

Опыт фронтального корпусного исследования конструкций с внутренним и внешним посессором

А.Е. Кибрик, М.М. Брыкина, А.Н. Хитров (МГУ)

1. Проблема

В работах [Кибрик 2000, 2003] были очерчены общие контуры так называемых конструкций с внешним посессором, противопоставленных синонимичным конструкциям с внутренним посессором, реализуемым в форме атрибутивной синтагмы. Например:

конструкция с внутренним посессором

конструкция с внешним посессором

1а. Вдруг Клара схватила **мою** руку].

1б. Вдруг Клара схватила **меня** [за руку]. (ЭО)

2а. С нас лились **ручьи горячего** пота].

2б. **Горячий пот** лился с нас [ручьями]. (ЭО)

3а. **Их мать**] умерла.

3б. **Мать** [у них] умерла. (ЭП)

4а. Я сунул все деньги **в его** карман].

4б. Я сунул все деньги **ему** [в карман]. (ЭП)

В (1б–2б) представлены конструкции с экстрапозицией обладаемого¹ (ЭО), в (3б–4б) – конструкции с экстрапозицией посессора (ЭП). При описании этих конструкций термины “посессор”, “обладаемое”, “посессивная конструкция”, восходящие к базовым атрибутивным отношениям (типа *моя рука, его карман, их мать*), используются расширительно, применительно ко всем атрибутивным словосочетаниям². Было установлено, что конструкции с внешним посессором совместимы почти со всем спектром атрибутивных значений, а не только с посессивным значением в узком смысле.

В сущности, в основе посессивной конструкции лежит абстрактное отношение когнитивной сопряженности, связывающее два концепта онтологически (по природе вещей) или ситуационно (в конкретной актуализованной ситуации). Типологически наиболее частотным синтаксическим способом выражения отношения когнитивной сопряженности является атрибутивная синтагма³. Наряду с этим когнитивная сопряженность может канонически выражаться парой синтаксически непосредственно не связанных именных групп. Типичным случаем является, например, наличие при глаголах типа *давать/отнимать* актанта-адресата/источника, онтологически когнитивно сопряженного с актантом-пациентом по собственно посессивному отношению (*Я знаю, кто дал моему братишке землянику // кто*

¹ Это сопровождается так называемым подъемом посессора, занимающего исходную позицию вершинного имени. В примерах посессор выделен подчеркиванием, обладаемое – курсивом.

² Включая такие, как *ручьи пота*.

³ Аtribuтивная синтагма состоит из пары именных групп, связанных подчинительным отношением. В русском языке она оформляется показателем родительного падежа (*дом отца*) в зависимом имени или посессивной формой зависимого местоимения (*мой, твой, наш, ваш, свой дом*). Возможно также ее оформление показателем в вершинном имени (изафетная конструкция в тюркских языках), в обоих именах (сопряженное состояние в арабском языке) или специальным служебным словом (английский предлог *of*). Аtribuтивная синтагма может также представлять собой контактное соположение зависимого и вершинного компонентов.

отнял у *братишки моего землянику*)⁴. Можно полагать, что наличие в языке независимых средств разрывного выражения когнитивно сопряженных концептов системно поддерживает формирование конструкций с внешним посессором (параллельных конструкциям с внутренним посессором).

Выбор типа посессивной конструкции невозможно описать в виде единого предписывающего правила, так как этот выбор является многофакторным: он одновременно контролируется несколькими автономными языковыми модулями (параметрами), значения которых совокупно влияют на предпочтительность той или иной посессивной конструкции. Факторы, обуславливающие невозможность / возможность / обязательность внешнего посессора, могут быть сгруппированы в характеристики трех типов: (а) ингерентные свойства посессора, (б) параметры на уровне предикации, (в) параметры на уровне дискурса. В настоящем исследовании рассматриваются только факторы, действующие в пределах предикации.

2. Ресурсы корпусного метода

Целью исследования было проверить и уточнить предложенную в [Кибрик 2000, 2003] модель на фиксированном представительном корпусе текстов таким образом, чтобы в рассмотрение были включены все имеющиеся в корпусе посессивные конструкции. Требование полноты корпусного описания придает ему ряд важных преимуществ перед традиционным описанием, основывающимся на прототипических образцах.

1. Корпусный подход реализует принцип первичности эмпирических данных. Не данные подбираются в соответствии с принятой теорией, а теория конструируется так, чтобы она адекватно описывала и объясняла произвольные данные.
2. К исследованию предъявляется требование адекватности лингвистического аппарата описания: для анализа любого релевантного события в корпусе (в нашем случае это посессивные конструкции) должно быть достаточно используемых в модели концептов и классификаций. Принцип адекватности предполагает, что ни одно встретившееся в корпусе событие не должно вступать в противоречие с предлагаемой теорией.
3. Корпусный подход обеспечивает доступ к статистическому распределению событий по значениям параметров, влияющих на тип события. Эта информация является существенной для понимания параметрической структуры факторов и их иерархической организации, см. ниже.
4. Доступ к кластеризации значений параметров. В рассматриваемой языковой области выбор конструкции является суммирующим результатом действия множества альтернативных значений различных параметров, и понять этот механизм невозможно, не исследуя поведение их всевозможных сочетаний. Собранный корпусным методом языковой материал позволяет представить эмпирическую информацию в виде эксплицитной базы данных, автоматизирующей такого рода исследование и дающей результаты, недостижимые традиционными эвристическими техниками.

Однако для того, чтобы иметь возможность полагаться на данные, прежде всего статистические, полученные в результате работы с корпусом, необходимо понимать, насколько мы можем им доверять, насколько они отражают реальное распределение событий, а не их распределение в отдельно взятом отрывке письменного дискурса. Другими словами, репрезентативность корпуса, безусловно, влияет на достоверность статистических данных. Возможность четкой формализации процедуры установления этой зависимости требует дальнейших исследований.

3. Построение корпуса и базы данных

Был сформирован экспериментальный корпус текстов объемом около 233 тысяч слов, в который вошли почти сорок нарративных литературных текстов различных авторов XX века, и проведена его разметка по основным релевантным параметрам.

⁴ Косвенным подтверждением такого анализа является существование языков типа фула, см. [Коваль, Нялибули 1997: 77, 203, 233], в которых при глаголах *давать [что, кому]* и *отнимать [что, у кого]* актант-адресат и актант-источник выражаются одинаково: это означает, что соответствующее средство кодирует не адресата и источник, а собственно отношение когнитивной сопряженности этих актантов с пациенсом.

Во-первых, были выделены все конструкции с внутренним и внешним посессором. Таких конструкций оказалось около 7000. Если атрибутивная конструкция в большинстве случаев обнаруживалась без осложнений⁵, то выделение конструкции с внешним посессором потребовало существенных уточнений самого этого понятия и отделения его от смежных явлений.

Во-вторых, каждой выделенной конструкции была приписана информация о всех ее релевантных свойствах в соответствии с разработанной классификацией. К релевантным свойствам относятся морфосинтаксические типы конструкций с внешним посессором и факторы, (не)благоприятствующие использованию этой конструкции.

Наряду с информацией о типе конструкции и морфологической форме экстрапонируемого элемента (для конструкций с внешним посессором) указывались параметрические значения следующих факторов: 1) семантический тип посессивного отношения, 2) тематический класс вершинного предиката, 3) тематическая роль актанта, выражаемого исходной посессивной конструкцией (конструкцией с внутренним посессором), 4) синтаксическая позиция актанта, выражаемого исходной посессивной конструкцией, 5) тип одушевленности посессора.

Для упрощения работы с размеченными конструкциями была создана база данных. Первичная разметка была осуществлена М.М. Брыкиной, С.В. Пискуновой и А.Н. Хитровым. В дальнейшем разметка многократно исправлялась и уточнялась на основании работы с базой данных и в связи с расширением числа параметров и уточнением их внутреннего категориального состава.

В-третьих, была проведена проверка корпуса на репрезентативность относительно исследуемого явления и каждого параметра. При этом использовалась немного модифицированная методика, главная идея которой описана в [Баранов 2001] и состоит в том, что репрезентативность можно оценить по изменению относительной частоты рассматриваемого явления при увеличении выборки. Если относительная частота явления от прибавления каждого последующего фрагмента текста будет изменяться все меньше и меньше, то это означает, что корпус в целом репрезентативен. На наш взгляд, эта методика не является идеальной по крайней мере потому, что не позволяет установить связи со статистикой, но формулируемое условие является, по-видимому, необходимым, хотя, возможно, и недостаточным для оценки репрезентативности корпуса.

Очевидно, что описанный экспериментальный корпус такого относительно небольшого объема не может считаться абсолютно репрезентативным. Однако, в результате оценки корпуса на репрезентативность, мы можем говорить о степени доверия к данным, относящимся к тому или иному параметру. В целом, данные, относящиеся к различным синтаксическим реализациям посессивной конструкции, а также к распределению синтаксических позиций и тематических ролей, представляются более достоверными, чем данные, касающиеся средних и нижних частей иерархий тематических классов предиката и семантических типов посессивных отношений. На данном этапе кажется нецелесообразным более подробно описывать эту область, так как пока не установлено строгой математической зависимости между репрезентативностью и статистикой.

4. Параметризация факторов

Как указывалось выше, факторы, влияющие на выбор типа конструкции, являются не атомарными сущностями, а параметрами, имеющими множество значений. Эти параметры играют существенную роль не только в грамматике посессивных конструкций, но и в самых различных модулях языка. Тем не менее, в процессе создания базы данных стало очевидно, что большинство параметров не имеют общепринятого и удовлетворительного описания. Поэтому параметризация факторов происходила в несколько этапов. В основу параметризации был положен принцип эмпирического соответствия выделяемых значений (типов, реализаций) параметров их статистическому поведению в базе. Текущее состояние параметризации (на февраль 2004 г.) представлено ниже.

4.1 Синтаксические отношения

Этот параметр в нашей классификации имеет следующие значения: **SU** субъект, **DO** прямой объект, **LOC** дополнение места, **NP** посессивная группа в составе объемлющей именной группы (совпадает с глагольным

⁵ Конструкции типа *машинное масло, собачий хвост, бабушкино кресло* не рассматривались.

классом), **PR** предикат (неэлементарная ИГ в составе именного сказуемого), **OBL** косвенное дополнение/обстоятельство.

4.2 Тематические роли

Стремясь к минимизации состава тематических ролей, мы ограничились следующими значениями:

AG_t агенс переходного глагола, **AG_i** агенс непереходного глагола, **EFF** эффектор (неодушевленный деятель, изменяющий ситуацию⁶), **EX** экспериенцер, **ST** стимул, **PA** пациенс, **TH** тема, **TAR** цель, **PL** место, **VEN** бенефактив, **PR** предикат (неэлементарная ИГ в составе именного сказуемого, совпадает с синтаксическим отношением), **NP** посессивная именная группа в составе объемлющей именной группы (совпадает с синтаксическим отношением), **OTH** прочее.

4.3 Тип посессивного отношения

Были выделены следующие 26 типов посессивных отношений (в названии отношения первый член – зависимое имя в генитиве (посессор), второй член – вершинное имя (обладаемое)):

1. AD: отношение смежности “сущность – смежная сущность” (*спутник земли, окрестности города, кобура пистолета*)
2. ATR: атрибутивное отношение “атрибут сущности – сущность” (*площадь Ногина, таблица умножения, закон инерции*)
3. BP: отношение части тела “личность – часть тела личности” (*его голова / рука / челка, моя душа, мой ум*)
4. CL: “личность – одежда / часть одежды личности”⁷ (*его рубаша / сапог / хлястик / рукав / штанина / ошейник / рюкзак*)
5. EXM: отношение экземпляра “явление / сущность – экземпляр ~” (*экземпляр книги, штука кафеля, пучок травы, початок кукурузы, глоток вина, капля крови, лист бумаги*)
6. FM: отношение формы “сущность – форма сущности” (*квадрат полей, круг зрителей, слой обоев, полоса света, пружина дыма, гроздь форели, обрубок хвоста*)
7. GR: групповое отношение “однотипные элементы группы – группа” (*их артель / группа / бригада / колонна / клуб / команда / партия / кружок / отряд / эскадра, их барак / родина, пара людей, колода карт, букет цветов, коллекция картин*)
8. GRsg: отношение включения в группу “элемент группы – группа” (*его команда / компания / семья / бригада*)
9. KIN: отношение родства “личность – родственник личности” (*отец Пети, жена президента*)
10. MAT: отношение материала “вещь – материал вещи” (*вата подкладки, бревна галереи, древесина удиллица*)
11. MB: отношение членства “совокупность / коллектив / организация – член ~” (*член группы, передовик производства, капитан корабля, чемпион мира*)
12. MES: отношение количества “вещество – количество вещества” (*литр вина, туча комаров, излишек энергии, процент успеваемости*)
13. MRH: метафорическое отношение “сущность – метафора сущности” (*вуаль печали, марля тумана, ягоды зрелости, сети рабства, змеи молнии, щетина забора*)
14. OB: объектное отношение “объект события – событие” (*вывих ступни, раздача / выпечка хлеба, ремонт пути, поглощение пищи*)
15. OB': объектное отношение “объект действия – субъект действия” (*искательница приключений, подносчик шлакоблоков*)
16. PAR: параметрическое отношение “сущность – параметр сущности” (*ее характер / внешность / поза / взгляды, история болезни, смысл слова, тематика совещаний*)
17. PLC: отношение места “сущность – место / время сущности” (*наш ресторан / край / завод; поле битвы, дата смерти*)
18. PLC': отношение “место / время сущности – сущность” (*обитатели дома, проза будущего*)
19. POS: посессивное отношение “обладатель – обладаемое” (*его дом / спальня / койка / тумбочка / чемодан / окурки / лодка*)

⁶ EFF был введен потому, что, как оказалось, роль эффектора существенно отличается от роли агенса по ее корреляции с внешним посессором, что говорит о том, что она была явно недооценена в теории тематических ролей.

⁷ Выделение этого типа было предложено и обосновано А. Леонтьевым.

20. POS*: посессивное отношение “обладаемое – обладатель” (*владелец чемодана / дачи, обитатели дома*)
21. PRD: отношение продукта “деятель / сущность – продукт ~” (его слова / речь / вопрос / идея / мысль / мнения / поступки; зола костра, дым папиросы, результат дела, следы разгула, плоды трудов, отходы производства, звук шагов, лучи солнца)
22. PRD*: отношение продукта “продукт деятельности – деятель / причина” (*автор цитаты, источник мучений*)
23. SOC: социальные отношения “личность – социальный партнер личности” (наш хозяин / знакомый / товарищ / противник / враг, друг мастера, ровесник Кузьменышей, помощник бригадира)
24. SUB: субъектное отношение “субъект события / состояния – событие / состояние” (его действия / занятия / старание / усилия, простой бригады, крушение парохода, прохлада комнат)
25. SYM: символическое отношение “сущность – символ сущности” (ее имя / фамилия / адрес / отражение, названия гор, заглавие рассказа, номер дела / магазина, список членов, карта Прованса, фотография дочери, изображение пики, знак радости, эмблема власти, гримаса разочарования)
26. WH: отношение часть-целое “сущность (целое) – часть сущности” (ограда парка, огонек папиросы, двери склада, ложе реки, глубина озера, правление колхоза, развалины санатория)

4.4 Тематический класс предиката

Была разработана классификация типов глаголов⁸ с точки зрения их влияния на поведение посессивных конструкций. Существенное значение для определения глагольных классов имело также единство тематических ролей глагольных актанта. В целом в этой классификации имеется порядка сорока типов глаголов.

1. act – <SU: од AG_т, неод EFF; DO: PA> действие: физическое воздействие, изменение состояния (*арестовать, давить, наполнять, рвать*)
2. actI – <SU: AG_т / PA [OBL: PA / TAR]> процесс: физическая активность (греметь, напирать, реветь, махать чем (PA), выпиться в (TAR), целиться в (TAR))
3. advp – ИГ в составе наречной группы (для отвода глаз, после отсылки письма, в пору зимы)
4. aid – <SU: од+перех AG_т, од+неперех AG_т, неод EFF; DO / OBL: TAR> помогать, влиять на контрагента речью или действиями (*благославлять, оскорбить, уважать, мешать чему, помогать кому*)
5. app[ear] – <SU: TH> появление, возникновение (*возникать, повториться, родиться, явиться*)
6. ask – <SU: AG_т, DO: PA (OBL: TH)> просить; изменение ментального состояния (речью), перлокутивная речь (вопросать кого, воспитывать, звать, приучать кого к чему (TAR), просить (кого: PA) (о чем TAR), уверять в чем (TH))
7. be – <SU: TH>; <OBL: PL, SU: TH> находиться что где, иметься что, существовать (*быть, значиться где, оказаться, покоиться*)
8. beat – <SU: AG_т, DO / OBL: PA, EX> контакт с телом (PA) с целью причинения боли лицу (EX) (*бить по/в, кого/кому, ткнуть в, кого, дать по, кому, отбить что, ранить что, ущипнуть что*)
9. cont[act] – <SU: пер.од. AG_т / непер.од. AG_т / неод. TH; DO / OBL: TAR> физический контакт (*брат/взять, гладить, задеть, прижиться к, хватать, целовать, касаться чего*)
10. dep[end] – <SU: TH, OBL: ST> (не)зависимость (*зависеть от, (не) обходиться без*)
11. dest[roy] – <SU: AG_т, DO: PA> разрушение (брить, сбрить, вычеркивать, душить, испортить, рвать)
12. disp[osition] – <SU: TH> состояние: положение в пространстве / времени (возвышаться, громоздиться, доходить до, обращен в / на, размещаться, свешиваться, стоять)
13. do – <SU: AG_т, DO: TAR> создание, реализация (варить, вязать, затевать, писать, приступить к, родить, создавать)
14. fall – <SU: PA> неконтролируемый процесс, изменяющий состояние (вздрагивать, голодать, застрять, испариться, опоздать, наполняться, ожить, падать, скользить, болен)
15. go – <SU: одуш AG_т / неодуш PA> процесс: двигаться в пространстве (*бежать, возвращаться, катиться, остановиться, проникать, трогаться*)
16. have – <SU: кто TAR, DO / OBL: что / чем TH> <SU: что TH, OBL: кому TAR> обладание (*иметь, владеть / овладеть, лишиться, терять, принадлежать, свойственно*)

⁸ Тип глагола влияет на посессивную конструкцию, когда эта конструкция непосредственно подчинена глаголу. Однако посессивная конструкция может непосредственно подчиняться не глаголу, а входить в именную или наречную группу. Этим ситуациям соответствуют классы *pr* и *advp*, условно включаемые в данную классификацию.

17. look – <SU: пер.од. AG_i / пер.неод EFF / непер.од AG_i; **DO / OBL:** TAR> активное восприятие: **ощущение** (*вдыхать, всматриваться, глядеть на, любоваться чем / на что, ожидать*), **память-процесс** (*воскресить в памяти, вспоминать, познавать*)
18. ment[al] – <SU: пер.од. AG_i / непер.од. AG_i / пер.неод EFF / непер.неод PA; **DO / OBL:** TH; [3-ий АКТАНТ] **OBL:** TAR / дат.од. BEN / дат.неод. TAR> ментальные действия и процессы **перех.** (*вешать, взвешивать, выдумывать, выяснить, избрать что чем TAR, измерить, обещать что кому BEN, подтвердить, принять за TAR, посвящать что кому BEN, разрешать что кому BEN*), **непер.** (*внять чему, думать о, интересоваться чем, противиться чему, ручаться за, судить о чем*)
19. mov[e] – <SU: AG_t, **DO:** PA > перемещение объекта (бросать, волочить, вращать, вытеснять, наливать, приблизить, собирать, тянуть)
20. mov` – <SU: AG_t, **DO:** TAR> движение относительно ориентира (достигать, на(до-)гнать, переехать что/за что, покинуть, пройти)
21. need – <SU: TAR, **OBLdat:** BEN> необходимость, желательность, полезность, реализуемость (*необходим, нужен, обязателен, полагается, удаваться*)
22. nr – посессивная группа в составе объемлющей именной группы ([*фляга [масла [откуда-то из запасов ОРСа]]*], [*проводник [в форме железнодорожника]*])
23. occ[ur] – <SU: од. AG_i /неод EFF; [OBL: TAR]> процесс: жизненная деятельность (вести себя, курить, жить, заниматься чем TAR, купаться, обедать, отдыхать, работать, упражняться в TAR, участвовать в чем TAR, учиться чему TAR)
24. opin[ion] – <SU: TH, **DO / OBL:** ST> мнение (быть (лучшего) мнения о брезговать чем, верить кому / во что, возлагать надежды на TAR, надеяться на TAR, согласен с, убежден в,ценить что)
25. [inter]pers[onal] <SUном: од EX, неод PA; **SUдат:** EX; **OBL:** ST> межличностное взаимодействие (*драться с, перегляднуться с, подружиться с, ссориться с, целоваться с*)
26. psy[ch] – <SUном: од EX, неод PA; **SUдат:** EX; **OBL:** ST> психическое состояние, непереходное ((не) выносить чего, бояться кого, влюблен в, восхищаться чем, горд чем, изнывать по, млеть от, опасаться за, пугаться чего, равнодушен к, сердиться на, скучать о, упиваться чем)
27. qual[ity] – <SU: TH> качество, свойство (виновен, вред, вреден, готов, неправ, тоньше, (никуда не) годится, иметь значение)
28. quan[tity] – <SU: TH> количество (*хватает, что было (у нее силы)*)
29. real[ize] – <SU: TH> реализация (достигнуто, завершено, получиться, происходить, разразиться, сбываться, совершиться, удаваться)
30. res[ult] – <SU: TH> результирующее качество, свойство (*мертв, виновен, окреп, повидал, погребен, разложен*)
31. save – <SU: AG_t / EFF, **DO:** TH> (не) сохранять существование (беречь, защищать, компенсировать, спасать кого от EFF, сторожить, чинить)
32. say – <SU: AG_t / AG_i, **DO:** TH, **OBL:** кому BEN> речь (болтать, вопрошать что, кого TAR, звать [=называть] кого TAR кем TH, подсказывать, сказать, упоминать)
33. see – <SU: од EX, неод PA; **DO / OBL:** ST> восприятие, чувство, память: **восприятие зрением, слухом, обонянием** (*видеть, видно, видать, встретить, дышать чем, обонять, ощущать, слышать, созерцать, чуют*), **ментальное состояние** (*быть X-го мнения о, ведать чем, знать что / о чем, иметь понятие о, помнить*), **чувство (перех)** (выдерживать что, любить, чувствовать, хотеть)
34. see` – <SU: од AG_t / неод ST; **DO / OBL:** од EX / неод TAR> стимулирование восприятия, эмоции, чувства (будоражить, возбудить, возмутить, волновать, ласкать, мучить, оживлять, поразить, пугать, терзать, грозить кому, непонятно кому, приелся кому, тепло кому, хотеться кому, ясно кому)
35. sym[metry] – <SU: TH; [OBL: OTH]> симметричный предикат (отличаться от, быть под стать чему, быть идти вразрез с, одинаков, походить на, сродни чему, противоречить чему)
36. take – <SU: AG_t, **DO:** TH> получение, передача собственности (брать / взять / отбирать, давать, делиться чем TH с кем BEN, дарить кому BEN, отнимать, покупать, наделять кого BEN, чем: TH, получать кто BEN, что TH)
37. tax[onomy] – <SU: TH, **PR:** PR> вхождение в класс объектов (мой дядя матрос, представлять собой, служить чем кому BEN, числиться кем)
38. use – <SU: пер. AG_t / непер. AG_i; **DO / OBL:** PA> расходование (есть, пить, поглощать, тратить, воспользоваться чем, пользоваться чем)
39. vis[ual] – <SU: од AG / неод EFF; **DO / OBL:** TAR> делать (не)доступным восприятию (выражать, демонстрировать, напоминать что / о чем кому EX, обнажать, оттенять, прикрывать, указывать на)
40. visI – <SU: ST; [OBL: EX]> быть / становиться (не)доступным восприятию (благоухать, виднеться, выступать, исчезать, открываться (взору), пахнет, славиться, греметь, маячить, мелькать, хрустеть, белеет, струиться)

5. Статистические тенденции

Поскольку значения каждого параметра различаются тем, в какой степени они благоприятствуют типу посессивной конструкции, естественным проявлением этого свойства является статистическое распределение значений параметров относительно конструкции с внутренним и внешним посессором. Созданная база данных позволяет получить эту статистику.

Возможны два равнозначных способа статистической оценки корреляции между типом посессивной конструкции и значением параметра.

а) От значения параметра к типу конструкции. Фиксируется значение параметра. Число всех событий в базе с данным значением параметра принимается равным 100%, и в пределах этой совокупности определяются частоты (в процентах) конкретных типов посессивной конструкции. Частота конкретного типа конструкции при конкретном значении параметра сравнивается со среднестатистическим процентом такой конструкции среди всех посессивных конструкций в базе. Если частное от деления первой величины на другую равняется единице, данное значение параметра нейтрально к данному типу конструкции, если оно меньше единицы, то данное значение параметра не благоприятствует этому типу конструкции.

Значение параметра тем более БЛАГОПРИЯТСТВУЕТ данному типу конструкции, чем процент таких событий выше среднестатистического.

б) От типа конструкции к значению параметра. Фиксируется тип конструкции с внешним посессором (ЭП, ЭО или ЭП+ЭО). Число событий с данным типом конструкции в базе принимается равным 100%. В этой совокупности событий определяются (в процентах) частоты событий, реализующих конкретные значения конкретного параметра. Далее эти величины сравниваются с независимо вычисленными процентами всех имеющихся в базе посессивных конструкций с данными значениями параметра. Если частное от деления первой величины на другую равно единице, то данное значение параметра нейтрально к данному типу конструкции, если меньше единицы – то оно не благоприятствует данной конструкции.

Чем больше частное от деления первой величины на другую, тем более данное значение параметра БЛАГОПРИЯТСТВУЕТ данному типу конструкции.

Ниже мы пользуемся первым способом как более наглядным, хотя для некоторых обобщений второй способ может быть более предпочтительным.

5.1 Относительная доступность синтаксических отношений внешнему посессору

В таблице 1 и следующих приводятся статистические данные о частоте внешнего посессора (ВнешП) в зависимости от синтаксической позиции посессивной именной группы. Синтаксические отношения упорядочены в соответствии с общей частотой ВнешП (первая строка таблицы: ЭП+ЭО). В колонке “Всего” указана процентная доля всех конструкций с ВнешП по отношению к общему числу посессивных конструкций в базе. Эта колонка разделяет синтаксические отношения, благоприятствующие конструкции с ВнешП (слева от данной колонки) и не благоприятствующие (справа от данной колонки).

Во второй строке указаны частоты конструкций с ЭП, в третьей – частоты конструкций с ЭО. Клетки, в которых эти частоты превышают среднестатистические значения, выделены заливкой.

В таблице 1 различаются синтаксические отношения: субъект, (прямой) объект, место, предикат, косвенное дополнение – обстоятельство, а также ИГ (если посессивная группа находится внутри объемлющей именной группы).

Таблица 1. Иерархия доступности синтаксических отношений для внешнего посессора

	LOC	SU	DO	Всего	PR	NP	OBL
ЭП+ЭО	23,7%	22,0%	16,9%	14,9%	3,9%	0,6%	0,6%
ЭП	23,7%	12,2%	5,7%	9,6%	3,9%	0,5%	0,5%
ЭО	0,1%	9,8%	11,1%	5,3%	0,0%	0,2%	0,1%

Примеры на ЭП:

- LOC:** 5а. Положа [Матрене] руку [*на плечо*], сказала строго. {BP/PL/mov}
 5б. Парень, стоявший рядом со мной, шарит [у меня] [*в кармане*]. {CL/PL/actI}

- SU:** 6а. [**Глаза**] [**у женщины**] были черные. {BP/TH/qual}
 6б. [**У меня**] хватило [**мужского самообладания**]. {SUB/TH/quan}

Примеры на ЭО:

- DO:** 7а. Мы все знали [**ее**] [**в лицо**] хорошо. {BP/ST/see}
 7б. Тогда он подошел к ней, взял [**ее**] [**за руку**] и ответил. {BP/TAR/cont}
 7в. Почему-то я решил, что правильной будет [**ее**] поцеловать [**в сияющее личико**]. {BP/TAR/cont}.
- SU:** 8а. [**Он**] [**ненавидящими глазами**] следил за нею. {BP/EFF/look}
 8б. [**Я**] чувствовал это [**кожей лица**]. {BP/EX/see}
 8в. [**Очкарик**] вздрогнул [**всем телом**] {BP/PA/fall}
 8г. И только [**лоб**] уходил [**лысым куполом**] в лысую просторную маковку. {WH/TH/disp}

В целом субъект, место и объект превышают среднюю частоту конструкций с ВнешП, то есть наиболее благоприятствуют этой конструкции. При этом субъект благоприятствует как экстрапозиции посессора (12,2%), так и экстрапозиции обладаемого (9,8%). Что касается места, то оно допускает только конструкцию с ЭП. Прямой объект, наоборот, благоприятствует экстрапозиции обладаемого (11,1%), хотя и допускает экстрапозицию посессора. Прочие отношения находятся в конце иерархии, которая имеет вид:

SU, LOC > DO > PR > NP, OBL

5.2 Относительная доступность тематических ролей внешнему посессору

Таблица 2. Иерархия доступности тематических ролей для внешнего посессора

	EFF	PA	PL	TH	TAR	Всего	AGi	EX	ST	PR	AGt	oth	NP	BEN
ЭП+ЭО	23,9%	23,3%	21,4%	21,2%	17,3%	14,9%	13,4%	10,3%	8,0%	4,3%	2,3%	1,1%	0,7%	0,0%
ЭП	1,1%	10,3%	21,2%	16,3%	7,7%	9,6%	1,6%	2,6%	1,4%	4,3%	1,1%	0,9%	0,4%	0,0%
ЭО	22,8%	13,0%	0,2%	4,9%	9,6%	5,3%	11,8%	7,7%	6,6%	0,0%	1,1%	0,3%	0,3%	0,0%

Примеры на тематические роли, благоприятствующие внешнему посессору:

- EFF:** 9. [**Ведро воды**] заменяет [**по калорийности**] сто граммов масла {MES/SU/save}
- PA:** 10а. Фадеев ... повернул [**меня**] [**лицом**] к небу носками своих сапог {BP/DO/mov}
 10б. [**Тебе**] солнце напекло [**затылок**] {BP/DO/act}
- PL:** 11. Капитан сопел и ковырялся в лодке [**у меня**] [**за спиной**] {BP/LOC/actf}
- TH:** 12. Начальник выписывал в кузницу [**сливочное масло**] – [**в ничтожном, конечно, количестве**]. {MES/DO/ment}
- TAR:** 13. Бедный человеческий мозг просто не в силах представить [**в конкретных образах**] [**тамошнюю жизнь**] {PAR/DO/look}

Пять первых ролей имеют ВнешП статистически чаще, чем в среднем по базе, то есть про них можно сказать, что они благоприятствуют конструкции с ВнешП.

Наиболее стабильно поведение пациенса, который благоприятствует как ЭО, так и ЭП. При нем оба типа ВнешП имеют значения выше среднестатистических (13% > 5,3% для ЭО и 10,3% > 9,6% для ЭП). Что касается темы и цели, то они дополнительно распределены: тема существенно предпочитает ЭП, а цель – ЭО, при том, что альтернативный тип ВнешП также допустим, но несколько ниже среднестатистической величины.

Высокое положение эффектора в Таблице 1 целиком зависит от того факта, что эффектор решительно благоприятствует ЭО при почти полной его неспособности к ЭП. Место, наоборот, лидирует по способности к ЭП при невозможности ЭО.

Пять ролей, благоприятствующих ВнешП, образуют первую группу в иерархии. Она покрывает 63% всех посессивных конструкций, представленных в базе.

Далее идет вторая группа ролей: непереходный агент, экспериенциер и стимул. Эти роли в целом реже имеют конструкцию с ВнешП, чем в среднем по базе, хотя они благоприятствуют ЭО (11,8% для непереходного агента, 7,7% для экспериенциера и 6,6% для стимула). Их низкие показатели для ВнешП в целом связаны с тем, что они слабо доступны ЭП. Тем не менее эту группу можно считать частично благоприятствующей конструкции с ВнешП. Она составляет 11% посессивных конструкций.

Наконец, третья группа, охватывающая пять ролей, представленных в 26% посессивных конструкций, слабо способна к конструкции с ВнешП. Это предикат, переходный агент, бенефактив, прочее и именная группа.

5.3 Относительная доступность посессивных отношений внешнему посессору

Таблица 3. Иерархия посессивных отношений для внешнего посессора

	BP	CL	PLC	FM	GR	PAR	Всего	POS	WH	SOC	GRsg	SUB	KIN	MES	PRD	EXM
ЭП+ЭО	40,0%	34,1%	26,7%	22,1%	17,6%	15,2%	14,9%	10,6%	7,7%	7,6%	7,4%	7,2%	6,8%	6,4%	6,0%	5,9%
ЭП	27,9%	22,0%	25,9%	0,0%	2,4%	9,3%	9,6%	10,5%	1,1%	7,6%	7,4%	3,9%	6,8%	0,0%	3,5%	0,0%
ЭО	12,0%	12,2%	0,8%	22,1%	15,3%	5,8%	5,3%	0,1%	6,6%	0,0%	0,0%	3,3%	0,0%	6,4%	2,5%	5,9%

	MB	OB	MPH	SYM	AD	ATR	MAT	OB`	POS*	PRD*
ЭП+ЭО	2,0%	1,5%	1,2%	1,2%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%
ЭП	2,0%	1,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%
ЭО	0,0%	0,5%	1,2%	1,2%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%

Проиллюстрируем типы отношений, не встречавшиеся в предыдущих примерах⁹:

PLC: 14а. В этот день мы ... пообедали [у меня] [в гостинице]. {LOC/PL/occ}

14б. А [у ээков] [день] большой {SU/TH/qual}

FM: 15. [Такие варианты] носились [роем] в голове Сашки. {SU/PA/go}

GR: 16. [Их] потом по Москве [колонной] водили. {DO/PA/mov}

Выше среднестатистической величины ВнешП располагаются шесть отношений. Отношения частей тела и одежды однозначно занимают первое место в иерархии, существенно превышая как среднюю частоту конструкций с ВнешП в целом, так и частоту конструкций с ЭП и с ЭО. Иными словами, одежда когнитивно объединяется с частями тела.

Отношения места, с одной стороны, и формы и групповое, с другой, дополнительно распределены с точки зрения использования ЭО и ЭП, также достоверно превышая среднестатистическую частоту ВнешП.

Параметрическое отношение незначительно превышает среднестатистическую величину, при этом оно допускает как ЭО, так ЭП с величинами, очень близкими к среднестатистическим. Группа следующих отношений имеет весьма близкие статистические значения (в пределах статистической погрешности). Наконец, имеется группа отношений, надежно не благоприятствующих конструкции с ВнешП.

Иерархия посессивных отношений по статистическим данным имеет вид:

⁹ Отношение BP встречалось в примерах (5а, 6а, 7а–в, 8а–в) и др., CL см. в (4а, 5б), PAR см. в (12, 13), FM см. в (2а).

BP, CL > PLC, FM, GR > PAR > POS, WH > SOC, GR_{sg}, SUB, KIN, MES, PRD, EXM > прочие

5.4 Относительная доступность глагольных классов внешнему посессору

Таблица 4. Иерархия доступности глагольных классов для внешнего посессора

	beat	quan	cont	qual	res	be	app	dest	act	see`	fall	disp	look	actl	Всего	mov
ЭП+ЭО	69,1%	60,7%	49,3%	34,7%	25,5%	23,8%	21,7%	20,4%	19,6%	18,6%	17,0%	16,2%	16,2%	15,0%	14,9%	14,4%
ЭП	11,3%	53,6%	4,9%	31,1%	21,8%	22,2%	21,7%	18,5%	12,3%	11,2%	13,9%	9,4%	11,8%	10,6%	9,6%	7,6%
ЭО	57,7%	7,1%	44,4%	3,6%	3,6%	1,6%	0,0%	1,9%	7,3%	7,5%	3,0%	6,8%	4,4%	4,4%	5,3%	6,7%

	ask	occ	sym	take	visl	go	vis	aid	say	use	see	opin	ment	tax	need	psy	real
ЭП+ЭО	13,8%	13,4%	13,2%	10,7%	10,6%	10,6%	9,5%	9,1%	8,5%	7,8%	7,7%	7,4%	6,9%	6,5%	6,3%	4,9%	4,8%
ЭП	0,0%	11,8%	2,6%	10,2%	3,3%	6,9%	3,8%	0,0%	7,2%	4,7%	2,6%	0,0%	2,8%	6,2%	6,3%	3,3%	4,8%
ЭО	13,8%	1,6%	10,5%	0,6%	7,3%	3,7%	5,7%	9,1%	1,3%	3,1%	5,2%	7,4%	4,1%	0,3%	0,0%	1,6%	0,0%

Таблица 4 приведена не полностью. Отсутствует ряд классов, имеющих весьма низкую частоту конструкций с ВнешП. Обращает на себя внимание тот факт, что достаточно большое число глагольных классов благоприятствуют ВнешП. Особенно выделяется класс глаголов причинения вреда, большинство членов которого допускают только конструкцию с ВнешП. Интересно также особое поведение глаголов контакта, явно предпочитающих ЭО. Большинство других классов, благоприятствующих ВнешП, преимущественно выделяют ЭП.

5.5 Относительная доступность ИГ по типу одушевленности внешнему посессору

Таблица 5. Иерархия доступности значений одушевленности для внешнего посессора

	2	3	HU	1	AN	Всего	IN	REFL
ЭП+ЭО	26,5%	25,5%	22,2%	21,3%	19,9%	14,9%	5,6%	5,6%
ЭО	3,3%	7,4%	5,6%	5,1%	9,3%	5,3%	4,9%	2,1%
ЭП	23,2%	18,0%	16,6%	16,2%	10,6%	9,6%	0,7%	3,4%

Расположение местоимений 1-3 лица имен людей в таблице нельзя считать статистически достоверным. Существенно, что во всех случаях имеет место явное предпочтение конструкции с ВнешП и, в частности, ЭП. Интересно, что при 1-2 лице частота ЭО ниже среднестатистической нормы. 3-е лицо сближается с личными (HU) и одушевленными (AN) посессорами в отношении нарастания конструкций с ЭО и превышении их над среднестатистическими значениями. Конструкции с неодушевленными (IN) и рефлексивными (REFL) посессорами не благоприятствуют ВнешП. В целом иерархия одушевленности имеет вид:

1 / 2 / 3 / HU > AN > IN > REFL

6. Кластерные характеристики благоприятствующих факторов

Особого внимания заслуживают данные о взаимодействии значений различных параметров. Исследовались кластеры значений первых четырех параметров. При астрономическом большом количестве теоретически возможных сочетаний реализаций, в исследуемом корпусе всего семь статистически достоверно благоприятствующих кластеров покрывают 20% всех конструкций с внешним посессором.

Таблица 6. Кластеры значений параметров, благоприятствующие конструкции с внешним посессором

№ пп	Параметры				%% конструкций с внешним посессором	Примеры (обладаемое выделено курсивом, посессор – подчеркиванием)
	синт. отн.	сем. роль	посс. отн.	глагол. класс		
1	DO	PA	BP	beat	100%	<i>Я вспомнил того, молодого, [которого] я [в ногу] ранил.</i>
2	DO	PA	BP	act	52%	<i>И в следующее мгновение часовой почувствовал, как крепкая сила рванула [его] [за ноги, за руки].</i>
3	SU	TH	BP	qual	62%	<i>[У меня] [душа] крепкая, - объяснял Шура.</i>
4	SU	TH	PAR	qual	52%	<i>[Прованс] неистощим [в своих красотах].</i>
5	SU	TH	SUB	qual	32%	<i>[Дыхание] [у меня] коротко, сердце томится. [Она] становится неуголимой [в своей щедрости].</i>
6	SU	PA	BP	fall	27%	<i>[Спина] [у меня] никогда не заживает.</i>
7	SU	ST	PAR	see`	24%	<i>[Материальность нашей психики] впервые представляла мне [во всей наглядности, во всей ощутимости].</i>

7. Заключение

Корпусные данные позволяют, с одной стороны, проверять существующие теоретические выкладки и обобщения, например, действие широко известной иерархии одушевленности Силверстейна, а с другой, принимать отнюдь не тривиальные, статистически обоснованные решения, отражающие взгляд на природу вещей из самого языка. К наиболее ярким из таких решений, диктуемых, подчеркнем, самим языковым материалом, можно отнести выделение тематической роли эффиктора или отождествление надетой одежды с частями тела. Есть ряд и других примеров, в частности разделение группового отношения на собственно групповое (GR) и отношение включения в группу (GR_{sg}), ведущих себя статистически неодинаково. Обнаружение подобных случаев стандартными исследовательскими методами может быть значительно затруднено, а при обнаружении недостаточно обосновано.

Нельзя, однако, не учитывать, что эффективность корпусного метода, обусловлена соблюдением ряда предварительных условий, среди которых на первом месте, несомненно, стоит репрезентативность. Проблема установления репрезентативности данных не нова, однако существующие решения этой проблемы, как представляется, ещё требуют своего уточнения и пока обозначают только основное направление продвижения в данной области, а не конечный пункт.

Литература

1. Баранов А.Н. Проблема репрезентативности корпуса данных. // Труды международного семинара Диалог '2001 по компьютерной лингвистике и ее приложениям. Аксаково, 2001.
2. Кибрик А.Е. Внешний посессор как результат расщепления валентностей. // Слово в тексте и словаре. К 70-летию Ю.Д. Апресяна. М.: ЯРК, 2000: 429–441
3. Кибрик А.Е. Внешний посессор в русском языке // Кибрик А.Е. Константы и переменные языка. СПб: Алетейя, 2003: 307–319.
4. Коваль А.И., Нялибули Б. Глагол фула в типологическом освещении. М.: Азбуковник. 1997.