ЛСО друг с другом и со всеми другими словами рассматриваемой группы (исключение составляет пара *здание* и *строе*). Эти слова вступают в ЛСО в ракурсах **1** и **4**, где они являются синонимами, не вступают в мини-ракурсах. Кроме того, в ракурсе **1** *здание* является гиперонимом *дома*, который имеет различительный признак «+жилой». В ракурсе **4** *дом* и *здание* являются квазисинонимами с различительными признаками «+жилой» и «+учреждение», соответственно.

Строение — дом. В ракурсе **1** *строение* — гипероним *дома*, в ракурсе **4** они — квазисинонимы, причем *дом* имеет различительный признак «+хозяин (проживающий)».

Постройка (мн. ч-ло) — *дом, здание.* В мини-ракурсах *постройки* являются гиперонимом *дома* и *здания*, причем *дом* употребляется в **2**, а *здание* — в **3** ракурсе. *Постройки* также являются квазисинонимом *дома* в двух ракурсах, причем в ракурсе **2** маркированным членом является *дом* с различительным признаком «+главн.», а в ракурсе **3** маркированным членом являются *постройки* с различительным признаком «+функциональн.». Наконец, в ракурсе **1** *постройки* являются маркированным квазисинонимом *здание* с различительным признаком «+несамостоятельность», «+древность».

Табл. 3. ЛСО для слов *дом, здание, строение, постройка, корпус, сооружение* в контекстах однородности

		1 застройка	2 мини-жил	3 мини-спец	4 ЗВ	
1	дом	↔; ← + жилой			↔ + жилой + учрежден.	здание
2	дом	←			+ хозяин _	строение
3	дом		← + главн ?		← ? + функциональн.	постройки (мн.)
6	здание	? + несам., древность		←		постройки (мн.)
4	дом, здание	← + результат процесса				сооружение
5	здание			0 + спец		корпус
6	—часть→		дом(ед.)	здание	дом(ед.)	корпус
7	строение			0 + глав		корпус
8	постройки				? ?	сооружения

Сооружение — *дом, здание*. *Сооружение* является гиперонимом и квазисинонимом слов *дом* и *здание* и имеет различительный признак «+результат процесса», обозначая любой результат созидательной («соорудительной») деятельности в соответствии с определением в (9).

Корпус — *дом, здание*. *Здание* и *корпус* встречаются только в ракурсе 3, причем *корпус* является или **частью** здания или его квазисинонимом, маркированным признаком «+спец.». *Корпус* и *дом* встречаются в ракурсах 2 и 4, причем *корпус* находится в отношении «часть» к *дому*. *Корпус* не используется в ракурсе 1.

Строение и *постройка* являются точными синонимами и разделяются по сфере употребления: в официально-деловой сфере используется слово *строение, постройка* — крайне редко. Поскольку совместного употребления в одном примере этих слов не встретилось, синонимия этих слов не указана в сводной табл. 3.

Строение — *корпус* встречаются в одном примере [1] в ракурсе 3 как квазисинонимы, причем *корпус* маркирован признаком «+главный».

Постройка — *сооружение* — квазисинонимы в ракурсах 3–4. В имеющихся двух примерах речь идет о хозяйственных объектах.

Приведенной таблицей можно пользоваться для генерации синтаксических конструкций, соответствующих указанным отношениям и для номинации объектов в определенных ракурсах. В работах (1) и (2) мы уже использовали понятия, уточняющие особенности дискурсивного присутствия объекта, называя их «когнитивные перспективы». Это более узкое, чем ракурс понятие, позволяющее как бы «поворачивать» объект разными его ипо-

стасями, например, *озеро* как точка и как зеркало. Когнитивные перспективы влияют на выбор глагола для реализации локативных отношений на изображении. В данной статье мы исследовали понятие ракурса, связанное с номинацией объектов. Не очевидно место ракурса в модели генерации текста, оно может быть связано как с онтологией ситуаций, так и со словарем, обеспечивающим обозначение понятий словами естественного языка.

8. Направления дальнейших исследований

Проведенное исследование является только первым шагом в изучении ЛСО слов с помощью корпусов текстов и требует продолжения для проверки и углубленного исследования сформулированных выше результатов. Требуется изучения понятие «ракурс». Необходимо сравнить результаты исследования в контекстах однородности с результатами в контекстах тождества. Требуется проверка по корпусу описаний изображений (см. в (3)) полученных результатов и выработка правил употребления средств языка для описания объектов на изображении. Продолжения и более тщательной разработки требует классификация синтаксических конструкций. Следует принимать во внимание, что в процессе порождения речи участвуют многие факторы, поэтому правила генерации не могут быть простыми. Возможно, из результатов исследований, подобных данному, можно выделить однофакторные правила, но это требует серьезных исследований.

Литература

1. *Болдасов М., Соколова Е.* Онтология и планирование в генерации описаний изображений // Вопросы ИИ, № 2 (в печати)
2. *Boldasov M. V. Sokolova E. G.* Ontology for image annotating // 17th International Conference on Conceptual Structures ICCS'09. Knowledge and ontology "ELSWARE-2009" (Moscow 26–31 July, 2009). Russia, M.: University higher school of economics, 2009. P. 13–21.
3. *Соколова Е. Г.* Об использовании семантических отношений для описания изображений // Вестник РГГУ №8/07 Серия «Языкознание», М.: Издательский Центр РГГУ, 2007. с. 131–144.
4. *Соколова Е. Г., Болдасов М. В.* Формализованное описание содержания изображения как данные для генерации текста // Компьютерная лингвистика и интеллектуальные технологии. Труды Международной конференции Диалог'2007, «Бекасово», 30 мая — 3 июня 2007, М.: Издательский Центр РГГУ, 2007. с. 508–515.
5. *Новый объяснительный словарь синонимов русского языка* (под общ. ред. Ю. Д. Апресяна). Первый выпуск. М: «Школа «Языки русской культуры»», 1999.
6. *WordNet* <http://wordnetweb.princeton.edu/perl/webwn>
7. *Словарь синонимов русского языка* в двух томах (главн. ред. А. П. Евгеньева). Л.: «Наука», 1970, 1971.
8. *Национальный корпус русского языка*: www.ruscorpora.ru
9. *Малый академический словарь* — Словарь русского языка в 4-х томах // Электронный ресурс ФЭБ. — [М.]: 2005. <http://feb-web.ru/feb/mas/mas-abc/0encyc.htm>