

ПАССИВ НА –СЯ В РУССКОМ ЯЗЫКЕ: КОРПУСНОЕ ИССЛЕДОВАНИЕ

Коломацкий Д. И.

Московский государственный университет им. М. В. Ломоносова, Москва

lemon_tree@list.ru

Предметом исследования является русский финитный пассив. Задачей было разбить массив корпусных данных на классы и изучить их свойства, а затем попытаться развить теорию, которая позволяла бы по набору свойств предложения с –ся-глаголом предсказать его не/пассивность и объяснить недопустимость некоторых пассивных предложений с глаголом на –ся.

0. Введение

Данная работа описывает результаты проведённого автором корпусного исследования русского финитного (называемого иногда также «глагольным») пассива — пассива на –ся. Почему исследование посвящено только постфиксу –ся? Дело в том, что обычно в работах лингвистов исследуется пассив в целом, даже в русистике. Можно встретить работы по страдательным причастиям. Гораздо меньше работ по постфиксу –ся в его пассивном значении. Когда же в работе эти два средства объединяются автором, это можно считать правомерным, поскольку они оба выражают пассив. В то же время часто забывают об их различиях, и не столько в распределении их по классам глаголов, сколько в их природе. Страдательное причастие имеет только одно значение — значение пассива. Постфикс –ся имеет много значений, и при этом исторически пассивное значение вовсе не являлось первичным (известно, что постфикс –ся произошёл от рефлексивного местоимения). Именно поэтому он особенно интересен для исследователя.

Основной задачей исследования было разбить массив корпусной выборки на два класса — пассивный и непассивный (такое деление вовсе не означает, что автор придерживается точки зрения о привативности залоговой оппозиции, а объясняется направленностью его на изучения пассива в противоположность «всему остальному») и исследовать различные свойства предложений каждого класса. Если для каждого класса удастся выявить характерные свойства, возможно будет создать предсказательную и объяснительную теорию, которая, с одной стороны, способна предсказать пассивное значение у предложения с глаголом на –ся, а с другой стороны, объяснить, почему так в пассиве сказать можно, а так нельзя.

Исследование состояло из трёх основных этапов:

- 1) Разработка инструмента распределения предложений по двум классам;
- 2) Обработка корпусных данных с помощью этого аппарата;

- 3) Анализ свойств пассивных и непассивных предложений;
- 4) Разработка теории

1. Определение пассива для разбиения предложений по классам

Определение пассива мы заимствуем из работы К. И. Казенина [Kazenin 2001], который, в свою очередь, ссылается на определение Мартина Хаспельмата [Haspelmath 1990]. Согласно К. И. Казенину, конструкция считается пассивной, если:

- 1) используемая в конструкции глагольная форма каким-либо образом морфологически образована от формы, используемой в немаркированной (активной) конструкции;
- 2) деятель не выражен или выражен факультативным косвенным дополнением с производной глагольной формой;
- 3) субъект, если таковой присутствует, не является деятелем;
- 4) конструкция каким-либо образом ограничена в своей дистрибуции по отношению к немаркированной (активной) конструкции;
- 5) пропозиционная семантика конструкции идентична пропозиционной семантике конструкции немаркированного (активного) залога; в частности, число партиципантов и их роли не меняются.

Это определение было мной выбрано не случайно. Во-первых, это определение вполне согласуется с тем, что было уже сказано по поводу значения пассива большинством лингвистов, а, во-вторых, достаточно эксплицитно и чётко сформулировано, чтобы служить набором параметров для обработки корпусных данных.

2. Корпусное исследование: общая статистика

Языковой материал набирался из двух источников: корпуса русских текстов Тюбингенского университета (Tübingen Russische Korpora, 727 предложений), в который входят материалы советской публицистики, и произведения художествен-

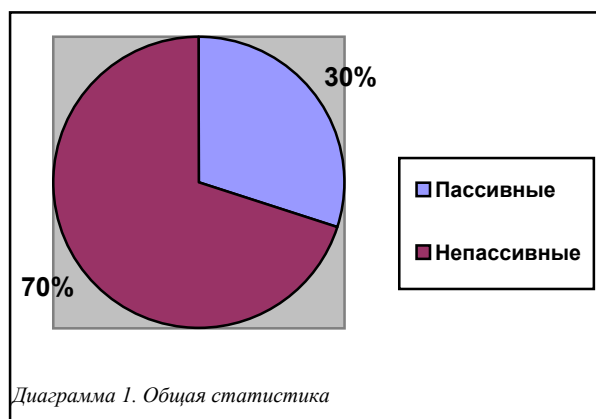
ной литературы — (Корпус русского языка Яндекс, 979 предложений). Всего в базе, таким образом, 1706 предложений, каждое из которых содержит один глагол с постфиксом -ся.

Обработка материала производилась следующим образом: каждому предложению приписывались значения параметров, заимствованных из определения пассива выше, после чего примеры сортировались по двум классам — пассивному и непассивному. В рамках непассивного выделялся класс примеров, содержащих недеривированный (то есть не выполнивший условие параметра 1) глагол на -ся. Зачем отдельно выделялись примеры с недеривированным глаголом, станет видно позже.

После разбиения всего массива предложений на классы были выписаны все глаголы, которые входили в предложения из каждого класса, после чего было проведено исследование на наличие совпадений между классами пассивных и непассивных предложений. Так как подобных совпадений было достаточно много, стало ясно, что помимо семантики глагола роль играет другой фактор. В силу того, что многие лингвисты отмечают роль семантических свойств актантов в пассивной конструкции, была проведена дополнительная классификация по свойствам синтаксического и семантического субъекта. Полученный результат и послужил основой для созданной теории.

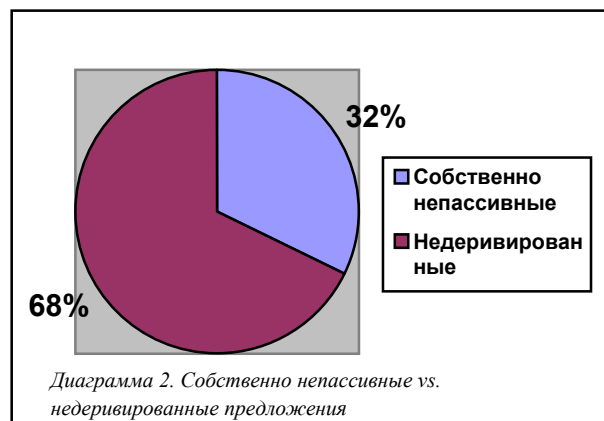
Общая статистика

Пассивных предложений	Непассивных предложений	Всего
514	1192	1706



Разделение непассивных глаголов на собственно непассивные и недеривированные

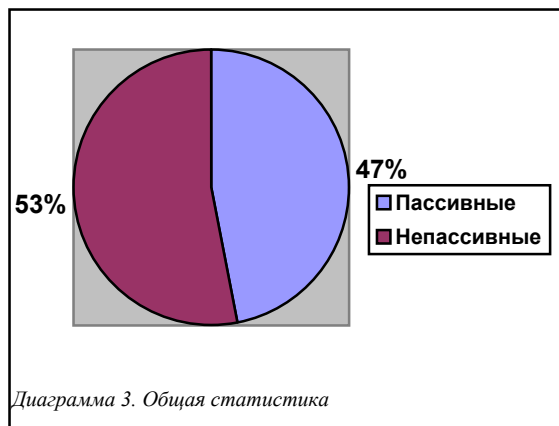
Собственно непассивных предложений	Предложений с недеривированным глаголом	Всего
384	808	1192



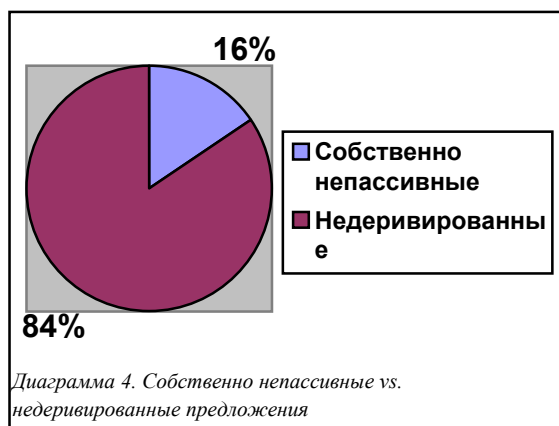
На данном этапе уже можно сделать несколько наблюдений. Во-первых, как многие могли ожидать, непассивных предложений по статистике намного больше, нежели пассивных. Однако при отдельном рассмотрении собственно непассивных конструкций (типа *Я умываюсь*) и конструкций, глагол в которых недеривированный (типа *бояться*), выясняется, что собственно непассивных предложений меньше, чем пассивных, а основной «приток» непассивным предложениям обеспечивают недеривированные конструкции, что, правда, в силу ряда причин (приводить эти рассуждения не представляется возможным в связи с известными ограничениями) не может служить аргументом в пользу первичности пассивного значения -ся.

Данные по различным источникам текстов
а. Публицистика (корпус Тюбингенского университета)

Пассивных предложений	Непассивных предложений	Всего
342	385	727



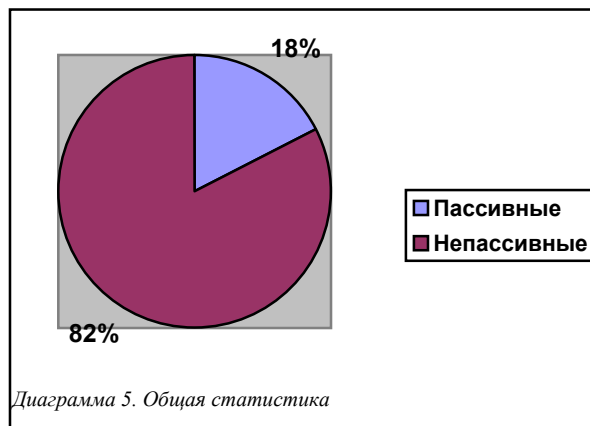
Собственно непассивных предложений	Предложений с недеривированным глаголом	Всего
60	325	385



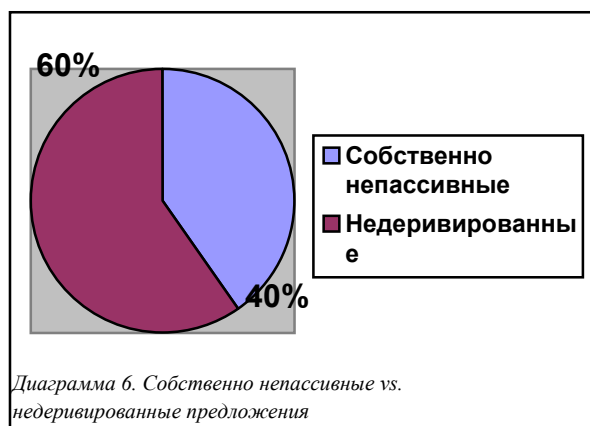
На данных двух диаграммах можно наблюдать, насколько выросла общая доля пассивных предложений с одной стороны, а также доля недеривированных предложений по отношению к собственно непассивным. Первый факт вполне объясним: многие учёные отмечали большую насыщенность пассивами официальных и научных текстов. Более частое употребление недеривированных глаголов можно связать с тем, что те глаголы, которые могут употребляться в собственно непассивных предложениях, встречаются реже в публицистике, нежели в художественном тексте.

б. Художественный текст (Корпус русского языка Яндекса)

Пассивных предложений	Непассивных предложений	Всего
172	807	979



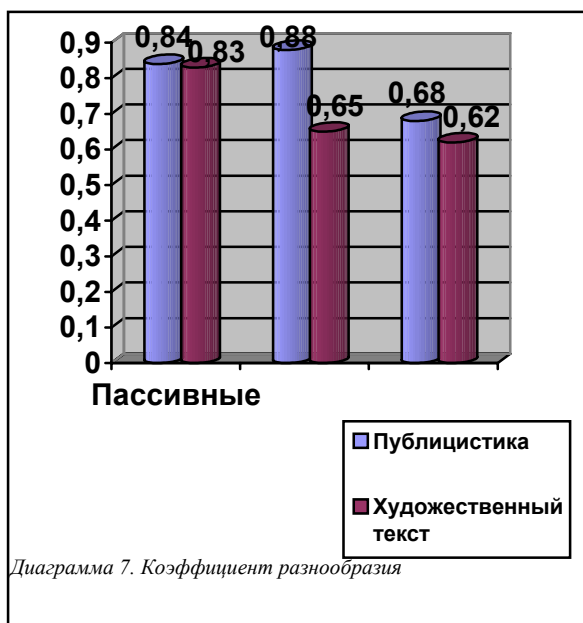
Собственно непассивных предложений	Предложений с не-деривированным глаголом	Всего
324	483	807



Можно видеть, насколько меняется статистика при переходе к художественному тексту — с одной стороны, увеличивается доля непассивных глаголов, с другой стороны, уменьшается доля недеривированных глаголов по отношению к собственно непассивным.

Помимо приведённой статистики, интересными могут быть результаты исследования следующего вопроса: сколько разных глаголов фигурирует в предложениях этих классов? Заметим, что на данные этого исследования искажены относительно реальной картины: сведения нескольких форм глагола в один вход не проводилось в силу того, что эта задача второстепенная, в то же время её должно

выполнение требует несравнимо бóльших затрат времени. На диаграмме даны значения коэффициента разнообразия (количество разных глаголов, делённое на общее количество). Чем он выше, тем сильнее разнообразие глаголов в рамках одного класса. Самое высокое значение (а именно, 1) означает то, что каждая словоформа встретилась ровно один раз.



Удивительно, насколько схожи результаты в двух разных источниках текстов: в двух случаях из трёх различие составляет менее десятой доли.

С помощью этой диаграммы мы можем объяснить такой «скачок» собственно неактивных глаголов в художественной речи. Видно, что их разнообразие существенно ниже, чем в публицистике, несмотря на то, что количество было большим. Большее количество в данном случае было уравновешено меньшим разнообразием.

Хочется также отметить относительно большое разнообразие пассивных глаголов. Становится понятно, что пассивных глаголов не так мало, как мы могли полагать.

3. Анализ свойств пассивных и неактивных предложений

Попытаемся установить некоторые свойства предложений пассивного и собственно неактивного классов (в одной из частей также будут привлечены предложения с недеривированным глаголом). Если то определение пассива, которое мы приняли и обосновали, являлось набором признаков пассива, то в данном случае речь идёт о его свойствах.

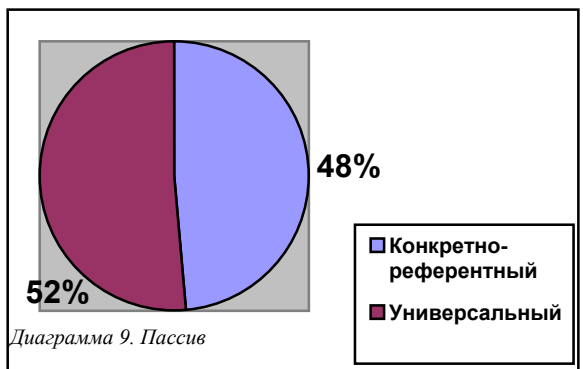
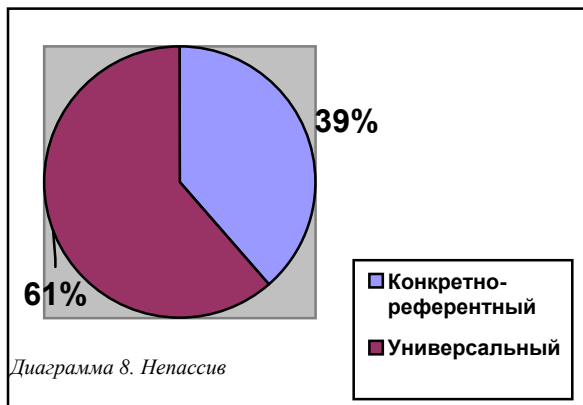
Синтаксический субъект

Синтаксические субъекты предложений корпусов изучались по двум параметрам: референциальный статус и одушевлённость/неодушевлённость.

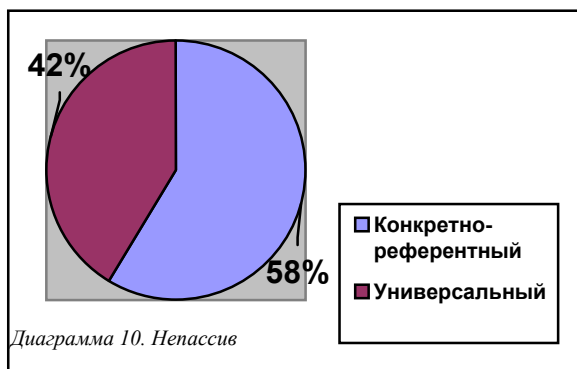
В данном исследовании деление референциальных статусов было достаточно грубым, с возможностью сделать его более подробным, если будут выявлены некие явные тенденции. Для этого же этапа было выбрано деление на два статуса: 'конкретно-референтный' и 'универсальный'.

Далее приведена статистика по данному параметру.

а. Публицистика

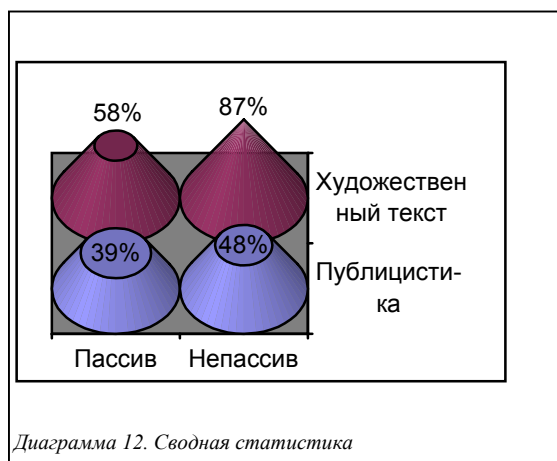


б. Художественный текст





в. Общая диаграмма процента конкретно-референтного синтаксического субъекта

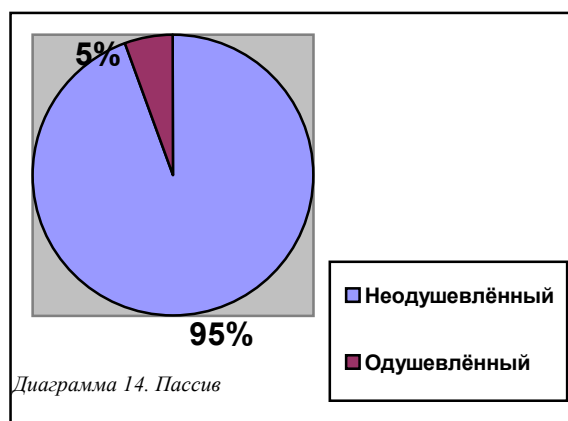
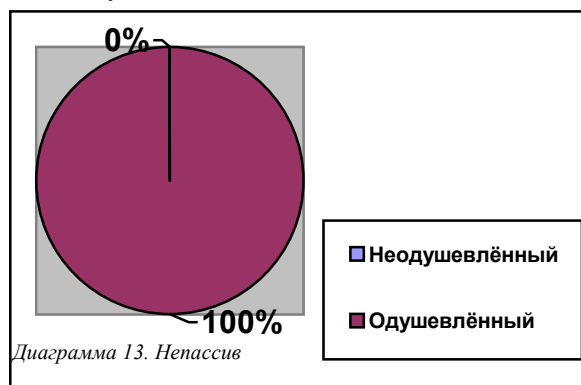


Из данных диаграмм видно, что выбор референциального статуса синтаксического субъекта зависит более от стиля текста, нежели только от класса предложения: вторая пара диаграмм показывает одинаковое увеличение доли конкретно-референтных субъектов по отношению к первой паре диаграмм. Можно, однако, видеть, что в обоих случаях доля конкретно-референтного синтаксического субъекта в пассивном классе всё же больше, чем в непассивном классе, но данный факт можно рассматривать только как тенденцию; нельзя также забывать, что проводить «прямое сравнение» двух диаграмм в каждой паре мы не вправе: в случае первой из диаграмм речь идёт о семантическом субъекте, а в случае второй — о семантическом объекте. Повторю: статистика, касающаяся синтаксического субъекта может быть полезна только в том плане, чтобы послужить предсказательной теории, которая получает на вход предложение с синтаксическим субъектом и, естественно, не знает, является ли он одновременно семантическим субъектом или семантическим объектом.

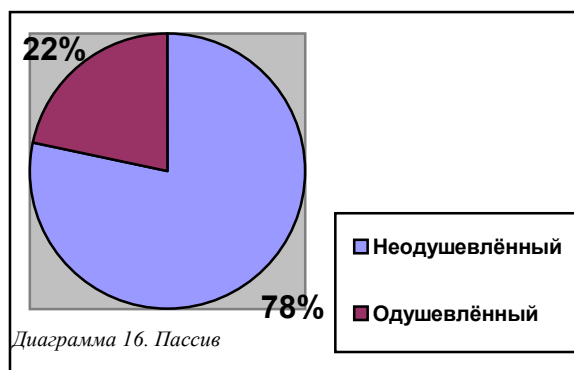
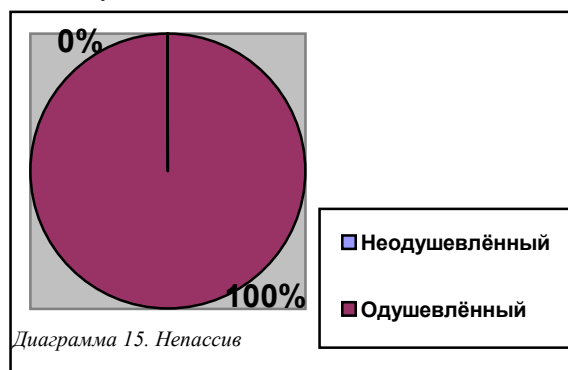
Далее рассматривался параметр одушевлённости и неодушевлённости синтаксического субъекта. То, что был затронут этот признак, легко объяснить: именно одушевлённость или неодушевлённость является самым последовательным (если не единст-

венным) средством выражения способности субъекта быть источником или контролёром действия.

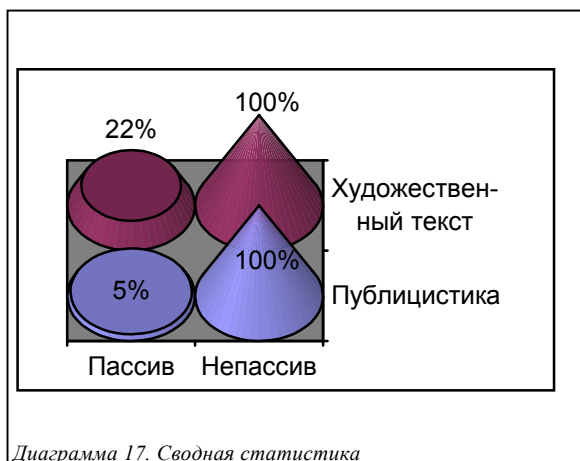
а. Публицистика



б. Художественный текст



в. Общая диаграмма процента одушевлённого синтаксического субъекта



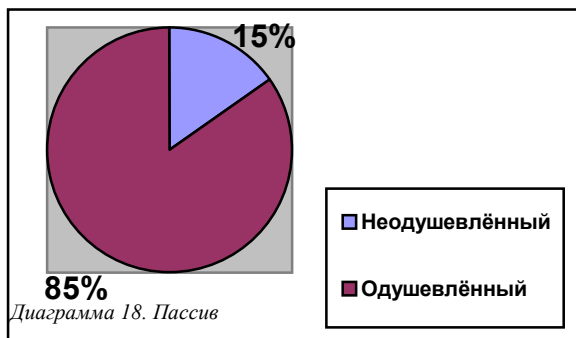
Можно видеть, что в данном случае наше предположение практически полностью подтвердилось: в неактивном классе все синтаксические субъекты — не-animate, а в активном классе в подавляющем большинстве случаев синтаксический субъект, наоборот animate. Остаётся только небольшая часть предложений, которые не укладываются в это правило — а именно, предложения с animate синтаксическим субъектом в активном классе.

Выяснилось, что один большой класс глаголов, образующий такие предложения (пусть и не все, вошедшие в эту группу) — это глаголы эмоции и ментального состояния. Практически во всех случаях глаголы этого класса образуют активные предложения, как это уже было отмечено выше; исключения описаны в заключительном разделе.

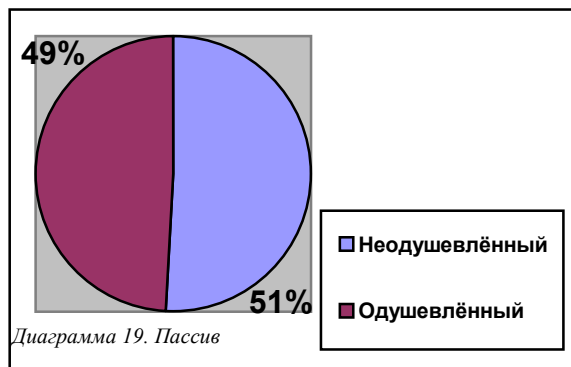
Семантический субъект

Далее исследовались свойства семантического субъекта пассивных предложений (только в пассивных предложениях он не совпадает с синтаксическим субъектом). Исследовалась animate и не-animate семантического субъекта (в силу выраженности семантического субъекта нельзя предоставить точные данные о его референциальном статусе).

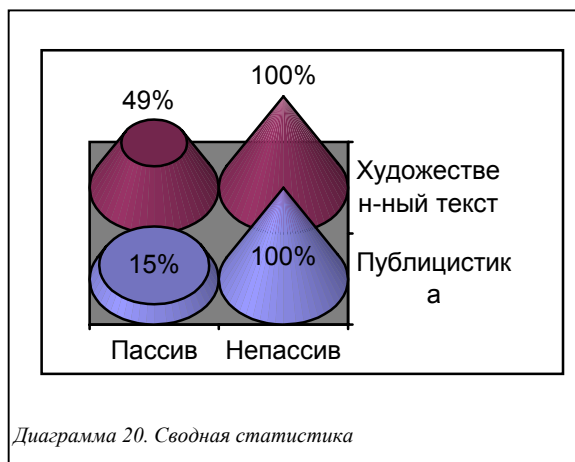
а. Публицистика



б. Художественный текст



в. Общая диаграмма процента одушевлённого семантического субъекта



Из диаграмм видно, что в публицистике animate семантический субъект встречается гораздо чаще, чем в художественном тексте; как видно, сам этот параметр не может служить самостоятельным признаком пассивности или неактивности: доля animate семантического субъекта снова во многом определяется лишь стилем и жанром и сама по себе может служить лишь отражением тенденции.

Но, быть может, более ясную картину мы получим, рассмотрев частоту встречаемости комбинаций описанных признаков.

Комбинации свойств

Далее комбинации признаков обозначаются следующим образом: $x:y:z$, где

x — статус синтаксического субъекта (1 — универсальный, 2 — конкретно-референтный), y — animate синтаксического субъекта (1 — animate, 2 — не-animate), z — animate семантического субъекта (то же обозначение). Всего комбинаций было восемь, и каждая из них оказалась представлена хоть каким-нибудь числом примеров. Приведём примеры каждой из комбинаций, а затем рассмотрим статистику.

Комбинация '1:1:1':

(1) Поскольку понятие «молодой» на сегодняшний день довольно абстрактное, не позволяющее установить какой-либо возрастной ценз, то печататься в альманахе должны молодые, не имеющие сборников и публикаций в толстых журналах.

Комбинация '1:1:2':

(2) Бывало, под предлогом необходимой надобности, прибежишь от урока в ее комнатку, начинаешь мечтать вслух, нисколько не стесняясь ее присутствием.

Комбинация '1:2:1':

(3) Демократизация и гласность настойчиво утверждаются в коллективах, в деятельности партийных и общественных организаций, спадает глянец с якобы однородной действительности.

Комбинация '1:2:2':

(4) Ценность любого телескопа определяется тем, насколько слабый небесный источник он может зафиксировать и насколько детализированной будет «картинка» исследуемого объекта.

Комбинация '2:1:1':

(5) Один раз я на нее рассердился.

Комбинация '2:1:2':

(6) Известная пианистка, художественная натура, тонкая, чувствительная и прикованная к дому, в цепях одиночества, ожесточившаяся и утончившая свое душевное зрение дочь.

Комбинация '2:2:1':

(7) Обед кончился.

Комбинация '2:2:2':

(8) — А где возьмете, скажем, металлы? — И эта проблема разрешима — в лунных породах содержатся титан, и железо, и алюминий. Правда, в связанном виде.

Итак, диаграмма встречаемости комбинации признаков:



Итак, можно выделить наиболее часто встречающиеся комбинации признаков для пассивных предложений: '1:2:1', '2:2:1', '1:2:2' и '2:2:2'. Их объединяет, что объяснимо, неодушевлённость синтаксического субъекта (вторая цифра «2» во всех четырёх случаях). Видно также, что для публицистики особенно «важно», чтобы семантический субъект превосходил синтаксический по статусу одушевлённости. Этим же объясняется меньшая популярность конструкций с одушевлённым синтаксическим субъектом. Интересно, насколько «поведение» художественного текста часто обратно поведению публицистического: можно заметить, что в художественном тексте конструкции с одушевлённым синтаксическим субъектом и неодушевлённым семантическим субъектом даже более популярны, нежели конструкции с одушевлённым семантическим субъектом, то есть даже «поощряется» более низкий в этом плане статус семантического субъекта.

И тем не менее такие предложения пассивны. В этой связи хотелось бы вспомнить точку зрения о том, что допустимость пассива определяется как раз статусом семантического субъекта. Обычно о его референциальном статусе говорить очень сложно, поскольку он чаще всего выражен нулевой лексемой, но когда он выражен ненулевой лексемой, можно наблюдать такую тенденцию: чем ниже статус семантического субъекта, тем выше вероятность, что пассивный пример будет приемлемым:

1) ?Сад охраняется сторожем.

2) *Сад охраняется моим другом Ваней.

Последнее рассуждение о комбинации признаков актантов и может считаться самым ярким и показательным результатом нашего исследования. Представляется, что это направление станет главным для дальнейших исследований.

Но если выбрать параметр для поверхностного распознавания пассивности, то наиболее чётким из параметров оказался признак одушевлённости синтаксического субъекта — для выполнения нашей третьей цели по созданию теории мы воспользуемся именно этим свойством.

4. Параметры определения пассивности высказывания

Далее я приведу параметры, которые могут быть использованы теорией для отнесения предложения к пассивному классу. Параметры теории применяются строго последовательно, от первого к последнему.

Параметр 1: Если смысловой глагол в предложении не является морфологически и семантически образованным от некоего глагола, отличающегося от данного исключительно отсутствием постфикса -ся (в неопределённой форме), то предложение пассивным не является.

Параметр 1': Если исходный глагол непереходен, то пример пассивен.

Параметр 2: Если синтаксический субъект является неодушевлённым, предложение пассивно.

Параметр 3: Если субъект одушевлённый, но исходный для употреблённого в конструкции глагол является глаголом эмоции или ментального состояния (деятельности), то предложение пассивно.

Параметр 4: В остальных случаях предложение не пассивно.

5. Исключения и сложные случаи

1. Контролируемость vs. одушевлённость

В большинстве случаев для данной проблемы контролируемость может быть приравнена к одушевлённости (не как синонимия, а как двухсторонняя импликация), но иногда встречались исключения.

2. «Навязанная контролируемость»

Существуют случаи, когда предложение с одушевлённым субъектом оказывается не пассивным, когда мы ожидаем, что оно (при таком глаголе и таком субъекте) будет пассивным. Это «навязывание контролируемости» может осуществляться как лексически, так и дискурсивно, и даже вследствие отсылки к экстралингвистическим знаниям.

6. Заключение

В заключение хочется сказать, что главная цель исследования ещё не достигнута, и эта работа описывает только начало его. На пути создания объяснительной теории сделано ещё немного, но изучение комбинаций признаков даёт определённые предварительные знания. Сама же предсказательная теория как таковая могла бы иметь только ограниченное применение — в автоматической обработке естественного языка.

Список литературы:

- 1) Haspelmath Martin. The grammaticization of passive morphology // *Studies in Language* vol. 14. Amsterdam – Philadelphia: John Benjamins, 1990. №1, С. 25 – 72.
- 2) Kazenin Konstantin I. The passive voice // *Language Typology and Language Universals*, vol. 2. Berlin – New York: Walter de Gruyter, 1989. С. 899 – 916.