

Автоматическое определение географических объектов в тексте с помощью данных онтологии

Антонов Егор Сергеевич

Аспирант Московского Государственного Университета им. М. В. Ломоносова, Москва, Россия

Одной из актуальных задач прикладной лингвистики является автоматическое определение именованных сущностей (далее NE, от “Named Entity”) в тексте. Под именованными сущностями подразумеваются различные объекты действительности, такие как организации, персоны, даты, географические объекты и т.д.

В рамках данного исследования была поставлена задача автоматического определения всех упоминаний географических объектов (далее гео-объекты) в тексте. В качестве базы данных NE была выбрана онтология [Freebase]. Этот выбор объясняется несколькими параметрами:

- к данной онтологии предоставлен свободный доступ;
- тематика онтологии не ограничена, т.е. имеется возможность экстраполировать результаты для гео-объектов на другие области действительности, не изменяя при этом формата обращения с данными;
- [Freebase] на данный момент содержит более 20 млн. сущностей, что примерно в 7 раз больше, чем в Wikipedia (~3,5 млн.).

Полигоном для проверки алгоритма послужил англоязычный корпус новостей [Reuters].

Концептуально [Freebase] представляет собой базу данных, в которой содержатся сущности, их характеристики и связи между сущностями. Физически это выглядит как таблица следующего вида:

Идентификатор	Свойство	Значение_1	Значение_2
---------------	----------	------------	------------

Подобная структура таблицы позволяет преобразовать объектную структуру, характерную для описания действительности, в плоский вид. Все гео-объекты в [Freebase] принадлежат к типу “/location/location” (всего около миллиона объектов). Исходя из определения [Freebase] гео-объект – это нечто имеющее фиксированное место на планете Земля.

Базовые свойства гео-объектов в [Freebase]:

- геолокация (пара <широта, долгота>);
- положение в иерархии вложенности (например, город Майами содержится в округе Майами-Дейд, в штате Флорида, в стране США и т.д.);
- соседи (например, Московская область граничит с Калужской областью);
- площадь;
- население.

Определение гео-объектов в тексте

Первый шаг определения объекта в тексте – его поиск по названию. Каждый объект может называться несколькими способами (New York, NY, Big Apple). В [Freebase] каждый объект имеет имя и псевдонимы (alias). Помимо этого, некоторые объекты имеют также идентификатор в Wikipedia (например, “New_York”), который можно привести к нормальной строке. Итого получается примерно 2,2 млн. названий на чуть менее 1 млн. гео-объектов. Для поиска такого большого количества подстрок необходим быстрый алгоритм с применением хеширования. Нами были рассмотрены алгоритмы Ахо-Корасик (см. [Aho, Corasick]) и Рабина-Карпа (см. [Karp, Rabin]). Алгоритм Ахо-Корасик при одинаковом времени работы потребовал 2 Гб оперативной памяти, в то время как алгоритм Рабина-Карпа потребовал 800 Мб, поэтому реализован был вариант последнего.

Ранжирование объектов

Реальная встречаемость названий невелика: при обработке всего корпуса (~0,9 млн. новостных сообщений) из всех названий встретилось только около 80 тыс., при этом 60 тыс. из них встретилось менее 10 раз. Такая статистика обусловлена следующими фактами:

- а) большинство гео-объектов встречается очень редко;
- б) многие из объектов имеют одинаковое название (самое «многозначное» название – Midway – соотносится с 215 гео-объектов);

При этом почти каждое слово с большой буквы является одновременно и названием города («The», «June», «March», «A» и т.д.). Т.о. имеется два вида проблем ранжирования объектов:

- 1) Название объекта совпадает с общеупотребительным словом.
- 2) Несколько объектов имеют одинаковое название.

Проблема (1) сводится к приписыванию каждой паре <название, объект> некоторого коэффициента надежности. Проблема (2) имеет более сложную структуру решения: приходится исходить из гипотезы, что объект встречается в контексте связанных с ним данных, а если связанных данных в тексте не присутствует – то нужно выдать в результате объект «по умолчанию», т.е. самый известный.

Такие понятия, как «коэффициент надежности», «связанные данные», «известность объекта», лежат в области нечеткой логики. Физически в базе присутствуют абстрактные и однотипные структуры вида <название; роль_1: множество участников; ... ; роль_n: множество участников>. Весь смысл структуры содержится в названии, которое может что-то сказать человеку, но для компьютера они не несут никаких дополнительных свойств. Поэтому онтологические данные требуют дополнительной концептуализации в соответствии с задачей. Так, для решения проблемы (1) было сформулировано следующее правило:

- каждой паре <название, объект> приписывается начальный коэффициент обратно пропорциональный частоте названия.

Примеры правил для решения проблемы (2):

- каждому объекту приписывается вес, пропорциональный логарифму количества встретившихся объектов из атрибута “/location/location/containedby” (атрибут иерархии вложенности);
- каждому объекту приписывается вес, пропорциональный количеству объектов в кластере геолокации радиусом до 10 градусов;
- каждому объекту приписывается вес пропорциональный населению;
- и т.п.

Подобные правила являются архитектурной прослойкой, которая концептуализирует онтологические данные в контексте разрешения онтологической неоднозначности.

В текущей версии программы коэффициенты правил подобраны вручную, в дальнейшем планируется применить машинное обучение для выявления закономерностей между результатами правил.

Литература

Aho A. V., Corasick M. J. Efficient string matching: An aid to bibliographic search // Communications of the ACM. 1975. №18 (6), Pp. 333–340.

[Freebase]: [http://\[Freebase\].com](http://[Freebase].com)

Karp R. M., Rabin, M. O. Efficient randomized pattern-matching algorithms // IBM Journal of Research and Development. 1987. № 31 (2), Pp. 249—260.

Reuters: <http://trec.nist.gov/data/reuters/reuters.html>

Chak j-lo' *lo'bol*
хотеть 1-есть фрукт
Я хочу поесть фрукты [Haviland: 329]

С точки зрения семантики невозможность охарактеризовать глагол по аспекту, который уже оформлен показателем ирреалиса, вполне обоснована. Аспектуальные значения, согласно классической формулировке Б. Комри, характеризуют различные взгляды на внутреннее развитие ситуации во времени («...aspects are different ways of viewing the internal temporal constituency of a situation» [Comrie: 3]). Ирреалис же предполагает, что ситуация принадлежит сфере воображаемого или гипотетического, и потому обозначает потенциальное или возможное событие, а не наблюдаемый факт действительности («...belongs to the realm of the imagined or hypothetical, and as such it constitutes a potential or possible event but it is not an observable fact of reality» [Elliott: 67]). Таким образом, обозначаемая глаголом ситуация не имеет денотата в реальном мире, так что нет ничего удивительного в том, что говорящий на цоциле не может приписать ей какое-либо аспектуальное значение. Поскольку отсутствует сама ситуация как таковая, то у неё не может быть и внутреннего развития во времени. Тем не менее, далеко не все языки мира идут по тому же пути, что и цоциль: в большинстве случаев аспект и ирреалис или ирреальная модальность могут свободно сочетаться друг с другом в составе единой глагольной словоформы.

Таким образом, в языке цоциль мы наблюдаем типологически довольно редкий факт взаимоисключительности грамматических категорий ирреалиса и аспекта, благодаря которому эти категории могут интерферировать таким образом, что одним из средств выражения ирреальной семантики является отсутствие аспектуального показателя. Семантические зоны аспекта и ирреалиса оказываются тесно взаимосвязаны, что отражается в грамматике языка.

В грамматических глоссах используются следующие сокращения: 1 – первое лицо, 2 – второе лицо, CLASS – классификатор, DEF – определённый артикль, DMN – диминутив, ENCL – энклитика, IPFV – имперфектив, IRR – ирреалис, Q – вопросительная частица, PFV – перфектив.

Литература

Comrie B. Aspect: An introduction to the study of verbal aspect and related problems. Cambridge, 1976.

Cowan M. M. Tzotzil Grammar. México D.F., 1969.

Elliott J. R. Realis and irrealis: Forms and concepts of the grammaticalisation of reality // Linguistic Typology. 2000. № 4.1. P. 55–90.

Haviland J. B. Sk'op Sotz'leb – El Tzotzil de San Lorenzo Zinacantán. México D.F., 1981.

Laughlin R. M. Of Cabbages and Kings: Tales from Zinacantán. Washington D.C., 1977.

Формальные маркеры функционально-стилевой принадлежности русскоязычных текстов

Головко Николай Вячеславович

Аспирант Ставропольского государственного университета, Ставрополь, Россия

Обращение к изысканию формальных маркеров функционально-стилевой принадлежности текста необходимо в целях построения лингвистического обеспечения оптимизированного алгоритма детектирования функционального стиля (ФС), разработку которого мы осуществляем в рамках реализации ФЦП «Научные и научно-педагогические кадры инновационной России» на 2009-2013 годы (ГК НК-523П). Определение ФС не является принципиально новой задачей; известен ряд исследований соответствующей тематики, предлагающих разнообразные алгоритмические решения для реализации детектирующего механизма – в частности, отметим работы П.И.

Браславского [Браславский: 1] , который предлагает оценивать такие параметры текста, как средняя длина слова, доля именных частей речи, глагольных форм, личных местоимений; иными словами, в рамках концепции предлагается традиционный подход – исчисление тех параметров текста, которые дают представление о его форме, но не о его содержании. Нам представляется, что данные воззрения, равно как и подобные им, не позволяют в полной мере реализовать комплексный подход к анализу текстов, что заставляет нас осуществлять поиск двух типов параметров: материальные (выразительные) и информационные (содержательные) маркеры. При этом оба типа маркеров должны являться формальными, т.е. имеющими внешнее выражение, поддающимися исчислению и формализации, доступными для машинного исследования.

Осуществляя отбор маркеров, необходимо учитывать сферы наиболее критического различия между строго формальными книжными стилями (научным и официально-деловым) и свободными книжными стилями (художественным и публицистическим). Если суммировать типические черты различных ФС и их групп, то становится очевидным, что на основании упомянутых черт могут быть выделены два базовых параметра.

С одной стороны, это количественный показатель, описывающий количество звуков и, соответственно, букв в употребляемых лексических единицах, т.е. средняя длина слова. Исходя из типических характеристик ФС, мы находим возможным выдвинуть предположение о том, что средняя длина слова может являться единственным количественным параметром текста, достаточным для определения его функционально-стилевой принадлежности.

С другой стороны, это качественный показатель, описывающий степень полисемантической текста. Отметим, что нами не обнаружено опыта исследования каких-либо формальных параметров текстов, которые позволяли бы составить представление о плане содержания текста и, соответственно, судить о потенциальной степени его полисемантической; в то же время нам представляется, что подобный показатель может представлять определенную ценность для процесса анализа. Исходя из типических черт тех или иных ФС, в особенности – характеристик, связанных с типом предпочитаемых слов, мы считаем допустимым утверждать, что существует возможная взаимосвязь между длиной слова, фактом его заимствованности / незаимствованности из других языков и количеством его потенциальных значений. Действительно, по мере увеличения длины слова возрастает вероятность того, что оно является заимствованным и обладает либо единственным денотативным значением, либо одним денотативным и одним коннотативным значением, т.е. представляет собой либо однозначную, либо слабо полисемантизованную лексическую единицу; справедливо также обратное утверждение [Головки: 75]. Тогда отделение свободных книжных стилей от формальных может быть осуществлено посредством выявления и исследования двух взаимосвязанных маркеров: средней длины слова и степени потенциальной полисемантической текста.

Таким образом, в наши задачи входит изыскание внешних признаков потенциальной полисемантической текста. Выяснив на основании теоретических рассуждений и практических исследований, что низкая степень полисемантической текста может свидетельствовать о высокой доле в нем заимствованных слов, в том числе терминологического характера, и наоборот, в процессе решения указанной задачи мы можем свести его к изысканию внешних признаков заимствованных и незаимствованных слов соответственно, соотнося получаемые данные с вопросами потенциальной полисемантической текста.

Так как нами уже была доказана возможность взаимосвязи между графическим обликом слова и фактом его заимствованности, а также со степенью полисемантической, мы не будем повторно обосновывать данную мысль, однако сформулируем общее

предположение в следующем виде: ФС, к которому принадлежит текст, может быть определен посредством оценки и интерпретации данных о массовой доле в указанном тексте слов, принадлежащих к определенным словарным группам. Данное предположение базируется на умозаключении о том, что, коль скоро те или иные звукосочетания, выраженные определенными графическими элементами, могут быть свойственны или не свойственны исследуемому языку, их наличие или отсутствие может являться косвенным свидетельством заимствованности или незаимствованности того или иного слова, а, следовательно, и степени его полисемантической. Потенциальная же полисемантическая мера неопределенности его толкования экстраполируется и на текст в целом, поскольку мера неопределенности системы (т.е. энтропия) по определению обладает свойством аддитивности.

Подобное исследование самым естественным образом позволит составлять представление о степени потенциальной полисемантической текста, т.е. о плане его содержания, посредством формального анализа – изучения плана выражения, а, следовательно, и определять его функционально-стилевую отнесенность на основании этого аналитического критерия.

Литература

Браславский П.И. Пути повышения эффективности поиска научной информации в Internet. http://www.nsc.ru/ws/show_abstract.dhtml?ru+2+44.

Головкин Н.В. Логико-квантитативный аспект теории фиксирования типов языковой информации // Вестник Ставропольского государственного университета. 2008. № 3. С. 72-79.

Некоторые способы маркирования перехода хода в русском жестовом языке

Кебец Павел Леонидович

*Студент Московского государственного университета имени М. В. Ломоносова,
Москва, Россия*

В данном докладе будет проведена попытка дать начальную характеристику стратегии многоканальных поведений, которыми характеризуются определённые стадии диалога, на материале русского жестового языка (далее РЖЯ).

Исследование маркеров перехода хода проводилось на материале американского жестового языка (American Sign Language, ASL) (см. [Baker 1977]). Данная работа представляет собой попытку провести подобное исследование на материале РЖЯ.

В [Sacks et al 1974] механизм смены коммуникативных ролей называется *turn-taking*. В данном докладе для обозначения этого явления используется термин *переход хода*. Собственно *ход* — это реплика одного из участников речевого акта, непрерываемая другим.

Далее будет показано, какие маркеры можно выделить в дискурсе РЖЯ и какое они занимают место в системе перехода хода.

1. Устройство диалога

С точки зрения взаимодействия адресанта и адресата можно выделить следующие стадии диалога:

- инициация (initiation) — как адресант начинает переход хода и как адресат вызывает отклик у адресанта;
- продолжение (continuation) — как адресант сообщает, что продолжит речь после небольшой паузы, и как адресат показывает, что слушает и что адресант может продолжать;
- переключение (shift) — как адресант сигнализирует об окончании высказывания и как адресат высказывает желание переменить свою роль.

2. Маркеры перехода хода

Маркерами в данном случае называются те кванты поведения артикуляторных органов, которые сами по себе не имеют лексического или грамматического значения,

но помогают регулировать отношения между локуторами. Этот набор не является строго регламентированным, однако некоторые маркеры встречаются достаточно регулярно.

В результате исследования в качестве маркеров перехода хода были выделены особые движения рук и направление взгляда.

2.1 Движение рук

Несмотря на то, что руки являются основным речепорождающим органом у носителей ЖЯ, некоторые движения рук могут также выполнять функции маркеров. В их числе движение рук относительно позиции покоя.

2.1.1. Позиция покоя

Периоды активности рук сменяются периодами покоя. Однако, как отмечается в [Baker 1977], существует три различные позиции покоя:

- *позиция полного покоя* (full-rest) — максимально расслабленное положение, при котором адресат внимателен к говорящему и не готовится физически к порождению жеста;
- *позиция полупокоя* (half-rest) — руки адресата находятся в районе пояса, а ладони повернуты к собственному телу;
- *позиция четвертьпокоя* (quarter-rest) — одна из рук адресата (руки всё ещё в состоянии покоя), чаще доминантная, находится выше пояса, недалеко от нейтральной позиции.

Перемещение руки из одной позиции в другую может свидетельствовать о намерении одного из локуторов сменить свою коммуникативную роль.

2.2 Взгляд

Взгляд как маркер перехода хода охватывает широкий спектр функций. Например, он может выступать в качестве просодического оформления высказывания.

Если для утвердительных высказываний и специальных вопросов более характерен –ВЗГЛЯД (случаи, когда взгляд локутора в момент общения отводится в сторону) в начале высказывания, то для общих вопросов более типичен +ВЗГЛЯД (случаи, когда взгляд локутора обращён к собеседнику) (см. [Baker 1977]).

3. Место маркеров в диалоге

В [Baker 1977] приводится подробное сопоставление конкретных маркеров и стадий диалога для ASL. В представленной далее схеме показана их корреляция в РЖЯ.

1. Инициация:
 - a. Адресант**
 - i. «поднятие» рук из позиции покоя
 - ii. помахивание рукой перед адресатом с целью привлечения внимания
 - iii. –ВЗГЛЯД с началом иллокутивного высказывания
 - iv. +ВЗГЛЯД при общем вопросе
 - b. Адресат**
 - i. +ВЗГЛЯД
 - ii. сохранение неактивности
2. Продолжение
 - a. Адресант**
 - i. –ВЗГЛЯД
 - ii. невозвращение в позицию покоя:
 1. заполнение паузы движением, показывающим размышление
 2. удержание последнего выполненного жеста на протяжении паузы
 - b. Адресат**
 - i. сохранение +ВЗГЛЯД
 - ii. краткое повторение некоторых жестов адресанта
3. Переключение
 - a. Адресант**
 - i. +ВЗГЛЯД

- ii. попытка вызвать отклик у адресата:
 1. открытая ладонь и опускание рук(и)
 2. удержание последнего жеста (как вопрос)
- iii. возвращение в позицию покоя

в. Адресат

- i. движение из состояния покоя с целью обратить внимание адресанта на себя
- ii. –ВЗГЛЯД в тот момент, когда адресант +ВЗГЛЯД
- iii. перебивание и повторение нескольких первых.

Выводы

По итогам работы был сделан вывод, что маркеры перехода хода присутствуют на всех стадиях диалога. Важно отметить, что похожая система маркеров перехода хода характерна не только для РЖЯ, но также и для американского жестового языка (см. [Baker 1977]). Однако между ними есть и отличия, как, например, +ВЗГЛЯД во время специального вопроса в РЖЯ, или использование адресатом открытой ладони как маркера переключения в ASL.

Литература

Baker C. (1977). Regulators and Turn-taking in American Sign Language Discourse. In L. A. Friedman. Ed. On the Other Hand: New Perspectives on American Sign Language. New York: Academic Press, 215-236

Sacks H., Schegloff E., Jefferson G. (1974). Language, Vol. 50, No. 4, Part 1, 696-735

Графическая типология знаков кохау ронгоронго: к постановке проблемы.

Коровина Евгения Владимировна

Студентка Российского государственного гуманитарного университета, Москва, Россия

Кохау ронгоронго (письму острова Пасхи) в настоящее время посвящено большое количество работ. Однако вопрос о сути письма и о возможности его прочтения по-прежнему стоит весьма остро. В виду отсутствия каталога, адекватно отражающего особенности письма, любые попытки анализировать письмо с помощью статистических инструментов натывается на почти непреодолимые трудности.

Причины отсутствия каталога разнообразны. С одной стороны это обусловлено тем, что это письмо богато лигатурными написаниями, но из-за малого объема корпуса во многих случаях трудно показать, что перед нами именно лигатура, а не сложный знак. Аналогичное имеет место и в иероглифическом майя: существуют как сложные знаки, состоящие из более, чем одного элемента, например, ОТОТ 'дом' состоит из двух элементов, но каждый из них не составляет отдельного знака, так и примеры собственно лигатурного написания, например, логограмма СНУК 'захватывать', составленная из двух слоговых знаков *chu* и *ku* [Давлетшин: 104]. Другой трудностью является то, что в письме острова Пасхи имеется, по-видимому, довольно большая вариативность написаний, т.е. у одного знака имеется часто более чем один вариант написания, а кроме того одному и тому же варианту написания может соответствовать более чем один знак и различение таких случаев очень затруднительно.

Однако кажется возможным даже при нынешнем состоянии каталога классифицировать знаки по месту их употребления в лигатуре. Частично такая классификация была предложена в статье Позднякова [Поздняков: 168], но здесь хотелось бы показать некоторые аспекты взаимодействия между этими классами. Предполагается, что знаки по этому признаку можно разделить на начальные, конечные, изолированные и неопределенные. Однако такое разделение имеет смысл только для высокочастотных знаков, в данном случае для знаков, употребленных больше 100 раз.

При описании результатов я придерживаюсь нумерации, имеющейся в каталоге Бартеля [Barthel: 348], как наиболее общепринятой.

Начальным знаком мы называем такой знак, который более чем в половине случае употреблен в лигатуре слева, т.е. в ее начале; соответственно, конечным называется знак, который в большей части случаев употреблен в лигатуре справа. Изолированным же называется такой, который употребляется по большей части вне лигатур.

В результате изолированными (ИЗО) получились следующие знаки: 280, 670, 40, 2, 70, 7, 700, 600, 22, 50, 11, 5, 48, 430, 8, 20, 90, 4, 1, 71. Начальными (НАЧ) знаками являются 380, 300, 200. И наконец, к конечным (КОН) знакам относятся 76, 3, 64, 74, 10, 62, 63, 6, 61. Остальные (ПРО) частотные знаки не показывают существенного предпочтения какой-нибудь одной позиции.

Если после этого посмотреть, как ведут себя эти четыре группы знаков в отношении сочетаемости между собой, можно увидеть следующее. Сочетаемость сильно выше случайной будет в парах НАЧ-КОН, КОН-ИЗО, ПРО-НАЧ, а в парах НАЧ-НАЧ, КОН-КОН, НАЧ-ИЗО сочетаемость будет сильно ниже случайной. Этот результат является довольно тривиальным, поскольку все эти классы были изначально определены по положению в лигатуре. Однако следует заметить, что редкая сочетаемость между собой внутри двух из четырех групп знаков, вероятно, несет некоторую информацию не только о графической структуре письма, но и о собственно языковых свойствах текста, стоящего за этим письмом.

Литература

Barthel T.S. Grundlagen zur Entzifferung der Osterinselschrift. Hamburg, 1958

Давлетшин А.И. Палеография древних майя : Дис. ... канд. ист. наук : 07.00.09 М., 2003

Поздняков К.И., Поздняков И.К. Рапануйская письменность и рапануйский язык: предварительные результаты статистического анализа // Антропологический форум. 2005. № 3. С. 167–204.

Дифференцирующие характеристики дейктических элементов при контекстуальном дейксисе в англоязычной научной коммуникации

Коровина Ирина Валерьевна

Сотрудник Мордовского государственного университета им. Н.П. Огарёва, Саранск,
Россия

Дейксисом традиционно считается использование языковых выражений и других знаков, которые могут быть проинтерпретированы лишь при помощи обращения к физическим координатам коммуникативного акта – его участникам, месту и времени [Кибрик: 87]. В речи дейксис реализуется через дейктические элементы (местоимения (в частности, личные и указательные) (*he, they, those*), некоторые наречия места и времени (*today, there, on the left*) и т.д.). Дейктики, как и любые другие элементы, в речи реализуют **референцию** – связь языкового знака и объекта/явления действительности (референта) [Падучева: 7]. Однако, до сих пор нет моделей, иллюстрирующих то, как они осуществляют референцию в речи.

Общепризнанными характеристиками дейктических элементов считаются:

- **ситуативность** (зависимость от конкретной коммуникативной ситуации)
- **субъективность** (выделение объекта по случайному признаку соотнесенности с говорящим)
- **эгоцентризм** (постоянная отнесенность к говорящему)
- **мгновенность и эфемерность значения** (изменение значения дейктиков от одного высказывания к другому) [Кацнельсон: 11].

Необходимо сказать, что полнозначные слова также обладают ситуативностью в речи. Кроме того, эгоцентризм и субъективность в определенной степени могут быть также приписаны полнозначным словам, поскольку говорящий всегда является продуцентом своей речи, и соответственно каждый речевой акт относительно эгоцентричен и субъективен. Мгновенность значения дейктиков имеет место быть, но

она не является их дифференцирующей характеристикой. Целесообразно рассмотреть дейксис как процедуру и на основе параметров этой процедуры выявить дифференцирующие признаки дейктических элементов, которые бы отличали их от полнозначных слов. В качестве материала исследования предлагаем избрать англоязычные научные тексты лингвистической тематики, поскольку научные тексты обладают четкими информационными координатами, которые позволяют более детально рассмотреть систему дейктических координат, в рамках которых функционируют все дейктические элементы.

Анализ показал, что чаще всего в научной коммуникации встречается **контекстуальный дейксис**, который представляет собой дейктическое указание, осуществляющееся через антецедент (вербально выраженный элемент предшествующего/последующего контекста) [Сребрянская: 23]. На примере контекстуального дейксиса дейктическая процедура просматривается более четко, чем в случае неконтекстуального дейксиса (без антецедента).

(1) *In this subsection instances of person and social deixis will be analyzed in a way that shows that **these** two categories of deixis cannot be adequately discussed without reference to **each other**. (these и each other – дейктики; person and social deixis – антецедент)*

Дейктическое указание в речи осуществляется всегда в рамках системы дейктических координат, которая в научных текстах состоит из следующих элементов:

- дейктический центр (в случае контекстуального дейксиса это точка текстуального пространства, содержащая дейктик),
- вектор дейктического указания, исходящий из дейктического центра, и
- дейктический антецедент.

Успешная референция, осуществляемая дейктиками в речи, всегда зависит от хотя бы одного из указанных параметров системы дейктических координат. Подобная зависимость не наблюдается у полнозначных и служебных слов. Пример (1) иллюстрирует зависимость дейктической референции от антецедента.

(2) *Let us stay for a moment with the opposition of form and substance which has made itself apparent **here**.* (зависимость от дейктического центра)

(3) *There are additional senses that apply to syntactic forms that derive from deictics by means discussed **earlier** in this chapter.* (зависимость от дейктического центра и вектора дейктического указания (анафора))

Для полноты картины необходимо обратиться к семантической структуре дейктических элементов. Дейктики, в отличие от полнозначных слов, не обладают денотативным значением; им присуще **грамматическое значение**, которое выражено как в категориях числа и лица (*they – he/she/it, he – she, this – these*), так и в их указательности. Имеется также другая разновидность дейктиков, имеющих определенное денотативное значение, которое однако характеризуется наличием так называемой семантической «лакуны», заполнение которой возможно только в речи/контексте (*to the right, yesterday, here*). Можно заключить, что дейктики обладают частичной семантикой.

Таким образом, процедуру дейктического указания в случае контекстуального дейксиса можно представить следующим образом: языковой знак – частичная семантика – антецедент – референт. При этом, успешная референция дейктиками осуществляется в рамках системы дейктических координат. Учитывая параметры процедуры дейктического указания можно вывести дифференцирующие характеристики дейктиков в случае контекстуального дейксиса:

1) наличие у дейктиков частичной семантики, выраженной либо грамматическим значением (не денотативным) (*he, she, this*) или денотативным значением с определенной семантической «лакуной», которая заполняется в контексте (*to the left, over there*).

2) зависимость успешной дейктической референции от хотя бы одного из параметров системы дейктических координат (дейктического центра, вектора дейктического указания или антецедента).

Приведенные характеристики дейктических элементов являются дифференцирующими для данной группы слов и оказывают существенную помощь в определении круга дейктических элементов при контекстуальном дейксисе.

Литература

Кибрик А.А. Человеческий фактор в языке: Коммуникация, модальность, дейксис. М., 1992.

Падучева Е.В. Высказывание и его соотносённость с действительностью. М.:, 1985.

Кацнельсон С.Д. Общее и типологическое языкознание. Л, 1986.

Сребрянская Н.А. Дейксис и его проекции в художественном тексте: монография. Воронеж, 2005.

Морфемная структура оценочных слов в английском языке (корпусное исследование)

Куликов Сергей Юрьевич

Студент Московского государственного областного университета, Москва, Россия

Для задач автоматического извлечения мнений необходим комплексный анализ оценочной структуры входного текста. В данной работе мы рассмотрим вопросы, связанные с оценкой слова. Важность точного определения оценки слова обусловлена тем фактом, что именно слово является отправной точкой анализа в системах автоматического извлечения мнений.

В современной компьютерной лингвистике зачастую не делается различия между словообразованием и словоизменением [Марчук: 61], но для точной интерпретации оценки в системах автоматического извлечения мнений выделение деривационных элементов необходимо. Во-первых, это приведет к экономии места в словаре системы, что увеличит ее быстродействие, а во-вторых, повысит качество работы системы.

Данное исследование было проведено на материале англоязычного корпуса *polarity dataset v2.0*. В ходе исследования были выделены 214 различных оценочных коллокаций со словом *film*. Выбор коллокаций со словом *film* обусловлен как тематикой корпуса (пользовательские отзывы о фильмах), так и частотой данного слова в корпусе (слово *film* является самым частотным знаменательным словом). Общее число коллокаций – 473. В ходе исследования все оценочные слова были разделены на следующие группы: однокорневые слова, сравнительные (в т.ч. супплетивные) формы, суффиксальные слова, префиксально-суффиксальные слова, сложные слова. В группе сложных слов выделение некорневых морфем не проводилось.

Однокорневые слова составили 60 различных коллокаций, что составляет 28% от общего числа различных коллокаций. А в абсолютном отношении (по числу словоформ) – 42%. Сравнительные формы составили 11 различных коллокаций, что составляет 5% от их общего числа. В то же время в абсолютном отношении они составляют – 12%. Сложные слова вошли в состав 7 коллокаций, что составляет 3% от общего числа. В абсолютном отношении их доля составляет только 1,7%. Суффиксальные слова образуют 129 коллокаций, т.е. 60%. А в абсолютном отношении – 42%. На долю префиксально-суффиксальных слов приходится 7 коллокаций, т.е. 3%. В абсолютном отношении – 2,3%.

Из приведенных выше данных видно, что наибольшее число коллокаций образованы по модели суффиксальной деривации. Доля сложных и префиксально-суффиксальных слов мала как в количественном, так и в качественном отношениях. Высокая доля однокорневых слов и сравнительных форм (суммарно более 50% от всех словоформ) свидетельствует о значительной роли словаря системы в качественном анализе данных. Анализ 5 самых частотных коллокаций показывает, что они входят

именно в данные две группы. При этом из сравнительных форм наиболее высокочастотными являются супплетивные формы, которые целесообразнее заносить в словарь, чем выводить правила их образования.

Значимость именно морфемного, а не морфологического, анализа заключается в различной семантической нагрузке морфем. Так один и тот же корень с различными суффиксами имеет разную как оценочную полярность, так и её степень. Так, например, английский корень *-worth-* при присоединении суффиксов *-у* и *-less* имеет противоположные значения. Сходное явление наблюдается и с другими сочетаниями корней и аффиксов (*usual* и *unusual*, *hearted* и *heartless*). В то же время присоединение различных корней может вести лишь к иной степени одинаковой оценки (например, *heartbreaking* и *heartwarming*). В случае префиксально-суффиксальной деривации на материале корпуса можно сделать вывод о том, что префиксальная часть слов, образованных по данной модели, отвечает за полярность оценки (префиксы *in-* и *de-* соответствуют отрицательной оценке всего слова, а префикс *en-/in-* – положительной), в то время как суффиксальная часть определяет степень полярности.

Таким образом, семантическая структура префиксов и суффиксов может описываться с помощью лексических функций, которые позволяют выявить тип оценки (лексические функции семейств *Воп* и *Мagn*). «Регулярные средства выражения *Воп* и *AntiВоп* в русском языке — префиксы и основы *без-, не-, благо-, добро-, зло-, перво-*» [Апресян: 49]. Для типологически несходного английского языка указанные лексические функции ведут себя несколько по-другому. Как показало наше исследование, проявление значений лексических функций внутри слов для английского языка зависит от морфемной структуры слова.

В случае только суффиксальной модели словообразования реализация возможна только у функций *Воп* и *AntiВоп*. Лексическая функция *Мagn* не актуализируется. При префиксально-суффиксальной деривации происходит дифференциация между различными аффиксами. У префиксов проявляются семантические свойства, присущие лексической функции *Воп* (или *AntiВоп*). У суффиксов – семантические свойства лексической функции *Мagn* (или *AntiМagn*). Важно учесть, что в некоторых случаях (например, в вышеописанном случае с *heartbreaking* и *heartwarming*) при общей словообразовательной модели (корень 1 “*-heart-*” + корень 2 + формообразующий суффикс “*-ing*”) за реализацию функции *Мagn* отвечает суффикс “*-ing*”, в то время как лексическое значение второго корня отвечает за проявление несамостоятельных лексических функций *Minus* и *Plus*.

В данном исследовании мы выявили словообразовательные модели, по которым образуются оценочные слова в английском языке. В ходе анализа структуры оценочных слов были выявлены типы семантических явлений. Для описания для этих типов семантических отношений подходит формализм лексических функций. Внедрение компонента морфемного анализа в системы автоматического извлечения мнений приведет к улучшению работы системы в целом за счет привлечения большего объема семантической информации.

Литература

Апресян Ю.Д. Избранные труды. Т.1. М., 1995.

Марчук Ю.Н. Компьютерная лингвистика. М., 2007.

polarity dataset v2.0 http://www.cs.cornell.edu/people/pabo/movie-review-data/review_polarity.tar.gz.

Исследование восприятия глаголов разных семантических классов методом функциональной магнитно-резонансной томографии

Малютина Светлана Александровна

Студентка Московского государственного университета имени М.В. Ломоносова,
Москва, Россия

Лауринавичюте Анна Кестучё
Аспирантка Московского государственного университета имени М.В. Ломоносова,
Москва, Россия

Работа посвящена определению нейрофизиологических коррелятов понимания глаголов двух семантических классов: физического действия и инструментальных. Глаголы физического действия обозначают действия, выполняемые какой-либо частью тела (например, *рвать*, *щекотать*). Инструментальные глаголы обозначают действия, для выполнения которых необходимо использование инструмента, не являющегося частью тела (например, *пилить*, *вязать*); являются подклассом класса глаголов физического действия. Представляет интерес то, связана ли обработка глаголов данных семантических классов не только с областями мозга, ответственными за языковую обработку, но ещё и с теми областями мозга, что и собственно выполнение соответствующего физического действия (или физического действия, выполняемого с помощью инструмента) или наблюдение за ним.

Теория воплощенного сознания (англ. Embodied Cognition Framework, [Pulvermüller 2005]) утверждает, что в обработку языковых стимулов будут вовлечены области мозга, активирующиеся при совершении физического действия / наблюдении за ним. Этой теории противопоставлен символический подход (англ. symbolic account), утверждающий, что значение слова не связано с сенсомоторными мысленными образами, а выводится исключительно из синтаксических отношений между абстрактными языковыми символами [Landauer & Dumas 1997] или статистического совместного появления лингвистических форм [Kintsch 2008]. Согласно символическому подходу, в обработку языковых стимулов будут вовлечены только области мозга, ответственные за языковую обработку, и не будут вовлечены области, связанные с выполнением соответствующего действия / наблюдением за ним.

Чтобы получить данные в пользу какой-либо из этих гипотез, был проведён эксперимент с применением метода функциональной магнитно-резонансной томографии. Это метод позволяет установить, с активностью каких регионов головного мозга связано выполнение различных когнитивных, в том числе языковых, заданий; основной принцип заключается в том, что испытуемый выполняет задания и в это время измеряются локальные изменения в кровоснабжении его головного мозга. Цель проведенного эксперимента – определить с применением этого метода, какие области мозга вовлечены в обработку глаголов, содержащих семантические компоненты физического действия и инструментальности.

Экспериментальный материал состоял из трёх групп глаголов: инструментальных глаголов физического действия (типа *пилить*), неинструментальных глаголов физического действия (типа *щекотать*) и абстрактных (обозначающих действие, напрямую не связанное с использованием частей тела или инструментов, типа *скупать*, *одалживать*). К каждому глаголу предлагался подходящий и неподходящий к глаголу объект (например, к *вязать* – *свитер* и *крышу*), из которых испытуемому давалось задание выбрать подходящий (например, для *вязать* это *свитер*). В качестве контрольного условия использовалось задание с последовательностями небуквенных знаков: испытуемому предлагалось выбрать, какая из двух последовательностей идентична третьей.

Эксперимент был разработан на материале немецкого языка и проведён на базе клиники Гросхадерн (Мюнхен, Германия). В эксперименте приняли участие 18 здоровых испытуемых. При сравнении глаголов всех трёх групп и контрольного условия (знаки) была обнаружена активация в лобных и задневисочных отделах, то есть классических речевых зонах, что подтверждает валидность эксперимента. При сравнении глаголов физического действия и абстрактных глаголов была обнаружена бо́льшая активация для абстрактных глаголов в лобных и задневисочных речевых зонах, передней височной области. Областей, более активированных для глаголов физического действия,

обнаружено не было. При сравнении инструментальных и неинструментальных глаголов была обнаружена бо́льшая активация для неинструментальных глаголов в лобной, нижней височной области, верхней теменной области. Областей, более активированных для инструментальных глаголов, обнаружено не было.

Таким образом, результаты данного эксперимента свидетельствуют в пользу символического подхода: разница при обработке глаголов разных семантических классов затронула только области мозга, традиционно связываемые с языковой обработкой. Гипотеза теории воплощённого сознания о том, что в обработку глаголов будут вовлечены в том числе те же зоны, что и выполнение соответствующего действия / наблюдение за ним, на данном материале не подтвердилась.

По количеству активного мозгового субстрата можно выстроить следующую иерархию (в порядке возрастания количества активации): инструментальные глаголы физического действия < неинструментальные глаголы физического действия < абстрактные глаголы. По материалам опроса носителей немецкого языка, ту же иерархию глаголы образуют и в порядке уменьшения их представимости (англ. *imageability*; параметр характеризует, насколько легко представить себе действие, обозначаемое глаголом). Таким образом, обработка данных семантических классов глаголов отличается по количеству активированного субстрата в речевых областях, и чем менее представимы глаголы данного класса, тем больше активация.

Авторы благодарят О.В. Драгой и М.В. Иванову, под руководством которых выполнено данное исследование.

Литература

- Pulvermüller F.* Brain mechanisms linking language and action // *Nature Reviews Neuroscience*. 2005. № 6. С. 576–582.
- Kintsch W.* Symbol systems and perceptual representations // De Vega M., Glenberg A., Graesser A. *Symbols and Embodiment*. Oxford, 2008. С. 145–164.
- Landauer T.K., Dumas S.T.* A solution to Plato's problem: the latent semantic analysis theory of acquisition, induction, and representation of knowledge // *Psychological Review*. 1997. № 104. С. 211–240.

Обыденное толкование как проявление агнонимичности слов

Маметьева Мария Андреевна

Студентка Кемеровского государственного университета, Кемерово, Россия

Современная научная парадигма характеризуется повышенным вниманием к обыденной семасиологии как направлению, выявляющему особенности восприятия и толкования слов рядовым носителем языка. Интерпретация значений слов, разные способы их осмысления приводят исследователей к выделению стратегий толкования значений слов, выявляемых в разных социально-языковых сферах, научному описанию такой составляющей речемыслительной деятельности, как понимание. Подобный подход обусловлен прежде всего ориентированностью лингвистических исследований на человека говорящего: толкование, как и понимание значений слов рядовыми носителями языка, антропоцентрично по своей природе. Особенности восприятия, понимания и толкования слов характеризуют языковую способность человека, раскрывают его индивидуальный внутренний лексикон.

Исследование значений слов в аспекте их понимания приводит к выделению такого понятия, как агнонимия. Термин агнонимы был предложен В.В.Морковкиным и А.В.Морковкиной в их фундаментальной монографии «Русские агнонимы (слова, которые мы не знаем)» (1997) для обозначения «совокупности лексических и фразеологических единиц родного языка, которые неизвестны, непонятны или малопонятны многим его носителям (ср. *обабок* – «*подберезовик*», *обжорный ряд* – «*место торговли горячей пищей*» и т.д.)» [Морковкин: 86].

Агнонимы, выявляемые, как правило, путём лингвистического эксперимента,

описываются с позиций «знаю – не знаю», «понимаю – не понимаю», «правильно – неправильно». Между тем, сегодня высказываются идеи не столько о предписательном, сколько об описательном аспекте рассмотрения «неизвестных» слов. Так, Н.Д. Голев считает, что понятие «агнонимия» нужно рассматривать на более глубоком уровне, основываясь не на знаниях субъекта о слове, но на степени понимания этого слова [Голев: 93].

Описательный подход к подобным словам позволяет выдвинуть гипотезу о связи степени агнонимичности слов и стратегий их толкования. Данная гипотеза может быть верифицирована экспериментальным путём. Проведённый в рамках настоящего исследования эксперимент по выявлению особенностей обыденного толкования слов позволяет установить определённую зависимость выбора стратегии толкования от степени агнонимичности слов.

Суть эксперимента состояла в следующем. Испытуемым, в качестве которых выступили студенты-филологи и медицинские работники, предлагалось определить значение сорока пяти слов религиозной тематики. Словами-стимулами явились три группы слов: слова с предметным значением *Икона, Кадило*; религиозные термины *Благовест, Канонизация*; абстрактные религиозные понятия *Благочестие, Милосердие*.

Как показывают результаты эксперимента, в группе слов с предметным значением коммуниканты чаще используют идентифицирующую стратегию при определении слов, находящихся на дальней точке агнонимичности, т.е. при характеристике общеизвестных и частотных в языке слов. Как представляется, выбор стратегии позволяет дать широкое определение известному предмету, описать его со всех сторон, выделить определённые признаки: *Кадило - сосуд с благовониями, с которым священник ходит по церкви*.

Среднюю степень агнонимичности отражает выбор классифицирующей стратегии, проявляющейся в тех случаях, когда коммуникант может заменить слово известным ему синонимом: *Придел - пристройка в храме*.

При толковании слов, входящих в зону ближней агнонимичности (нечёткое, поверхностное представление о значении слова) продуктивной является мотивационная стратегия, представляющая собой толкование слова через его внутреннюю форму. Подобным образом, как правило, интерпретируются заимствованные и сложные слова: *Самопожертвование - принесение себя в жертву*.

Слова, входящие в агнонимическое ядро (полное незнание значений слов), в большинстве случаев толкуются путём ассоциативной стратегии: *Я думаю, похоть - неудержимое желание*. Однако ассоциативная стратегия частотно используется и в комплексе с другими стратегиями: в объективное толкование коммуникант привносит субъективное значение слова, фиксирует свои ассоциации.

Самая низкая степень агнонимичности обуславливает идентифицирующую и классифицирующую стратегии. Опрашиваемый обладает значительным объёмом знаний о слове, может дать его характеристику, выделив частные признаки, или заменить слово известным ему синонимом. Среднюю степень агнонимичности отражает выбор иллюстрирующей стратегии. Коммуникант может воссоздать только ситуацию, характерную для употребления слова, но затрудняется перечислить его признаки и свойства. Контекст в данном случае выполняет распознающую функцию. На высокую степень агнонимичности указывают мотивационная и ассоциативная стратегии. Используя мотивационную стратегию, коммуникант обращается к форме слова, пытаясь через этимологический анализ придти к содержанию. Однако нередко встречается явление, называемое «народная этимология»: опрашиваемый, руководствуясь лишь своим языковым чутьём и обыденными представлениями, ошибочно определяет деривационные связи слова. Выбор ассоциативной стратегии обусловлен не точными знаниями коммуниканта значения слова, а теми ассоциациями, которые находятся в подсознании и представляют собой набор субъективных понятий. Полными агнонимами

можно назвать слова, интерпретируемые с помощью отрицательного способа, т.е. опрашиваемый либо затрудняется привести значение слова, либо даже не пытается его определить.

Таким образом, проведённый анализ подтверждает гипотезу о зависимости стратегии толкования от степени агнонимичности слов, т.е. от объёма знаний о слове, который есть у коммуниканта.

Литература

Голев Н.Д. Лингвистические тупики юридической техники // Юрислингвистика-8. Межвузовский сборник научных трудов. Кемерово-Барнаул, 2007. С.92-104.
Морковкин В.В., Морковкина А.В. Русские агнонимы (слова, которые мы не знаем). М., 1997.

Характеристика участников ситуации в русском языке

Митева Полина Степановна

Аспирантка Московского государственного университета имени М. В. Ломоносова,
Москва, Россия

Для выражения характеристики событий и участников события языки мира используют следующие конструкции: *конструкция вторичной предикации*, или *депиктивная* (DepC, пример (1)) и *адвербиальная* (AdvC, пример (2)):

(1) И он **сердито** ушел с площади, величественно запахнувшись в свою зеленую мантию. [НКРЯ]

(2) Он так рассердился на дедушку, что даже в садик пришёл **сердитый**. [НКРЯ]

Возможна также *атрибутивная* конструкция, но в данной работе она рассматриваться не будет. DepC отличается от AdvC тем, что в DepC морфосинтаксически маркируется связь с тем выражением, которое характеризуется на семантическом уровне (*контролёром*): в русском языке с помощью согласования по роду, числу и, в части конструкций, по падежу.

Прототипически AdvC используется для характеристики события, а DepC – для характеристики участников события. Однако это верно не всегда: в (2) и (3) наречия характеризуют участника: его эмоциональное состояние в (2) и свойство характера в (3):

(3) Петя **благоразумно** промолчал.

Традиционно AdvC и DepC изучаются отдельно. Однако при внимательном рассмотрении оказывается, что в языках мира эти конструкции покрывают непрерывную семантическую область ([Himmelman & Schultze-Berndt]). В некоторых языках одно и то же значение может передаваться с помощью обеих конструкций, т.е. они вступают в конкуренцию, как, напр., в примере (4) из английского языка:

(4) a. The boy returned to his parents **hungry**.

b. The boy **hungrily** returned to his parents.

Рассмотрим основные группы выражений, характеризующих участников.

1. Характеристика агенса

К выражениям, характеризующим агенса, относятся наречия *благородно*, *легкомысленно*, *неосторожно*, *трусливо* и др. Эти выражения дают характеристику агенту ситуации на основании совершаемого им события. В примере (3) Петя характеризуется как *благоразумный* человек на основании того, что он промолчал. Свойство, которым наделяется участник ситуации, рассматривается как его постоянное свойство. В языках мира характеристика агенса выражается с помощью AdvC, но не DepC:

(5) *Петя вернулся **неосторожным**.

Наречия данного класса могут иметь две интерпретации: «верхнюю» (6а), при которой характеристику получает участник события, и «нижнюю» (6б), которую можно условно назвать интерпретацией образа действия. (6а) ~ ‘Вася ответил на вопрос, так как он неосторожный (человек)’, (6б) ~ ‘Вася говорил неосторожным образом’.

(6) a. Вася **неосторожно** ответил на вопрос.

б. Он говорил **неосторожно**.

Наречия ориентированы на семантическую роль агенса независимо от его синтаксического выражения. Так, в (7) агенс выражен дополнением в твор. падеже.

(7) Ключ был легкомысленно оставлен **нами** дома.

Самой частотной синтаксической позицией для наречий данного класса с верхней интерпретацией является позиция между подлежащим и сказуемым. При этом практически единственный допустимый коммуникативный статус – *комментарий* (термин из [Рудницкая: 39]). Наречия как бы стоят за пределами деления на тему и рему: дополнительно к тому, что говорящий хотел сообщить о событии, он дает характеристику агенту этого события. Постглагольная позиция наречия и фокус на наречии делают предпочтительной его нижнюю интерпретацию.

Характеристика состояния субъекта

Состояние субъекта во время совершения какого-либо действия может быть выражено как с помощью DepC (пример (1)), так и с помощью AdvC (пример (2)). К выражениям, характеризующим состояние субъекта, можно отнести прилагательные *голодный, голый, пьяный, веселый, печальный* и др. и образованные от них наречия. Рассмотрим факторы, от которых зависит распределение AdvC и DepC.

1) *Синтаксическая позиция контролёра*

В то время как наречия состояния ориентированы на семантическую роль субъекта, для вторичных предикатов важна синтаксическая позиция контролёра (DepC с вторичным предикатом в твор. п. может относиться к подлежащему и прямому дополнению, DepC с согласующимся предикатом может относиться также к другим членам предложения). Поэтому невозможно сказать *Петя печально_i встретил Васю_j*, так, чтобы *печально* характеризовало Васю, но возможно *Петя встретил Васю_i печальным_i /печального_i*.

2) *Коммуникативный статус контролёра*

Коммуникативные статусы темы и узкой ремы нехарактерны для наречий состояния, для вторичных предикатов они естественны и частотны.

3) *Семантика характеризующего выражения*

Если первые два фактора разрешают использование обеих конструкций, вступает в действие семантический фактор.

Наречия состояния имеют две интерпретации: «верхнюю», характеризующую состояние субъекта, и «нижнюю», характеризующую внешние проявления этого состояния. В подавляющем большинстве случаев две интерпретации совмещаются (пример (8)), однако возможны контексты, в которых возможна только одна из этих интерпретаций: только верхняя в (9), только нижняя в (10).

(8) Вера **сердито** сдёрнула наволочку и бросила на пол. [НКРЯ]

(9) Когда Леон засыпает в моем сне, я **радостно** просыпаюсь. [НКРЯ]

(10) Он **печально** скривился, демонстрируя огорчение. [НКРЯ]

В примерах типа (9) AdvC не может быть заменена на DepC без потери части значения (ср. *Вера сдёрнула наволочку сердитая* – теряется компонент ‘Вера сдёрнула наволочку сердитым образом’). В примере (10) возможно заменить наречие на вторичный предикат, так как нижняя интерпретация наречия недоступна (нельзя проснуться радостным образом). Для некоторых наречий (напр., *голодно, голо*) верхняя интерпретация невозможна или нехарактерна, поэтому состояние участника может быть передано только с помощью DepC.

Итак, в русском языке практически не возникает конкуренции AdvC и DepC. Их распределение зависит от рассмотренных факторов.

Литература

Рудницкая Е.Л. Сентенциальные наречия в русском языке. Диссертация ... кандидата филологических наук. М., 1993.

Himmelman N., Schultze-Berndt E.F. (eds.) Secondary predication and adverbial modification. The Typology of depictives. Oxford, 2005.

Национальный корпус русского языка: <http://ruscorpora.ru/>

Модель семантического описания деепричастной системы на примере финно-пермских языков

Муравьев Никита Алексеевич

Студент Московского государственного университета имени М.В. Ломоносова,
Москва, Россия

Рассматривая существующие описания семантики деепричастных форм в языках финно-пермской группы, можно заметить, что финно-пермские деепричастия в большинстве случаев являются, по классификации в [Недялков 1995], специализированными, то есть имеющими одно или два значения. При этом деепричастия, как правило, делятся на два семантических класса. К первому классу относятся деепричастия, выражающие временной контекст совершения главного действия, а ко второй – деепричастия, выражающие различные «невременные» обстоятельственные значения. Различие между ними подробно описано в работах [Недялков 1995] и [Храковский 2009]. Ср. пример 1 а) и б), где представлены деепричастия первого и второго класса.

1. а) pukal-yg-as sya n'ikod-ked oz s'ornit (кз)
сидеть-CONV-POSS3SG он никто-COMIT NEG.NPST.3 разговаривать
Сидя (дома), он ни с кем не разговаривает.

- б) sya rušj-is ki-nas mašit-yg-moz (кз)
он бежать-PST.3SG рука-INSTR.POSS3SG размахивать-CONV-ATT
Он бежал, размахивая руками.

Также считается, что если деепричастия, относящиеся к первому классу, выражают три типа таксисных отношений, то деепричастия второй категории, исходя из классификации [Недялков 1995] таксисной соотнесенности не имеют. Однако в работе [Храковский 2009] указывается на то, что такие деепричастия имеют фоновый таксис, иными словами, выражаемое ими действие в любом случае хронологически соотнесено с главным действием. Материал финно-пермских языков склоняет нас принять именно такую точку зрения и рассматривать таксисную соотнесенность в качестве одного из потенциальных параметров модели описания. Поэтому возникает сразу два вопроса: можно ли считать таксисную соотнесенность обязательным признаком любой деепричастной формы? Если да, то как относиться к случаям, когда одна форма может выражать более одного таксисного значения?

Первый вопрос возникает в связи с существованием в коми-зырянском и удмуртском языках, наряду с деепричастием следования особой формы на -мон/-тэп (пример 2), которая также характеризует второстепенное действие, совершаемое после главного, но таксисная соотнесенность является фоновым компонентом значения, что дает основания относить эту форму ко второй категории деепричастий.

2. та пи-ед ужа-мон буд-эм (у)
твой сын-POSS2SG работать-CONV вырасти-PST.3SG
Твой сын вырос настолько, что может работать.

Однако для ответа на первый вопрос нужно также рассмотреть все остальные деепричастия второго класса и определить таксисные значения, которые могут выражаться этими формами. В каждом из трех языков к таковым относятся, прежде всего, деепричастные формы сопровождающего обстоятельства (пример 3) и отсутствующего действия (пример 4), первые из которых могут обозначать только одновременное действие, а вторые – как одновременное, так и предшествующее действие.

3. tuj vyl-ət gurət-əmən mun-ə gus'enič'nəj traktor (кп)
дорога верх-PROL гудеть-CONV идти-PRS.3SG гусеничный трактор
По дороге, гудя, идет гусеничный трактор.

4. соос	дор-ысь	выль	машина-зЭС	адЪЫ-тЭК
они	край-EL	новый	машина-POSS3PL	увидеть-CAR
ой		кошки (у)		
NEG.PST.1		уйти		

Я не ушел от них до тех пор, пока не увидел их новую машину.

Кроме того, ко второй категории относится удмуртское деепричастие со значением причины на -emen, имеющее свою отрицательную форму на -mteen (таксис предшествования или одновременности) и коми-зырянские формы сопровождающего состояния на -ana/-anaen/-anarugus' и положения на -sa (таксис одновременности).

Таким образом, форма на -mon, выступает аналогом деепричастия сопровождающего обстоятельства для таксиса следования. Но у удмуртского деепричастия причины, на первый взгляд, нет подобных аналогов. Проблема заключается в том, что значение причины нельзя считать просто подтипом значения сопровождающего обстоятельства, поскольку значение сопровождающего обстоятельства является внешней характеристикой главного действия, а значение причины указывает на логическое отношение между действиями и, тем самым, является скорее внутренней характеристикой главного действия. Таковую же внутреннюю характеристику, только с таксисом следования, выражает инфинитив в целевом значении.

Ответ же на вопрос, как относиться к случаям, когда одна форма может выражать более одного таксисного значения, весьма прост: такие формы, в противоположность специализированным, следует относить к контекстным [Недялков 1995]. Но поскольку каждое деепричастие характеризуется по двум признакам: таксисному отношению и типу характеристики, то и контекстные формы бывают трех типов: контекстные по таксису, контекстные по характеристике и контекстные по обоим признакам.

Таким образом, в описанной модели финно-пермские деепричастия различаются по двум параметрам: таксисное отношение (одновременность, предшествование, следование) и тип характеристики (временной контекст, внешняя характеристика, внутренняя характеристика). Из всех девяти логически возможных комбинаций значений в финно-пермских языках отдельные формы существуют только для шести из них – всех характеристик отношения одновременности и следования. Отношение предшествования же выражается контекстно некоторыми формами одновременности. Сокращения: кз – коми-зырянский, кп – коми-пермяцкий, у – удмуртский, АТТ – аттенуатив, CAR – каритив, COMIT – комитатив, CONV – деепричастие, EL – элатив, INSTR – инструменталис, NEG – отрицание, NPST – непрошедшее время, PL – множественное число, POSS – посессив, PROL – пролатив, PRS – настоящее время, PST – прошедшее время, SG – единственное число, 1, 2, 3 – лицо.

Литература

- Nedjalkov, V. P.* The meaning of converb constructions // *Converbs in Cross-Linguistic Perspective*, ed. Martin Haspelmath, Ekkehard König. Berlin, Mouton de Gruyter, 1995
Храковский В. С. Типология таксисных конструкций. М. «Знак», 2009

Особенности функционирования глаголов оценки

(по данным лингвистического и психолингвистического эксперимента)

Никитина Янина Владимировна

Студентка Московского городского педагогического университета, Москва, Россия

Пожалуй, желание и умение давать оценку является одним из фундаментальных свойств человеческой психики. Человеку свойственно оценивать предметы и явления окружающего мира, других людей, собственные мысли, слова и действия. Вне всякого сомнения, эта функция так или иначе реализовывалась уже на этапе первобытного общества, усложняясь по мере развития человека как биологического вида и человеческого общества. Одновременно с появлением новых, более сложных

аксиологических характеристик, расширением и усложнением системы оценок, усовершенствовалась и система средств их вербальной репрезентации. Современный язык богат единицами разных уровней, так или иначе репрезентирующих оценку. На разных этапах развития лингвистической науки исследователей привлекали те или иные параметры оценочных суждений, появлялись новые методы анализа способов языковой реализации оценки, аккумулировались знания о ее природе, структуре и принципах организации.

Представляется, что одной из наиболее сложных и на данный момент плохо изученных лексических групп, связанных с репрезентацией оценки, является группа глаголов оценки, представляющих собой единицы метакоммуникации (например, *assess, evaluate, appraise, value, rate, estimate*). Эти глаголы широко используются в тех случаях, когда речь идет об оценке чего-либо / кого-либо и фактически всегда сопровождаются другими языковыми средствами, эксплицитно или имплицитно репрезентирующими оценку. Однако в лексикографических источниках, как правило, предлагается очень приблизительное описание семантики данных глаголов, что приводит к неучету их специфики и ошибкам в их использовании. В ряде случаев они трактуются друг через друга, что приводит к возникновению «порочных» кругов.

Для того чтобы дифференцировать значение этих глаголов и определить сферу их функционирования, было проведено экспериментальное исследование. Следует отметить, что глаголы оценки могут изучаться по крайней мере с двух перспектив: с одной стороны, возможно проведение семантического эксперимента для уточнения значения анализируемых единиц и выявления совокупности интегральных и дифференциальных семантических признаков; с другой стороны, представляется целесообразным проведение психолингвистического эксперимента (с опорой на ассоциативный эксперимент), направленного на выявление тех актуальных связей, которые обнаруживаются в языковом сознании носителей языка. Для того чтобы сравнить те результаты, которые можно получить с опорой на разные типы эксперимента, и установить корреляцию между ними, проведенное нами исследование проводилось в два этапа.

На первом этапе был проведен семантический эксперимент, в ходе которого испытуемым-носителям языка предлагалось оценить предъявленные высказывания и прокомментировать свою оценку. При моделировании стимульных высказываний были использованы разные глаголы оценки при полном совпадении контекста. Исследование проводилось в несколько этапов, каждый из которых предварялся корректировкой выдвинутых гипотез.

В результате исследования были сделаны следующие выводы относительно семантики рассмотренных глаголов оценки:

Assess: 1. оценивать нейтрально (неодушевленные и одушевленные предметы); 2. оценивать в числовом выражении;

Appraise: 1. давать самооценку (категоричную); 2. оценивать негативно (неодушевленные и одушевленные предметы); 3. экспертная оценка в числовом выражении (дорогостоящие вещи: недвижимость, ювелирные украшения);

Estimate: оценивать в числовом выражении (мнение дилетанта);

Evaluate: оценивать положительно (конкретные явления);

Value: 1. давать самооценку (положительную); 2. оценивать положительно (абстрактно); 3. оценивать в числовом выражении;

Rate: 1. давать самооценку (объективную); 2. оценивать объективно (однозначная шкала оценок, набор определенных критериев) в относительном формате; 3. оценивать в числовом выражении

На втором этапе исследования был проведен ассоциативный эксперимент, направленный на выявление особенностей структуры и содержания образа сознания (концепта) ОЦЕНКА и установление корреляций между компонентами образа сознания

и «овнешняющими» их языковыми единицами (в частности, глаголами *assess, evaluate, appraise, value, rate, estimate*). Иными словами, результаты психолингвистического эксперимента позволили уточнить выводы, сделанные на основе данных, полученных в ходе лингвистического исследования, и дополнить их.

Факторы, влияющие на наличие подлежащего местоимения при сказуемом в русском языке

Павлова Елизавета Кирилловна

*Студентка Московского государственного университета имени М. В. Ломоносова,
Москва, Россия*

Русское подлежащее может быть выражено полной именной группой или редуцированным способом, т.е. местоимением, а также может быть не выражено посредством именной группы.

А.М. Пешковский писал об употреблении подлежащего так: «в первых двух лицах подлежащее все же большей частью употребляется, несмотря на свою ненужность. <...> Таким образом, присутствие личных слов можно рассматривать в русском языке как норму, а отсутствие их — как отступление от нормы, имеющее всякий раз свои причины и свой смысл» [Пешковский: 190].

Цель данной работы — выявить факторы, влияющие на наличие или отсутствие личного местоимения при глаголе, а также установить, когда местоимение чаще присутствует, когда нет.

Методом исследования данной проблемы является корпусный анализ [Национальный корпус русского языка]. Каждому примеру при обработке давалась характеристика по разным критериям, касающимся как типа клаузы в целом, так и ее главных членов. После обработки примеров производился количественный анализ.

Особенностью референциальной системы современного русского языка является то, что в прошедшем времени глагольная форма несет в себе информацию о роде и числе подлежащего, а в настоящем времени — о числе и лице.

Гипотезы в работе выдвигались следующие. Во-первых, тот факт, что в настоящем времени лицо выражено в глаголе, а в прошедшем нет, заставляет предположить, что личные местоимения чаще употребляются в прошедшем времени. Во-вторых, предполагается, что в 3 лице настоящего и будущего времен местоименное подлежащее реже опускается, чем в 1 или 2, что связано с тем, что местоимения 3 лица единственного числа дают дополнительную информацию о референте: о его роде.

Подробно исследовано каждое из времен. Ниже приведен пример исследования предложений, содержащих в себе форму глагола в настоящем времени.

Лицо сказуемого

В предложениях с 1, 2 и 3 лицом преобладает наличие подлежащего.

Примеры с нулевым подлежащим в 3 лице или являются следствием деления конструкций с однородными членами на более мелкие части (в разговорной речи):

(1) *А он сейчас не у неё живёт... Съехал... Дом снимает.*

или возможны случаи, когда опущено подлежащее в сентенциальном актанте:

(2) *Он с неохотой сказал, что **хочет** наконец **иметь** нормальную обычную семью...*

Число сказуемого

И в единственном, и во множественном числе местоимение чаще есть, чем нет, причем множественное число стимулирует выражение подлежащего.

Лицо и число сказуемого

И в единственном, и во множественном числе преобладают предложения с подлежащим.

Характеристики текста

Местоимение чаще всего отсутствует в устной речи (разговор праздный – устная непубличная речь). В разговорной речи важное место занимают локуторы – участники диалога, поэтому сказуемое в 3 лице не так часто по сравнению со сказуемыми в первом и втором.

Опущения местоимения в специальном стиле, стремящемся к максимально правильному и понятному выражению, практически не бывает.

Тип клаузы

В таких типах клауз, как сентенциальные актаны, сентенциальные обстоятельства и простые предложения в составе сложносочиненного, предложений с опущенным подлежащим почти не бывает, хотя на первый взгляд в придаточных предложениях следовало бы ожидать преобладания опущенных подлежащих, если референты в главном и зависимом предложениях совпадают:

(3) *Я сказал, что **лицу я** только своим врагам, а она для меня уже никто.*

Нулевое подлежащее часто бывает в простых клаузах:

(4) *И не всегда так будет... **Подумываю** всерьёз о поиске работы...*

Порядок следования подлежащего и сказуемого

Местоимение обычно предшествует глаголу. В немногочисленных примерах, где местоимение следует после глагола, акцент делается на глагол (выделяемым с помощью необычного порядка слов) (пример (4)).

В процессе исследования выяснилось, что на наличие подлежащего особенно влияло лицо сказуемого, тип клаузы, а также характеристики текста, такие как тип текста и сфера функционирования. Такие критерии, как тип сказуемого и его род (в прошедшем времени), особого влияния на наличие подлежащего местоимения не имели.

Местоимение чаще всего отсутствует в устной непубличной речи и в простых предложениях (которые, в свою очередь, преобладают в разговорной речи), а также во вводных.

Личные местоимения чаще употребляются в прошедшем времени, потому что в настоящем времени лицо выражено в глаголе, то есть первая гипотеза подтвердилась, что видно по таблице 1.

	местоименное подлежащее	подлежащее не выражено	Всего
настоящее-будущее	56%	44%	100%
прошедшее	67%	33%	100%

таблица 1. наличие подлежащего в зависимости от времени

В настоящем и будущем временах местоимение опускается чаще в 3 лице вопреки ожиданиям, хотя изначально предполагалось, что в 3 лице есть дополнительная мотивация использования местоимения: оно характеризует род референта. Итак, вторая гипотеза не подтвердилась, результаты видны на таблицах 2 и 3.

	местоименное подлежащее	подлежащее не выражено	Всего
настоящее-будущее	58%	42%	100%
прошедшее	69%	31%	100%

таблица 2. наличие подлежащего в зависимости от времени (1-2 лицо единственное число)

	местоименное подлежащее	подлежащее не выражено	Всего
настоящее-будущее	54%	46%	100%
прошедшее	62%	38%	100%

таблица 3. наличие подлежащего в зависимости от времени (3 лицо единственное число)

Литература

Пешковский А.М. Русский синтаксис в научном освещении М., 2001

Национальный корпус русского языка: <http://www.ruscorpora.ru/>

Типы взаимодействия импликативных свойств предикатов второго порядка с контекстом

Попова Дарья Павловна

Аспирантка Московского государственного университета имени М. В. Ломоносова, Москва, Россия

Вводные замечания. Данная работа посвящена эмпирическому исследованию взаимодействия импликативных свойств предикатов второго порядка со значениями других морфем и фраз в предложении. Под импликативными свойствами понимаются следствия, порождаемые предикатами второго порядка. То, что импликативные свойства предикатов второго порядка необходимо рассматривать в контексте, было впервые отмечено в [Зализняк]. Данная работа отличается от этого пилотного исследования как более широким охватом материала, так и теоретическим взглядом на то, почему предикаты второго порядка имеют разные следствия в различных контекстах. В [Зализняк] это явление рассматривается как *изменение* импликативных свойств предикатов под влиянием факторов контекста. Я покажу, что, на самом деле, имеет место *взаимодействие* импликативных свойств предикатов второго порядка с такими факторами контекста, как семантика вида, тип союза, лицо и время. Я продемонстрирую, что такого рода взаимодействие бывает двух типов, которые я буду называть композициональным и лексическим взаимодействием соответственно. Ни один из этих типов взаимодействия не предполагает *изменения* импликативных свойств предикатов второго порядка, в чём состоит основное отличие от позиции [Зализняк]. Ниже будут кратко охарактеризованы два типа взаимодействия на примере глаголов *стесняться* и *видеть*.

Метод исследования. Данное исследование опирается на два корпуса для глаголов *стесняться* и *видеть* соответственно. Каждый корпус представляет собой коллекцию из около ста релевантных примеров из Национального корпуса русского языка. Примеры были аннотированы по ряду параметров, таких как коммуникативный тип предложения, форма предикации, тип союза, морфологические признаки главного и зависимого предикатов, показатели субъективной и объективной модальности и др. В качестве статистического метода, позволяющего определить, какие параметры контекста из перечисленных выше взаимодействуют со следствиями, порождаемыми предикатами второго порядка, использован *метод главных компонент* [Ваауен: 127–135].

Композициональный тип взаимодействия. Данный тип взаимодействия характеризуется тем, что следствие из предложения с предикатом второго порядка складывается композиционально из следствия, порождаемого предикатом второго порядка, и из следствия, порождаемого видом инфинитива подчинённого предиката. Метод главных компонент показывает взаимную корреляцию между формой совершенного вида подчинённого инфинитива и следствием истинности подчинённого предложения (1) и между формой несовершенного вида подчинённого инфинитива и отсутствием однозначного следствия (2):

(1) *Потом он исчез, и она стеснялась спросить о нём.*

Следствие: *она не спрашивала о нём.*

(2) *Один мой друг стесняется знакомиться с девушками на вечеринках.*

Следствие: *может быть, знакомится, может быть, нет.*

Однако взаимодействие с импликациями других элементов в предложении позволяет специфицировать следствие в случаях с несовершенным видом подчинённого инфинитива (такая спецификация наблюдается в 20% случаев):

(3) *Сам я тогда стеснялся выпивать и заказывал себе только морс.*

Сам-импликация: *другие люди, противопоставленные говорящему и активированные в контексте, не стеснялись выпивать.*

Следствие из сам-импликации: *другие люди, активированные в контексте, выпивали.*

Тогда-импликация: *сейчас говорящий не стесняется выпивать.*

Только-импликация: *говорящий не заказывал себе никаких алкогольных напитков.*

Следствие: *говорящий не выпивал.*

В докладе будет подробно рассмотрен механизм взаимодействия импликаций контекста и следствий, порождаемых предикатами второго порядка.

Лексический тип взаимодействия. Лексический тип характеризуется тем, что разные лексические значения предиката второго порядка имеют разные следствия. Так, глагол *видеть* как глагол чувственного восприятия имеет следствием истинность подчинённой предикации (4), тогда как глагол *видеть* как глагол мнения не обладает таким следствием (5):

(4) *Недавно ехала в метро и видела, как/что девушка подпиливает себе ногти.*

Следствие: *девушка подпиливала себе ногти.*

(5) *Я вижу, что/*как Ване страшно приятно.*

Следствие: *может быть, Ване приятно, может быть, нет.*

Метод главных компонент показал, что тип союза может служить надёжной диагностикой для определения следствий и, соответственно, различения двух значений глагола. Глагол *видеть* как глагол чувственного восприятия сочетается с союзами *что* и *как* и с союзными словами. Глагол *видеть* как глагол мнения сочетается с союзом *что*, который при отрицании меняется на союз *чтобы*, и не сочетается с союзом *как* и с союзными словами. В докладе будут подробно рассмотрены возможные причины такого взаимодействия.

Заключение. Данная работа представляет собой обобщение обширного эмпирического материала по взаимодействию между следствиями, порождаемыми предикатами второго порядка, и следствиями и импликациями контекста. Как показано, такое взаимодействие бывает двух типов: композициональное и лексическое. Эмпирические данные, а также попытка их анализа представляют интерес как для исследования свойств предикатов второго порядка и проблемы проекции значений, так и для решения задач автоматического выявления следствий и импликаций из контекста и снятия семантической омонимии.

Литература

Зализняк Анна А. О понятии импlicative типа (для глагола с пропозициональным актантом) // Логический анализ языка. Знание и мнение. М., 1988. С.107–121.

Baayen R. H. Analyzing Linguistic Data: A Practical Introduction to Statistics using R. Cambridge, 2008.

Употребление артиклей носителями безартиклевого языка

Пушкова Варвара Владимировна

Студентка Пермского государственного университета, Пермь, Россия

Категория определенности-неопределенности – это понятийная категория, универсальная для всех языков, однако средства ее выражения различны: в одних языках для нее существуют грамматические показатели (артикли), в других они отсутствуют. Несмотря на универсальность категории, носителю безартиклевого языка нелегко понять закономерности, которым подчиняется артикль. Эта проблема всегда существовала при изучении иностранных языков (в частности, английского). Для психолингвиста здесь возникает множество вопросов: как воспринимается артикль носителем безартиклевого языка? различается ли понимание значения артикля у билингва и монолингва? Какие факторы влияют (у билингвов) на расстановку артиклей в тексте?

Н.В. Боронникова и Ю.А. Левицкий в работе «Артикль. Вопросы теории и типологии» рассматривают три аспекта значений артикля: показатель предметности; показатель функционального членения; синтаксический показатель. Мы рассмотрим первые два аспекта.

Для эксперимента нами был взят текст английской народной сказки *Catskin* с общеизвестной лексикой и простой грамматикой. Учитывая характер текста, можно сказать, что на расстановку артиклей не влияет авторская редакция. Информантами были носители учебного билингвизма – 15 студентов 3 и 4 курсов, изучающие английскую филологию. Им было предложено расставить в тексте артикли (среди 58 случаев употребления артиклей в тексте сказки встретились 29 определенных артиклей, 21 неопределенных и 8 нулевых).

Рассмотрим данные ошибочной расстановки артиклей билингвами. К ошибкам, связанным с артиклем как показателем предметности можно отнести следующие.

Конкретизация: подстановка *a => the* происходит, когда предмет, в данном тексте имеющий значение «один из многих, представитель класса» понимается как «конкретный, единственный», например, в сочетаниях: *a cloak with a hood, a nutshell, a hollow tree* (They were so light and so magic that she could pack all three into **a** nutshell).

Генерализация: наоборот, нечто единственное превращается в «типичное»: *a king, a palace*.

Определенный артикль опускается при числительных- субстантивах в 87% случаев: **the** second was to be silver like **the** moon (можно предположить, что для билингвов обозначение предметности через артикль не первостепенно)

Известно, что артикль в английском языке указывает на число существительного: неопределенный артикль не ставится при существительных во множественном числе. На нашем материале можно сделать вывод о том, что определенный артикль легко опускается именно при существительных во множественном числе: от 30 до 90% случаев. В некоторых ситуациях причиной может быть смысловая избыточность артикля (см. ниже), в других этот факт может быть приравнен к оценке существительного как неопределенного.

Интересен случай употребления неопределенного артикля там, где его не должно быть, т.е. при неисчисляемых существительных. Приписывая ему артикль *a*, билингв фактически превращает его в исчисляемое.

На расстановку артикля в тексте оказывает сильное влияние фактор сочетаемости: артикль в сочетаниях с числительными, прилагательными опускается значительно чаще, чем при существительных без определений.

Наиболее проблемным оказался артикль в коммуникативной функции. При повторном употреблении слова (то есть, когда оно переходит из категории нового в категорию данного) определенный артикль опускается (от 27 до 80% случаев). Вообще повтор чего-либо, например, определения (*a gold ring, a gold necklace, a gold brooch*) склоняет к тому, чтобы пропустить артикль (20%). То есть артикль, как определенный, так и неопределенный, маркирует появление чего-то нового в тексте; там, где он не выполняет интродуктивной функции, он легко «выпадает». Наши результаты говорят о том, что определенный артикль в теме опускается намного чаще, чем неопределенный в реме. Получается, что «данная» информация (с точки зрения билингва) не нуждается в особом оформлении, что для говорящего важно привлечь внимание к новой информации и подчеркнуть ее с помощью артикля.

Анализ показал, что особое значение при расстановке артикля билингвами имеют фактор сочетаемости и коммуникативный фактор, а самые частотные ошибки - это пропуск артикля перед прилагательными и числительными и пропуск артикля в теме.

Литература

Боронникова Н.В., Левицкий Ю.А. Артикль. Вопросы теории и типологии. Пермь, 1999

Репрезентация лексемы «деньги» в сознании череповецких школьников (по материалам ассоциативного эксперимента)

Рыкова Валерия Руслановна

Студентка Череповецкого государственного университета, Череповец, Россия

Цель нашей работы – проанализировать ассоциативное поле лексемы *деньги*.

Описание эксперимента. Эксперимент проводился в письменно-письменной форме. Респондентом предлагалась анкета, инструкция к которой была следующей: «Пожалуйста, на приведенные ниже слова ответьте первыми пришедшими на ум ассоциациями. Постарайтесь затратить на заполнение всей анкеты не более 5 минут».

В эксперименте приняли участие 78 человек. Всех респондентов можно разделить на две группы – учащиеся 5-х классов (11-12 лет) и 9-х классов (14-15 лет).

Стимульный ряд состоял из 30 слов. В докладе мы остановимся лишь на интерпретации реакций, полученных на стимул *деньги*.

Структурный аспект анализа.



На диаграммах видно, что в процентном соотношении преобладают тематические реакции. Это свидетельствует о том, что структура ментального лексикона респондентов достаточно сложна и расширена.

Школьники 9-х классов чаще дают синтагматические реакции, чем пятиклассники. На наш взгляд, в первую очередь это связано с нехваткой сведений о природе грамматических связей.

Что касается парадигматических реакций, то у школьников 9-х классов их гораздо меньше, чем у учащихся 5 класса.

Семантический аспект.

Самые частотные ответы девятиклассников связаны с количественной характеристикой стимула: *большие* – 4, *много* – 2, *много денег заработать*. Эту группу реакций можно определить как ядро ассоциативного поля учащихся 9-х классов.

Далее можно отметить, что достаточно широко представлены отождествительно-кваликативные реакции: *зло* – 2, *счастье* – 2, *сила*, *радость*, *власть*, *хорошо*, которые со стимулом образуют потенциальную синтаксическую модель (N_1 – это N_1).

К этой группе тесно примыкают реакции, расширяющие понимание денег как власти: *власть*, *правят миром* – 2, *могут многое*, т.е. некоторые подростки уже в таком возрасте отождествляют между собой деньги и власть.

Достаточно широко представлены реакции, характеризующие источник получения (*награда*, *много денег заработать*, *работа*) и целевое назначение (*покупка*, *покупки*, *траты*) стимула *деньги*.

Реакции, характеризующие деньги как обилие материальных ценностей, представлены в меньшем количестве: *богатство*, *достаток*, *обеспеченность*. Эту

группу реакций условно можно соотнести с первой (количество), но здесь помимо количества добавляется значение достатка, благополучия.

К первой группе (количественной) условно можно отнести и единичный ответ – *хотелось бы больше и нескончаемые*. Но здесь появляется значение желательности, и в целом подобная реакция имеет более личностный и экспрессивный характер.

Еще менее разнообразны реакции, характеризующие зрительное восприятие стимула: *бумажки, зеленые*.

Среди реакций появляется и жаргонная сниженная лексика, заменяющая основную номинацию стимула: *бабло*. Наличие подобного ряда реакций весьма характерно для респондентов данной аудитории

Ассоциации пятиклассников свидетельствуют, что у них еще нет четко сформированного отношения к деньгам. Для них деньги – это в первую очередь обилие материальных ценностей: *богатство - 7, богач, богатый человек*. Здесь количество не определено (даже размыто – *черезмного*), важен лишь сам факт наличия.

Определительных реакций почти нет: *все, хорошо*. Зато ситуация зрительного восприятия стимула представлена более широко, чем у учащихся 9-х классов: *зеленые бумажки, бумага (дорогая), банкнота, золото*. Думается, это, прежде всего, связано с тем, что у подростков данного возраста конкретно-чувственное восприятие играет более значительную роль, чем у учащихся 9-х классов.

Ситуация целевого назначения (*покупки – 2, тратить, обмен на вещи, средство оплаты, магазин, машина*) представлена гораздо более широко, чем источник получения. Вспомним, что у школьников 9-х классов подобных реакций было гораздо меньше, и они количественно соотносились с источником получения. Таким образом, можно сделать вывод, что младшие респонденты пока более озабочены вопросом, что сделать с деньгами, чем тем, откуда они берутся.

Появляется и совершенно новая группа реакций – название денежных единиц: *доллар – 2, рубли – 2, 10 копеек*. Отмечен и субъект ситуации: *мне*. В качестве реакции также была представлена числовая запись: *1000 000*.

Реакции, характеризующие источник получения денег, у пятиклассников и девятиклассников смещены к периферии, причем у учащихся 9-х классов подобные реакции представлены более разнообразно (у них в качестве источника получения выступает не только работа, но и *награда*), что, вероятно, связано с большим опытом респондентов.

В заключение заметим, что сопоставление возрастных вариантов ассоциативных полей дает возможность интерпретировать материал, преодолевая зоны нечетких ассоциативных смыслов.

Соотношение семантических признаков английских отадъективных существительных с их формальными характеристиками

Савочкина Елена Валерьевна

Ассистент, Смоленский государственный университет, Смоленск, Россия

Настоящая работа выполнена в рамках исследований, ведущихся школой количественной лингвистики профессора, доктора филологических наук Г.Г. Сильницкого в Смоленском государственном университете. Настоящая работа посвящена рассмотрению одного из относительно мало исследованных участков лексической системы современного английского языка – отадъективных существительных.

К исследованию привлекается словарный список отадъективных существительных, полученных путем 25% выборки адъективных основ из словаря *Advanced Learner's Dictionary of Current English*, и составивший в итоге 1412 единиц [Hornby].

Целью исследования является определение количественных характеристик отношений между семантическими признаками отадъективных существительных исходной выборки и фонологическими и морфематическими характеристиками данных основ. Задача установления межуровневых связей решается в рамках корреляционного анализа по критерию Пирсона. Обработка и анализ данных осуществлялся про помощи программы STATISTICA 6.0. Задаваясь уровнем надежности $p=0,95$ при указанном выше объеме выборки, мы считали значимыми коэффициенты корреляции Пирсона, для которых $r_{xy} \geq |0.06|$

К формальным признакам мы традиционно относим фонологические и морфологические признаки.

В качестве релевантных для целей данного исследования отобраны три категории **фонологических признаков**: 1) силлабическая структура (учет количества слогов), 2) акцентная структура (позиция ударного слога от зачина и исхода основы), 3) характер инициальной и финальной фонем исходного существительного (консонантизм или вокализм). При определении фонологических признаков мы пользовались словарем Cambridge English Pronouncing Dictionary.

Привлекаемые к настоящему исследованию **морфематические** признаки (гипоним, объединяющий понятия морфологического и словообразовательного анализа) являются внутрисклассовыми и отражают морфемную и деривационную структуру основы существительного. Обобщенные морфематические признаки отражают общие структурные характеристики основы: количество морфем, морфемная членимость/нечленимость, наличие аффикса, наличие одного/двух префиксов, наличие одного/двух/трех/четырёх/пяти суффиксов, наличие двух корней в составе основы, способ словообразования (суффиксацию, конверсию, словосложение).

При отборе **семантических признаков** для данного исследования мы опирались на общие, универсальные семантические классы, предложенные Г.Г.Сильницким первоначально для глагола [Сильницкий, Андреев: 10], а затем успешно перенесенные на другие части речи.

Энергетические признаки (ЭНРГ) отображают различные состояния и преобразования физической энергии: пространственные перемещения и местонахождение, изменение формы и структуры, физические и физиологические процессы. *Информационные признаки* (ИНФ) отображают различные виды переработки информации в психике человека и животных. К данному классу будут относиться существительные со значениями сенсорного и интеллектуального восприятия, эмоционального и волевого состояния, речи. *Онтологические признаки* (ОНТ) Они включают в себя экзистенциальные процессы и состояния, темпоральные и количественные характеристики, посессивность, наличие, социальные процессы.

Рассмотрим теперь корреляции семантических признаков с признаками остальных уровней. В таблицах представлены только значимые коэффициенты корреляции, пустым ячейкам соответствуют незначимые корреляции.

Табл. 1.1. Корреляции семантических, силлабических и фонемных признаков

	1 СЛ	2 СЛ	3 СЛ	4 СЛ	5 СЛ	6 СЛ	7 СЛ	8 СЛ	ГЗ	СЗ	ГИ	СИ	ГЗ- ГИ	ГЗ-СИ	СЗ- СИ	СЗ- СИ
ЭНРГ			0,13		-0,07	-0,07			0,09	0,09				-0,09	0,07	
ИНФ		-0,06	-0,09		0,09											
ОНТ			-0,07													

ГЗ- гласный зачин

СИ –согласный исход

Табл. 1.2. Корреляции семантических и акцентных признаков

	1	2	3	4	5	6	1'	2'	3'	4'	5'	6'	7'	2 уд
ЭНРГ	-0,16	-0,13		-0,06				0,05		-0,07	-0,05		-0,05	

ИНФ	0,10	0,06						-0,06		0,05				0,06
ОНТ		0,06							-0,05	0,05			0,07	

1- ударение от зачина
2' - ударение от исхода
2 уд - 2 ударения

Между парами признаками «энергетический»-«информационный», «энергетический»-«онтологический» и «информационный»- «онтологический» выявлено меньше оппозиций, чем случаев их отсутствия, что позволяет говорить о корреляционной гомогенности данных комплексов в пространстве фонологических признаков.

Табл. 2.1. Корреляции семантических и морфологических признаков

	1 морф	2морф	3морф	4морф	5морф	6морф	Чл	Нчл
Энрг	0,05	0,08		-0,09	-0,05			0,06
Инф		-0,10	-0,06	0,10	0,07			
Онт								

Табл. 2.1. (продолжение)

	Аф	Пф1	Пф 2	Сф 1	Сф 2	Сф 3	Сф 4	Сф5	2 корня	Сф	СС	Конв
Энрг	-0,06	-0,14	-0,06	0,07		-0,06	-0,06			-0,05		0,06
Инф		0,11				0,06	0,08		-0,05	0,05		-0,05
Онт	0,07		0,05						0,06			

Аф –наличие аффикса
Пф1 – наличие одного префикса
СС – существительное образовано словосложением
Конв - существ. образовано по конверсии
Сф - существ. образовано суффиксацией

Между парами признаков «энергетический»-«информационный» «энергетический»-«онтологический» и «информационный»-«онтологический» зафиксировано больше случаев оппозиций, чем их отсутствия, что позволяет говорить о **корреляционной гетерогенности** сопоставляемых признаков в пространстве морфологических.

В результате анализа корреляционных картин семантических признаков отадъективных существительных в пространстве их формальных признаков мы получили крайне неравномерное распределение корреляций между ними: для энергетического признака было зафиксировано 27 значимых корреляций, тогда как для информационного – 18 , а для онтологического всего 7. Это дало возможность отметить более тесную связь энергетического семантического признака и на наименьшую связь онтологического признака с привлеченными формальными характеристиками.

Литература

Сильницкий Г. Г., Андреев С. Н., Кузьмин Л. А., Кусков М. И. Соотношение глагольных признаков различных уровней в английском языке. Минск, 1990.
Hornby A. S. et al. Oxford Advanced Learner's Dictionary of Current English, the 3rd ed. Oxford, 1980.

Моделирование догадки о значении незнакомого слова иностранного языка (экспериментальное исследование)

Саркисова Элина Владиславовна

Студентка Тверского государственного университета, Тверь, Россия

При опознании незнакомого слова индивид прибегает к различным схемам действия, начиная с восприятия слова вне контекста, уделяя тем самым внимание форме слова, и заканчивая подключением контекста, энциклопедического знания и личностных факторов. Для проверки данной научной гипотезы нами был проведен эксперимент: на первом этапе от участников эксперимента (студентов V курса английского отделения

ИЯ и МК ТвГУ) требовалось дать реакции – предполагаемый перевод или ассоциативную реакцию на слово-стимул (SLIMNASTICS, NUDNIK, BOATEL, PARAKITING, NOGOODNIK, PROMISCUOUS, PULLMAN, BLIZZARD, MIDDLESCENT, EXCISE); второй этап предполагал опознание этих же слов, но уже в контексте.

Материал исследования: неологизмы/окказионализмы, слова английского языка.

Группу испытуемых (далее – Ии.) составили 10 студентов; на первом этапе работы от них получено 79 реакций на 10 слов. Наибольшее количество реакций вызвали слова: PULLMAN (10), BLIZZARD (8), MIDDLESCENT (8), что объясняется расширенным рассмотрением сети семантических отношений в лексике Ии., о чем свидетельствует анализ полученных результатов. Так, Ии. использовали не только синонимы (как в PARAKITING – *parashuting, парашлан, воздушный змей*), но и полисемию, родовидовые отношения (когда SLIMNASTICS рассматривалось как общее понятие – гипероним – для *aerobics, shaping, gymnastics, fitness*, выступающих в качестве гипонимов, NOGOODNIK оценивалось как *профессия*; PARAKITING – как *вид спорта*).

Обнаружены также ассоциативные и тематические отношения, выводящие на индивидуальную картину мира, что подчеркивается тем, что они не являются регулярными, имеют индивидуальный, часто окказиональный характер, о чем также свидетельствуют попытки Ии. проинтерпретировать слово путем «словотворчества», «словосочинительства»: NOGOODNIK – *новогодник*, MIDDLESCENT – *недонюхи*.

Возвращаясь к тематическим и ассоциативным отношениям, стоит отметить, что при интерпретации предложенных слов некоторые Ии. подключили свои фоновые знания: на слово-стимул BLIZZARD была получена реакция *Entertainment* от двух Ии. (*Blizzard Entertainment* – компания-разработчик компьютерных игр, реакция Ии. также записана с заглавной буквы); ср. также: PULLMAN – *Bill* (*Bill Pullman* – американский актер), EXCISE – *excel* (компьютерная программа), PROMISCUOUS – *Nelly Furtado* (канадская певица, “Promiscuous” – название песни). В одном из бланков реакция на слово-стимул PULLMAN выглядела следующим образом: *autor's name* (при смешении форм соответствующих слов английского и немецкого языков), очевидно, имелась в виду фамилия английского писателя Филиппа Пулмана.

При качественном анализе реакций отмечается обращение Ии. и к формальной составляющей предъявленных слов, а именно – к фонетическому, графическому и морфологическому образу слов. Анализ полученных вне контекста реакций свидетельствует о том, что в сознании и/или подсознании индивида существуют звуко-ассоциативные лексические поля [Сазонова: 51]: EXCISE – *exercise, excuse, excel, excise*; BLIZZARD – *lizard, близорукий*; SLIMNASTICS – *gymnastics, aerobics*; NOGOODNIK – *озорник, негодник* и др.

Второй этап эксперимента значительно сузил диапазон предложенных реакций, а также исключил поиск «общих» вариантов интерпретации типа *вид спорта, профессия*, о которых речь шла выше; это можно объяснить тем, что теперь слова были вписаны в контекст, который понимается Л.Г. Ким как «субъект, характеризующийся сложной семантической организацией и предопределяющий направления его осмысления и интерпретации читателем» [Ким: 1]. Теперь реакции Ии. во многих случаях были тождественны, например, слово PULLMAN определялось в основном как название средства передвижения (*She would take a boat, would catch a train and in Rome a pullman: тележка, теплоход, телега*). Значение слова BLIZZARD, вызвавшее многообразие интерпретаций на первом этапе эксперимента, теперь было определено как явление погоды (*The snow began to spin down again with the vindictive swirl of a blizzard: нурга, молния, вихрь*).

Нельзя не отметить, что контекст не всегда является объективной опорой в силу незнания других слов в предложении, в результате чего Ии. подключали свою фантазию для

вычленения смысла и продолжали использовать опору на внутреннюю структуру слова, как на первом этапе эксперимента.

На основе полученных данных можно выделить две качественно различные группы факторов, влияющих на опознание незнакомого слова, которые, как свидетельствуют полученные результаты, нельзя рассматривать отдельно, так как именно переплетение двух этих групп факторов обеспечивает доступ к верному значению слова:

1. Собственно языковые структуры: слова, их графические, фонологические, морфемные элементы; межсловесные связи; грамматика предложений; контекст с индикаторами, выводящими на выявление значения незнакомого элемента;
2. Неязыковые структуры, предполагающие непосредственное воздействие внешнего мира; влияние энциклопедического знания, индивидуальной картины мира посредством опыта; личностных особенностей; фоновых знаний, также знание нескольких языков, обуславливающее действие интерференции.

Итак, комбинация различных опор в процессе идентификации незнакомого слова обеспечивает более успешный доступ к значению слова. Установление связи между *словами-стимулами, единицами контекста* и наличными *единицами индивидуального лексикона* по сходству некоторых формальных признаков обеспечивает вхождение нового/незнакомого слова в контекст предшествующего опыта индивида и понимание значения слова.

Литература

Ким Л.Г. Субъективность текста как фактор вариативности его интерпретации // [http://proceedings.usu.ru/?base=mag/0066\(01_\\$042009\)&xsl=showArticle.xslt&id=a15&doc=../content.jsp](http://proceedings.usu.ru/?base=mag/0066(01_$042009)&xsl=showArticle.xslt&id=a15&doc=../content.jsp)

Сазонова Т.Ю. Моделирование процессов идентификации слова человеком: психолингвистический подход. Тверь, 2000.

Выбор референциального выражения в ситуации потенциального референциального конфликта: эффект «второго персонажа».

Успенская Анна Михайловна

*Студентка Московского государственного университета имени М. В. Ломоносова,
Москва, Россия*

Данное исследование относится к области экспериментального исследования дискурса – направления, к задачам которого относится изучение речевого взаимодействия собеседников в процессе коммуникации с помощью экспериментальных методик.

Работа опирается на теорию референции, согласно которой тип выбираемой говорящим номинации (полная именная группа, местоимение или нуль) связан с уровнем активированности референта в его кратковременной памяти. На уровень активированности влияют различные факторы (линейное и риторическое расстояния до antecedента, расстояние в абзацах, а также его семантико-синтаксический статус и внутренние свойства.), в сумме дающие коэффициент активации, при увеличении которого возрастает вероятность употребления редуцированного референциального выражения – местоимения или нуля.

Однако при этом становится возможным возникновение референциального конфликта - ситуации, когда референциальное выражение может быть отнесено более чем к одному референту. Эта неоднозначность может быть устранена при помощи различных средств, чаще всего лексико-грамматических [Кибрик 1987].

В работе рассматривается ситуация потенциального референциального конфликта, когда возможный конфликт компенсируется полом персонажа. Наши

эксперименты основаны на англоязычном исследовании [Arnold, Griffin 2007], в котором было описаны аналогичные эксперименты на порождение коротких историй по картинкам.

Эксперимент был проведен по методике рассказывания коротких историй по картинкам. В качестве стимулов были использованы картинки с легко узнаваемыми героями мультфильмов. Каждый из 15 сюжетов состоял из двух картинок, составляющих историю, которая начиналась на первой картинке и заканчивалась на второй. Сначала испытуемый в течение 2 секунд видел на экране две картинки, расположенные одна под другой, потом нижняя картинка исчезала, и он слышал предложение, которое описывало то, что происходило на первой картинке (напр., *Винни-Пух с Совой подошли к большому дереву*). Сразу после этого испытуемый должен был повторить только что услышанное предложение, а затем, вновь посмотрев на вторую картинку, закончить историю (например, *потом Винни-Пух подумал, что он хочет съесть мед*). Как и в англоязычном эксперименте, мы просили испытуемых говорить просто и ясно, чтобы было понятно и пятилетнему ребенку. По количеству персонажей каждый сюжет имел три варианта – на обеих картинках изображен: 1) один персонаж; 2) два персонажа разного пола; 3) на первой картинке изображены два персонажа разного пола, на второй – только один из них. Истории составлялись таким образом, чтобы независимо от количества персонажей и на первой, и на второй картинке главным действующим лицом был один и тот же персонаж.

Эксперимент был проведен с 24 испытуемыми. Как и в предшествующих работах, мы учитывали только те канонические ответы испытуемых, в которых главный персонаж был упомянут первым и являлся подлежащим.

кол-во персонажей на первой/второй картинке	канонические предложения, в %	тип референциального выражения, в %		
		нуль	местоимение	полная ИГ
1/1	90	34	37	29
2/2	72	1	2	97
2/1	78	0	1	99

Результаты эксперимента показали, что в первом случае, когда на обеих картинках был один персонаж, говорящий с равной вероятностью может использовать как синтаксический нуль и местоимение, так и полную именную группу (была проверена статистическая значимость результатов, ответы распределены одинаково, $p\text{-value} = 0.4169$). В двух других вариантах испытуемые практически всегда выбирали полную ИГ ($p\text{-value} < 0.01$). Отсутствие значимых различий между вторым и третьим типами ($p\text{-value}$ от 0.2364 до 0.7281) говорит о том, что при описании картинки испытуемый опирается прежде всего не на текущий визуальный ряд, где может быть как два персонажа, так и один, а на свою дискурсивную рабочую память, в которой содержится информация о двух активированных персонажах.

Результаты эксперимента не могут дать однозначный ответ на вопрос о том, какую стратегию использует говорящий, то есть, опирается ли на себя самого, или ориентируется на адресата. С одной стороны, согласно [Arnold, Griffin 2007] сам факт возможности референциального конфликта снижает уровень активированности, что приводит к необходимости употребления полной именной группы, по мнению исследователей, подобные результаты говорят об эгоцентрической стратегии. Однако в модели референциального выбора в работе [Kibrik 2011] действие фильтра референциального конфликта происходит после выбора референциального выражения и не влияет на коэффициент активации. Для продолжения исследования возможно проведение новых экспериментов с измененными условиями, которые помогут более

точно определить место и роль фильтра референциального конфликта в модели референциального выбора.

Литература

Кибрик А. А. Механизмы устранения референциального конфликта // Александр Е. Кибрик и Александр С. Нариньяни (ред.), Моделирование языковой деятельности в интеллектуальных системах. М., 1987. 128–45.

Федорова О.В. Основы экспериментальной психолингвистики: Принципы организации эксперимента. М., 2008.

Федорова О.В., Деликишикина Е.А., Малютина С.А., Успенская А.М. и Фейн А.А.. Экспериментальный подход к исследованию референции в дискурсе: интерпретация анафорического местоимения в зависимости от риторического расстояния до его антецедента // Александр Е. Кибрик (ред.), Компьютерная лингвистика и интеллектуальные технологии. Материалы международной конференции "Диалог 2010". 2010. 525–30.

Arnold J. E., Griffin Z. M.. The effect of additional characters on choice of referring expression: Everyone counts // Journal of Memory and Language. 2007. 56(4). 521–36.

Kibrik A.A. Reference in discourse. Oxford, 2011

Концептуализация явлений с диффузией референтной и предикатной составляющих, постановка проблемы и направления анализа

Филиппов Василий Олегович

Соискатель Московского государственного университета имени М. В. Ломоносова,
Москва, Россия

Пропозициональное содержание лежит в основе законченного высказывания, являясь его объективным компонентом (противопоставленным субъективному, прагматическому). Пропозиция обладает определенной внутренней структурой, в которой «выделяются два основных типа составляющих — термы, или имена, которые отображают сущности, объекты, являющиеся участниками ситуации, и предикаты, обозначающие свойства объектов и отношения между ними» [Кобозева 2007, с. 219]. «Термы, или имена», отображающие объекты, можно назвать референтными членами пропозиции или референтами (ср., например, понятия акта референции и акта предикации в [Сёрль 1986, с. 155, 156]). Референтно-предикатный дуализм, вероятнее всего, и является тем аналитическим механизмом, который позволяет нам запечатлевать переживаемые впечатления в словесной форме. На первый взгляд, критерии формирования референтной и предикатной частей высказывания могут показаться интуитивно понятными и непротиворечивыми. На самом же деле существуют целые группы явлений (назовем их диффузными), вербализация которых в разных языках обладает рядом специфических особенностей, указывающих на сложности с референтно-предикатной дифференциацией. Невозможность или затрудненность однозначного деления пропозиции на референтную и предикатную части, как особенность концептуализации ряда явлений из различных когнитивных сфер, мы называем *референтно-предикатной диффузией* (РПД). Диффузные ситуации противостоят таким, при концептуализации которых референтно-предикатная структура обнаруживается без каких-либо сложностей; назовем последние *эталонно-дифференцируемыми* (ЭД). ЭД-концептуализация дает высказывания типа R*P, где R, референтная часть, представляет собой предмет, а P — предикат, в наиболее простом, на наш взгляд, случае есть движение предмета P. Например: *Лист падает*. Референтная часть ЭД-высказывания является *фигурой*, это значит, что она легко и однозначно выделяется среди *фона*. Понятие фигуры коррелирует с понятием *фокуса внимания*, как «модели типа центр-периферия, в которой более интенсивное внимание приходится на центральную часть, а менее интенсивное — на окружающие ее части» [Талми 1999]. К предикату мы также предъявляем требования простоты, целостности и конкретности.

РПД-ситуация *Листопад* отличается от рассмотренной прежде всего количественно — ЭД-ситуация *Лист падает* повторяется в *Листопаде* множество раз (множественность — один из признаков РПД), а также по признаку занимаемого пространства (тоже признак РПД), в результате, становясь для нашего восприятия из фигуры — фоном.

Предлагаются следующие направления исследования РПД:

Типологизация РПД, изучение механизма концептуализации посредством признаков. Есть два, отчасти пересекающихся типа типологизирующих признаков: а) позволяющих провести границу между РПД и остальными ситуациями, б) ответственных за типологизацию внутри области РПД. Вот важнейшие из признаков: единичность ↔ множественность референта и предиката, размер референта, длительность предиката, дискретность ↔ континуальность (непрерывность) референта, гомогенность ↔ гетерогенность референта, когерентность предиката (принципиальное сходство частичных предикатов), сегментность (наличие функциональных блоков в референтной группе), реципрокность и т. д. Система признаков позволяет выделить внутри области группы явлений, обособленные от остальных по принципу сигнификативного сходства, то есть, характеризуются общим набором параметров.

Лексические функции Модели «Смысл — Текст» применительно к РПД. Исследуется возможность применения лексических функций МСТ [Мельчук 1999]. В области РПД особенно важны некоторые функции МСТ: Magn («очень»), Sing («квант, штука»), Func («осуществить ситуацию»), Incep («начать ситуацию»), Caus («каузировать») и другие. Вводятся специализированные функции, ориентированные на анализ РПД: например, Quant(ситуация), возвращает ситуативный квант, где $Quant(X) = QuantRef(X) * QuantPred(X)$, т. е. результат взаимодействия референтной и предикатной частей. Quant(Дождь) = Капля падает. Необходимость дифференциации референтной и предикатной частей квантовой ситуации обусловлена потребностью в более точном анализе РПД-ситуаций.

Концептуальная метафора и РПД. РПД, на наш взгляд, является такой областью концептуализации, на материале которой можно продуктивно изучать механизмы метафорической проекции. Диффузные явления сложны по своей природе, и их семантический анализ уже предполагает четкое их структурирование. Сравнивая схематическое представление о концептах разных диффузных явлений, можно представить себе характер метафорических связей, которыми оперирует наше сознание. Как правило, язык выбирает в сфере повседневного опыта некоторый хорошо обозначенный в сознании носителей языка фрейм, обладающий ясной структурой, и делает с него «концептуальный слепок» нужной степени абстракции — мы называем это суперкатегоризацией концепта-донора, далее проецируя его на другие области, содержание которых может быть упорядочено при помощи донорных схем.

Литература

- Кобозева И.М. Лингвистическая семантика. // М., 2007.
 Мельчук И.А. Опыт теории лингвистических моделей «Смысл ↔ Текст». // М., 1999
 Серль Дж.Р. Что такое речевой акт? // Новое в зарубежной лингвистике, вып. XVII. Теория речевых актов. М., 1986. С.151–169.
 Талми Л. Отношение грамматики к познанию. // Вестник МГУ. Серия 9. Филология. 1999.

Метаязык описания предикатов типа *to have a + конверсив*: классификационные параметры

Холодова Дарья Дмитриевна

Студентка Московского городского педагогического университета, Москва, Россия

В докладе рассматриваются некоторые аспекты метаязыка описания предикатов типа *to have a + конверсив*, необходимые для выявления различий семантики, а следовательно, особенностей функционирования этой модели. Исследуемые языковые

единицы представляют собой одну из наиболее продуктивных моделей в современном английском языке. При этом отсутствие исчерпывающего семантического описания исследуемой конструкции, а также отсутствие исследований синтаксического потенциала анализируемой модели определяют актуальность настоящего исследования.

Исследование семантики языковых единиц принято начинать с разработки соответствующего метаязыка описания. Поскольку в докладе речь идет о такой сложной конструкции как *to have a + конверсив*, представляется целесообразным ввести ряд классификационных параметров, которые в дальнейшем могут быть заложены в основу семантического эксперимента.

Один из важнейших семантических признаков, свойственных изучаемой модели, вносит информацию о том, что осуществляемое действие представляет собой «бесперспективную» деятельность.

Предикаты «бесперспективного протекания» характеризуют действие без указания на его окончание; действие, не направленное в своем протекании к определенному сущностью самого события пределу и, следовательно, означаемое действие не может исчерпать себя и затем прекратиться. При этом данное действие может перестать иметь место, однако его кульминация будет связана с некоторыми событиями, происхождение которых не предусматривается семантикой описывающего это действие предиката. Ср.: *Let's have a sing (Давай пооем)*, – где действие представляется как такое, которое вскоре будет завершено, при этом временная граница, по достижении которой действие должно прекратиться, не зависит от характера самого действия и обуславливается внешними по отношению к нему факторами (например, порцией или волей субъекта). Следует отметить, что субъект действия сознательно прикладывает усилия, чтобы реализовать это действие, и в этом смысле оно будет иметь цель. Однако цель этого действия заключается не в завершении процесса, а в самом процессе: *X had a read*, – каким бы коротким ни было время выполнения действия, субъект все же почитал немного. Иными словами, действие это начнет реализовываться уже в тот момент, когда оно будет инициировано субъектом, и достигнет своего «завершения» (в том смысле, что X позанимается этим), когда часть описываемого процесса будет уже завершена.

Второй семантический признак, характеризующий анализируемую модель, вносит информацию о том, что осуществляемое действие представляет собой «контролируемую» деятельность.

Под «контролируемостью» понимается приложение субъектом волевых усилий в виде сознательного управления начинанием, продолжением, качественным изменением, прекращением и завершением некоторого события или действия [Сулейманова 2009]. Исходя из этого, субъект рассматривается как осознанно прикладывающий усилия для инициирования, поддержания, изменения и завершения некоторого действия. Отсутствие проявления волеизъявления субъекта при управлении действиями свидетельствует об отсутствии в семантике описываемого события признака контролируемости.

Так, в предложении *Can I just have a smell? – (Можно мне просто понюхать?)* сообщается о приложении субъектом усилий для претворения события (обоняние) в жизнь; а в предложении *The seeds of most of the plants have a smell of vanilla (Семена большинства растений пахнут ванилью)* сообщается о том, что событие совершается независимо от приложения субъектом усилий, и X, таким образом, не представлен как источник энергии, необходимой для существования описываемого события. Как показано Т.Д. Шабановой [Шабанова 1998, 49], любое событие может быть представлено в виде последовательно сменяющихся друг друга фаз, при этом фактор контролируемости может вступать в силу на любом из этапов. Как представляется, события, описываемые конструкцией типа *to have a + конверсив*, являются контролируемыми на всех стадиях своей реализации. Ср. *X decided to have a walk and in*

an hour got back to work (*X погулял и вернулся к работе*), где субъект по собственному желанию инициирует и завершает описываемую деятельность.

Для проверки гипотезы о том, что обозначенные выше признаки реализуются применительно к конструкции *to have a + конверсив*, планируется провести семантический эксперимент, в рамках которого информантам – носителям языка будет предъявлена экспериментальная выборка примеров: идентичных текстов с единственным различием в выборе предиката для оценки правильности/неправильности этих высказываний.

Так, например, для выявления в семантике предиката признака «бесперспективности» действия информантам могут быть предложены следующие примеры:

Have a read through them at your leisure (*почитай их на досуге*), где, согласно гипотезе, X-у предлагают попробовать почитать книги, при этом не подразумевая, что все материалы должны быть прочитаны полностью;

Read through them at your leisure (*прочитай их на досуге*), где X-а просят полностью изучить все предложенные ему материалы.

Результаты эксперимента могут быть далее подвергнуты статистическому и интерпретационному анализу. В том случае, если наличие разницы в значениях описываемых моделей будет отчетливо осознаваться носителями языка, гипотеза получит свое подтверждение.

Литература

Сулейманова О.А. В развитие метаязыка лингвистического описания: абстрактность и контролируемость. М., 2009.

Шабанова Т.Д. Семантическая модель англ. Глаголов зрения: Теоретико-экспериментальное исследование. Уфа, 1998.

Методы автоматического транскрибирования (краткий обзор)

Широкова Анна Михайловна

Соискатель Московского государственного университета имени М. В. Ломоносова,
Москва, Россия

Системы автоматической обработки речи с открытым словарем строят модели распознаваемых слов, используя более мелкие единицы, в общем случае равные фонемам. Часто для перевода входной орфографической последовательности в фонематическую форму такие системы используют произносительный словарь, включающий слова с соответствующими транскрипциями. Однако, чтобы обеспечить покрытие слов, новых для системы, необходимо иметь возможность автоматически, предпочтительно в реальном времени, преобразовывать орфографический текст в фонемную запись. Здесь рассматриваются два принципиальных подхода к решению данной задачи: традиционный, основанный на лингвистических правилах, и альтернативный, предполагающий самообучение системы на основе некоторого словаря.

Традиционные системы, основанные на правилах

Формализм контекстно зависимых правил, получивший популярность в рамках создания первых систем синтеза речи, базируется на положении, что достаточное знание и адекватное описание контекста позволяет предсказать верную фонемную реализацию. Общий вид этих правил обусловлен сильным влиянием идей порождающей фонологии Хомского и Халле:

$$A[B]C \rightarrow D$$

где буквенное подвыражение В при левом контексте А и правом контексте С получает транскрипцию D.

Методы, ориентированные на данные, управляемые данными

Объединяющей основой рассматриваемых научно-прикладных подходов к автоматическому преобразованию «графема-фонема» является *ориентация на данные* (в англоязычной терминологии *data-oriented methods*), *управляемость данными* (*data-driven methods*). Это предполагает самообучение системы на сравнительно большом корпусе данных, который в общем случае представляет собой произносительный словарь, содержащий списки слов и сопоставленных им транскрипций.

Для решения данной проблемы применяются различные техники машинного обучения (*machine learning*). В некоторых системах необходимо предварительное соотнесение графем и наиболее вероятных для них фонемных реализаций, как в [Pagel и др.]; другие же, например [Bisani, Ney], наоборот, успешно функционируют без использования начальных экспертных знаний и последующей «ручной» коррекции на том или ином этапе работы.

Техники, основанные на локальной классификации

В достаточно большой группе методов G2P обязательным предварительным условием или отдельным шагом предобработки является построение *ряда сопоставлений* (*alignment*) графем и фонем. Сопоставительный ряд типично организован так, что каждый из его членов представляет собой одну букву; число соотнесенных с ней фонем может быть разным: от нуля до нескольких. Предсказание выходной фонемы (группы фонем) строится на базе знания контекста текущей буквы. Поскольку решение по каждой позиции принимается до перехода к следующей, удобно именовать это семейство техник *локальной классификацией*. Наиболее известные из подобных методов – нейронные сети и деревья решений.

Транскрипция по аналогии

Общая идея подхода состоит в том, что в процессе работы происходит сканирование обучающего словаря с целью нахождения в нем слов или частей слов, которые в определенной степени схожи со словом, транскрипцию которого нужно получить. Фонемное представление неизвестного слова неизбежно формируется по аналогии с примерами, на основе которых проводится обучение. Транскрипция по аналогии представляет собой более сложную стратегию, превосходящую локальную классификацию благодаря рассмотрению слова как некоего целого.

Вероятностный подход

Некоторые исследователи рассматривают задачу буквенно-звукового преобразования с точки зрения теории вероятностей.

В работе [Bisani, Ney 2002] рассматриваются модели соединенных последовательностей (*joint-sequence models*). Как отмечают авторы, эта техника полностью ориентирована на данные и не зависит от обрабатываемого языка.

Главная идея данного метода состоит в том, что соотношение входного текста и транскрипции может быть выведено на основе общей последовательности соединенных единиц, содержащих одновременно орфографические (входные) и транскрипционные (выходные) символы. Соединенная графемно-фонемная мультиграмма, иначе графон [Bisani, Ney 2002: 1], определяется как пара буквенной и звуковой последовательности потенциально различной длительности. Инвентарь графонов может быть выведен автоматически из обучающих данных или может быть уточнен вручную. Предполагается, что в модели соединенных мультиграмм орфографическая форма слова и его транскрипция порождаются общей последовательностью графонов. Например, слово «mixing» может быть рассмотрено как последовательность четырех графонов:

$$\begin{array}{l} \text{"mixing"} \\ [mɪksɪŋ] \end{array} = \begin{array}{|c|} \hline m \\ \hline [m] \\ \hline \end{array} \begin{array}{|c|} \hline i \\ \hline [i] \\ \hline \end{array} \begin{array}{|c|} \hline x \\ \hline [ks] \\ \hline \end{array} \begin{array}{|c|} \hline ing \\ \hline [ɪŋ] \\ \hline \end{array}$$

Для создания графонов буквенные и звуковые последовательности группируются по одинаковому числу сегментов. Для некоторой пары входной и выходной цепочек может быть создано несколько вариантов сегментации на графоны.

Обучающие произносительные словари

Необходимо отметить, что принципиальным свойством словарей для обучения систем G2P является их репрезентативность с точки зрения иллюстрации правил произношения данного языка. Для обучения G2P моделей используются доступные произносительные словари, покрывающие наиболее частотную лексику анализируемого языка. Несмотря на наличие слов-исключений, более вероятные, регулярные образцы буквенно-звуковых соответствий, несомненно, преобладают.

В настоящее время программное средство, основанное на методе соединенных последовательностей, тестируется на материале русского языка, предполагается, что полученные результаты будут представлены для обсуждения.

Литература

Pagel, V., Lenzo, K., Black, A.W. Letter-to-sound rules for accented lexicon compression. In: Proc. Internat. Conf. on Spoken Language Processing, Vol. V, Sydney, Australia, 1998.

Bisani, M., Ney, H. Investigations on joint-multigram models for grapheme-to-phoneme conversion. In: Proc. Internat. Conf. on Spoken Language Processing, Vol. 1, Denver, CO, USA, 2002.

О некоторых принципах разработки признаковой базы для типологической базы данных по количественным конструкциям (NNC-Database)

Шматова Мария Сергеевна

*Аспирантка Московского государственного университета имени М. В. Ломоносова,
Москва*

В работе обсуждаются некоторые проблемы, возникающие при разработке признаковой базы для типологической базы данных (далее: БД) по количественным конструкциям в языках мира. Под количественными конструкциями (англ. NNC или Numeral-Noun Constructions) понимаются конструкции, состоящие из существительного и количественного числительного и имеющие количественное значение.

Разрабатываемая БД призвана стать, прежде всего, удобным инструментом для обобщения данных по количественным конструкциям в языках различных семей. Планируется, что на завершающей стадии работы над проектом будет создана онлайн версия БД, получив доступ к которой, каждый пользователь сможет внести данные исследуемого им языка, а также ознакомиться с данными других языков для получения актуальной типологической картины. В связи с этим в ходе разработки также делается упор на теоретическую нейтральность БД.

Работа над БД такого рода обычно включает в себя несколько этапов. В данной работе будет рассмотрен первый этап – разработка признаковой базы, за которой последует подбор программного обеспечения для ее реализации, а также разработка пользовательского интерфейса.

Надо признать, что полное описание количественных конструкций встречается далеко не так часто, как хотелось бы. Во многих грамматических очерках дается лишь несколько примеров количественных конструкций в разделе об употреблении количественных числительных. В литературе, специально посвященной вопросу количественных конструкций в одном или нескольких языках, рассматриваются, как правило, только отдельные аспекты (например, падежное маркирование в русском языке, вопрос о характере синтаксической связи внутри количественной конструкции в языках мира). В связи с этим возникает потребность в ресурсе, который содержал бы как можно более полную характеристику описываемых конструкций. Поэтому нашей задачей на первом этапе работы явилось создание максимально полной признаковой базы.

Специфика разработки признаковой базы состоит, во-первых, в том, что для получения полной картины сочетаемости количественных числительных с существительными и синтаксического поведения конструкций необходимо

рассматривать признаки всех членов конструкции на всех уровнях языка. Во-вторых, следует помнить, что в языке возможны «исключения из правил», поэтому в рамках работы над БД для большинства параметров следует предусмотреть не только базовое, но и другие возможные значения. В-третьих, важно сопроводить каждый параметр окном комментариев и примерами, как это делается в проекте типологической БД по синтаксическим ограничениям [Сердобольская и др.: 440].

Ниже кратко рассмотрим основные предлагаемые параметры.

Одним из наиболее полно описанных к настоящему моменту параметров является порядок следования составляющих внутри конструкции. Согласно данным «Всемирного атласа языковых структур» (WALS), где по описываемому параметру рассмотрено 1002 языка, существительное может как предшествовать числительному (что наблюдается в большинстве языков), так и следовать за ним. Причем порой оба варианта могут иметь место в рамках одного языка (например, в баскском, абхазском языках) [Dryer]. Именно для таких случаев в описываемой БД предусмотрена возможность зафиксировать «исключения из правил».

Далее, не всегда состав количественной конструкции ограничивается числительным и существительным. Во многих языках (например, в австронезийских языках и языках семьи майя) при существительном в составе количественной конструкции обязательно употребление «специальных синтаксических лексем (реже морфем)» [Плунгян: 159] – классификаторов. Поэтому в рамках настоящей БД предусмотрен параметр «наличие/отсутствие классификатора».

Следующую группу параметров можно условно назвать «морфосинтаксическим составом» конструкции. Сюда включаются такие признаки, как словоизменительные характеристики всех членов конструкции, тип синтаксической связи между членами конструкции. К словоизменительным характеристикам в разных языках могут относиться категории числа, падежа, посессивности. Поскольку БД рассчитана на языки самых различных типов, исследователю будет предоставлена возможность при вводе собственных данных указать каждую из этих характеристик.

Во многих языках существуют семантические ограничения на употребление числительных с некоторыми существительными (ср. вещественные существительные в русском языке), что также планируется отразить в описываемой БД.

Выше были рассмотрены параметры, так или иначе относящиеся к внутренней структуре количественных конструкций. Последняя группа параметров относится уже к уровню предложения. Здесь стоит обратить внимание как на те предложения, где рассматриваемая конструкция находится в позиции субъекта, так и на те, где она занимает место прямого и непрямого объекта. В первом случае в качестве параметра будут рассмотрены синтаксические отношения «субъект – предикат», а во втором – «предикат – объект» соответственно. На этапе реализации именно этой группы параметров сложнее всего соблюдать принцип теоретической нейтральности, постулированный выше, так как именно с синтаксической точки зрения рассматриваемые конструкции являются предметом наибольших споров среди исследователей разных языков.

Литература

Плунгян В.А. Общая морфология. М., 2003.

Сердобольская Н.В., Циммерлинг А.В., Аркадьев П.М. Проект типологической базы данных по синтаксическим ограничениям на передвижения // Компьютерная лингвистика и интеллектуальные технологии: По материалам ежегодной Международной конференции «Диалог» (2010). Вып. 9. С.437-441

Dryer M. Order of numeral and noun // Haspelmath, Martin & Dryer, Matthew S. & Gil, David & Comrie, Bernard (eds.) The World Atlas of Language Structures Online: <http://wals.info/feature/89>

Определение оптимальных параметров контекстного окна для автоматического разрешения многозначности и выделения конструкций с опорой на многоуровневую разметку корпуса (на материале НКРЯ)*

Шурыгина Александра Сергеевна
Студентка Санкт-Петербургского государственного университета,
Санкт-Петербург, Россия

Работа выполнена
при финансовой поддержке РФФИ
(в рамках проекта 10-06-00586-а).

Исследование посвящено определению роли ширины контекстного окна в разрешении лексико-семантической неоднозначности русских существительных. Эксперименты проводятся на материале выборок контекстов из НКРЯ (<http://ruscorpora.ru/>) с лексико-семантической разметкой (<http://ruscorpora.ru/corpora-sem.html>).

Пристальное внимание уделяется контекстным маркерам, отраженным в лексико-семантической разметке контекстов НКРЯ.

Результаты экспериментов иллюстрируются на материале существительных *вид* и *глава*, характеризующихся достаточно высокой частотностью и многообразием значений, среди которых встречаются как конкретные, так и абстрактные. Множество значений данных слов отражено в представительной выборке из НКРЯ объемами 2866 и 1073 контекстов. Каждому из значений слов *вид* и *глава* соответствует набор тегов, принятых в системе разметки НКРЯ (см. напр. табл. 1).

Таблица 1. Филиация значений слова *вид*

Значения	Лексико-семантическая аннотация	Примеры	Число контекстов ($\Sigma=2866$)
<i>m1</i>	r:abstr t:perc der:v	Вид на озеро	1144
<i>m3</i>	r:concr t:workart	Альбом с видами Кавказа	10
<i>m5</i>	r:abstr t:ment	Виды на урожай	10
<i>m11</i>	r:abstr der:shift	Подсолнухи в виде букета	1075
<i>m12</i>	r:concr t:doc	Вид на жительство	7
<i>m21</i>	r:abstr r:concr pt:set sc:X	Отряды и виды животных	617
<i>m22</i>	r:abstr	Вид глагола	3

В экспериментах используется компьютерный инструмент, позволяющий создать векторную модель экспериментальной выборки. Контекстные вектора формируются для каждого из значений и служат образом соответствующего значения. Координаты вектора определяются частотами встречаемости лексических маркеров значения (lex), лексико-семантических (sem) или морфологических тегов (gr) контекстных элементов в эталонной выборке. Также предоставляются дополнительные статистические данные, в том числе о точности и полноте результатов тестов.

В ходе оценки оптимальных критериев автоматического разрешения лексико-семантической неоднозначности слов в контекстах НКРЯ было доказано, что наилучшие результаты достигаются при выборе в качестве контекстных маркеров значений тегов леммы (lex), комбинации тегов леммы и лексико-семантических тегов (lex + sem) или комбинации всех трех тегов (lex + sem + gr). Поэтому эксперименты проводились только с учетом этих комбинаций тегов.

Согласно условиям экспериментов, не рассматривались тривиальные параметры контекстного окна (то есть те варианты, при которых контекстное окно слева или справа было равно нулю).

Анализируя данные, полученные в экспериментах для слов *глава* и *вид* и систематизированные в таблицах и гистограммах, мы пришли к следующим выводам:

- 1) наилучшее разрешение неоднозначности по контексту происходит с учетом комбинации всех трех тегов (lex + sem + gr) — 87-89%.
- 2) эксперименты по разрешению неоднозначности с учетом комбинации тегов леммы и лексико-семантических тегов (lex + sem) показывают хорошие результаты, но более низкие, чем с учетом комбинации всех трех тегов — 84-87%.
- 3) в экспериментах с учетом комбинации всех трех тегов практически нет ошибок на полноту (в нашем исследовании ошибки на полноту — это те случаи, в которых программа не смогла принять никакого решения). Однако ошибок на точность меньше в экспериментах с учетом комбинации тегов леммы и лексико-семантических тегов (lex + sem). (В нашем исследовании ошибки на точность — это те случаи, в которых программа приняла неправильное решение).
- 4) наилучшее контекстное окно с учетом комбинации тегов (lex + sem) — [-4; +3], а с учетом комбинации всех трех тегов (lex + dem + gr) — [-2; +3] и [-4; +3]. Хорошие результаты были также в экспериментах с контекстными окнами [-2; +4], [-3; +3] и др.
- 5) правосторонние коллокаты более информативны для идентификации значений слова *глава*, чем левосторонние, в то время как для слова *вид* левосторонние коллокаты также информативны.
- 6) присутствие ряда тегов в контексте устойчивым образом связано с проявлением у слов определенного значения. Например для слова *вид*, теги *t:humq* («качества человека»: *озабоченный, неприглядный, задумчивый* и т.д.), *t:psych:emot* («эмоциональное состояние»: *равнодушный, довольный, удивленный* и т.д.) и *t:hum* («лица»: *знаток, победитель*) указывают на значение *m1*, а теги *pt:set, pt:aggr* («множества, совокупности»: *транспорт, вооружение, оружие* и т.д.) и *hi:class* («имена классов»: *оружие, растение, животное* и т.д.) — на значение *m3*.

Восприятие семантически аномальных предложений русского языка в норме и при афазии

Юрченко Анна Николаевна

*Студентка Московского государственного университета имени М. В. Ломоносова,
Москва, Россия*

Способность понимать значения отдельных лексем и интегрировать их в более широкий контекст является одним из необходимых условий нормального использования языка. Встает вопрос о том, какие механизмы лежат в основе семантического компонента языковой обработки и каковы его нейрофизиологические корреляты.

Применение метода вызванных потенциалов, основанного на измерении биоэлектрической активности головного мозга, позволяет анализировать процессы, обеспечивающие отдельные этапы языковой обработки. Сопоставление вызванных потенциалов, сопровождающих языковое восприятие у пациентов с различными типами языковых расстройств (афазии) и здоровых носителей языка позволяет изучать процессы, лежащие в основе того или иного компонента функционирования языка. Однако, обеспечивая высокое временное разрешение, метод вызванных потенциалов не дает возможности локализовать источник мозговой активности. Расширить знания о зонах мозга, вовлеченных в процессы языкового восприятия, позволяет анализ вызванных потенциалов, характеризующих языковую обработку у пациентов, противопоставленных по локализации поражения в мозге.

Целью настоящего исследования было проанализировать процессы, обеспечивающие семантическую обработку языка. Для этого впервые на материале русского языка было проведено сравнение вызванных потенциалов, характеризующих восприятие семантически правильных и семантически аномальных предложений у здоровых носителей языка и у пациентов с афазией, сгруппированных по диагнозу и локализации поражения мозга. В эксперименте приняли участие 8 здоровых испытуемых, 8 пациентов с небеглой афазией и 8 пациентов с беглой афазией.

Материал эксперимента включал 40 семантически правильных предложений и 40 предложений, содержащих семантические аномалии (пример (1)). Кроме того, в эксперимент вошли 160 отвлекающих предложений (80 правильных и 80 предложений, содержащих синтаксические и морфосинтаксические аномалии).

1. Малыш набирает **песок** / * **звонок** в ведёрко.

Предложения предъявлялись на слух. Испытуемых просили внимательно слушать предложения, и после того, как предъявление стимула заканчивалось, нажимать на правую кнопку двукнопочного манипулятора, если они считали услышанное предложение правильным, или на левую, если они считали услышанное предложение неправильным. Запись ЭЭГ-сигнала производилась с помощью 128 электродов, закрепленных в эластичной сетке Geodesic Sensor Net (Electrical Geodesics Inc.).

Статистический анализ данных о вызванных потенциалах осуществлялся с применением метода дисперсионного анализа с повторными измерениями ANOVA. В первую очередь, результаты эксперимента характеризуют восприятие аномальных предложений у здоровых носителей русского языка: восприятие семантически аномальных предложений по сравнению с правильными предложениями характеризуется эффектом N400 (негативность, достигающая максимума примерно через 400 мсек после предъявления стимула) – стандартным маркером лексико-семантической интеграции [Kotz, Friederici: 44-45].

Что касается противопоставления по типу афазий, восприятие семантически аномальных предложений сопровождается у пациентов с небеглой афазией, как и в группе нормы, эффектом N400. Однако этого эффекта не было обнаружено у пациентов с беглой афазией. Восприятие семантически аномальных предложений у этой группы пациентов характеризовалось эффектом P600 (позитивность, достигающая максимума примерно через 900 мсек после предъявления стимула), возникновение которого связывают, в том числе, с процессами интеграции разноуровневой лингвистической информации [Friederici, Weissenborn: 54].

Наконец, в результате анализа восприятия семантически аномальных предложений у групп пациентов, противопоставленных по локализации поражения, нормативный эффект N400 был обнаружен у пациентов с зоной поражения, ограниченной теменной долей левого полушария. У пациентов с поражением, затрагивающим также лобную и височную доли, эффекта N400 для семантически аномальных предложений не наблюдалось. Восприятие предложений с семантическими аномалиями у этой группы пациентов характеризовалось эффектом P600.

Таким образом, в проведенном исследовании впервые показано, что у здоровых носителей русского языка восприятие предложений, содержащих семантические аномалии, по сравнению с правильными предложениями сопровождается эффектом N400. Такой же эффект был обнаружен у пациентов с небеглой афазией, что свидетельствует об относительной сохранности у них процессов, обеспечивающих семантическую обработку языка. Однако при попытке интеграции семантической информации у пациентов с беглой афазией, вероятно, задействуются компенсаторные механизмы, в результате чего наблюдается увеличение потенциала P600. В соответствии с результатами анализа вызванных потенциалов у пациентов, различающихся по

локализации, теменная доля не играет ключевой роли в генерации потенциала N400, тогда как поражение лобной и височной долей ведет к нарушению механизмов, обеспечивающих процесс семантической интеграции.

Литература

Friederici A.D., Weissenborn J. Mapping sentence form onto meaning: The syntax-semantic interface // *Brain Research*. 2007. № 1146. С. 50-58.

Kotz S.A., Friederici A.D. Electrophysiology of normal and pathological language processing // *Journal of Neurolinguistics*. 2003. № 16. С. 43-58.