

КОГНИТИВНАЯ НАУКА

В МОСКВЕ



НОВЫЕ ИССЛЕДОВАНИЯ

МАТЕРИАЛЫ
КОНФЕРЕНЦИИ
2019

Под ред. Е.В. Печенковой, М.В. Фаликман

УДК 159.9
ББК 88.25
К57

Когнитивная наука в Москве: новые исследования. Материалы конференции 19 июня 2019 г. Под ред. Е. В. Печенковой, М. В. Фаликман. – М.: ООО «Буки Веди», ИППиП. 2019 г. – 656 стр.

ISBN 978-5-4465-2346-7

УДК 159.9
ББК 88.25

ISBN 978-5-4465-2346-7

©Авторы статей, 2019

ВЕРБАЛЬНЫЕ СРЕДСТВА РЕПРЕЗЕНТАЦИИ СЕМАНТИКИ ПОНЯТИЙ – ТОЧКИ ОПОРЫ СМЫСЛОВ ГОВОРЯЩЕГО И СЛУШАЮЩЕГО В МЕНТАЛЬНОМ ЛЕКСИКОНЕ

О. И. Пилатова

oiipilatova@yandex.ru

СПбГУ, Санкт-Петербург

Аннотация. Целью данного исследования являлось выяснение того, каким образом при восприятии устной естественной речи может быть обеспечен доступ к семантической информации, релевантной для успешного прогнозирования и распознавания определенных лексических единиц. Для получения сведений о том, какими вербальными средствами представлены существенные (семантические) с точки зрения носителей языка признаки десяти использованных понятий (выраженных существительными: *собака, книга, дверь* и т.д.), была использована методика толкования. Далее было проведено корпусное исследование для выявления представленности таких лексем-признаков в контекстном окружении этих понятий при естественной коммуникации. Третья часть исследования заключалась в оценке предсказательной силы контекстов в отношении целевых слов в зависимости от представленности в них полученных лексем-признаков (семантических ассоциаций) соответствующих понятий. В результате была установлена взаимосвязь между вербальными репрезентациями семантики понятия, ассоциациями на соответствующее слово и его лексическим окружением при употреблении, способствующим успешности прогнозирования его появления в речи. Выявленные элементы ментального лексикона (слова определенных словообразовательных гнезд) могут быть рассмотрены как составляющие конвенциональную основу и семантических репрезентаций использованных в исследовании понятий: являются вербальными признаками их семантических признаков, используемыми при речевой деятельности как говорящим, так и слушающим.

Ключевые слова: лексическая семантика, ментальный лексикон, восприятие устной речи, вероятностное прогнозирование, вербальные ассоциации, русские существительные, словообразовательное гнездо, семантические признаки

Введение

Семантическая информация оказывает значительное влияние на успешность и скорость распознавания речевого акустического сигнала (Vrouwer et al., 2013; Риехакайнен, 2016); прояснение того, как именно обеспечивается такая семантическая поддержка, является одной из актуальных задач моделирования процессов речевой деятельности. Одним из традиционных способов описания лексической семантики является компонентный (или логический)

анализ, основанный на представлении о содержании понятия как совокупности более простых понятий или сем. При этом семантическими признаются, как правило, не все, а только существенные компоненты значения. Предлагаются самые разнообразные наборы таких компонентов, а также процедуры их выявления и определения значимости, преимущественно основанные на языковой интуиции и экстралингвистических знаниях исследователя (Найда, 1983; Вежбицкая, 1996). Попытками преодоления субъективности результатов компонентного анализа являются, например, учет грамматических свойств соответствующих лексем или использование корпусов (наблюдение за их функционированием) (Апресян, 1995; Методы когнитивного анализа..., 2015). Помимо того, что интерпретация наблюдаемых фактов оказывается не менее субъективной, более важным представляется отметить то, что таким образом обнаруживается широкая вариативность возможных (и логически обоснованных) семантических описаний одних и тех же понятий, демонстрирующая условность разделения существенного/несущественного, языкового/речевого, лингвистического/экстралингвистического в значении (Лич, 1983; Кронгауз, 2001; Алефиренко, 2005; Ирисханова, 2014). Логично предположить, что не только набор и значимость тех или иных компонентов, но и форма их представления могут различаться при реализации разных задач. Для осуществления речевой деятельности структура и содержание ментального лексикона, являющегося динамическим образованием, представляют ключевой интерес; имеющие основания разделять «словарь говорящего» и «словарь слушающего» (Касевич, Ягунова, 2006). Поскольку порождающий и воспринимающий речь существует в одном сознании, а материалом устной коммуникации является содержимое ментального лексикона, задача исследования может быть сформулирована следующим образом: выявление взаимосвязи между вербализуемой частью внеконтекстной семантики понятия и его лексическим окружением в речи, способным обеспечивать успешное прогнозирование лексической единицы, соответствующей этому понятию. Цель данного исследования – выявление таких вербальных признаков содержания понятий, которые релевантны как для процесса порождения, так и для восприятия речи.

Еще одно направление в изучении лексической семантики – это исследование значения слова через анализ его связей с другими понятиями естественного языка: как ассоциативных, так и контекстных. Результаты применения дистрибутивного подхода подтверждают идеи о наличии взаимозависимости значения слова и его контекстного окружения (Turney, Pantel, 2010), которая проявляется в том числе и в его вербальных ассоциациях. С другой стороны, с помощью данных методов невозможно объективно выделить (среди ассоциаций и в содержимом контекстов), какие конкретно лексические единицы обеспечивают эту связь именно на семантическом уровне: известно, что и вербальные ассоциации могут быть обусловлены различными факторами, не только семантическими (Lucas, 2000; Залевская, 2011), а контексты содержат также ассоциативные несемантические, коллокационные и другие признаки, которые могут быть свойственны семантически близким понятиям. Кроме того, даже принимая широкое понимание семантики, сведения о контекстной схо-

жести близких по значению слов не дают информацию о связи этих контекстов с содержимым ментального лексикона.

Стимульный материал

Были использованы десять существительных русского языка (целевые слова, далее — ЦС): *собака, ножницы, человек, книга, ребенок, кровать, деньги, рубашка, купальник, дверь*. Поскольку планируется использовать полученные данные в функциональной модели восприятия речи, выбирались те конкретные существительные, которые присутствуют в Корпусе транскрибированных русских устных текстов (<http://narusco.ru>), разрабатываемом для этой модели (Венцов и др., 2013). Кроме того, так как основной целью исследования было не только и не столько обнаружение конкретных лексических средств репрезентации семантики того или иного понятия, сколько выяснение их представленности при решении различных задач, связанных с речевой деятельностью, были выбраны существительные разных семантических категорий (в том числе попарно).

Методика толкования понятий

Методика толкования понятий применяется, как правило, при изучении семантической памяти. При этом в качестве недостатка методики отмечается, что таким способом выявлены могут быть только те признаки, которые легко (или в принципе) вербализуются (McRae et al., 2005). При исследовании же лексической семантики данное ограничение представляется, наоборот, одним из достоинств методики. Участникам (46 женщин и 44 мужчины) давалось задание записать существенные признаки каждого из 10 ЦС. Инструкция содержала подробное перечисление разнообразных типов признаков (физические характеристики, внешние и внутренние части, функциональные свойства и многое другое). Слова предьявлялись в случайном порядке. Каждый признак необходимо было записывать с новой строки; в среднем каждый из участников записывал около 7–10 признаков для каждого понятия.

Результаты. Наблюдалась зависимость формы ответов от выбранной тем или иным участником стратегии: перечисление признаков через определения, через глаголы действий, начальные формы слов и т.д. При этом ответы разных информантов преимущественно содержали слова с повторяющимися корневыми морфемами. То есть носители языка склонны не только перечислять похожие признаки для того или иного понятия, но и использовать для этого одинаковые слова (и/или слова одних и тех же словообразовательных гнезд), в том числе в тех случаях, когда ответы в целом могли указывать на разные типы признаков, быть нетипичными или носить личностный характер (напр., признак *очки для чтения* единичный среди ответов, но признак *читать/чтение* на втором месте по частотности для ЦС *книга*). Полученные списки признаков были сгруппированы соответственно упомянутому в них лексему одного и того же словообразовательного гнезда и ранжированы по частотности. Сопоставление списков лексем-признаков с данными Русского ассоциативного словаря (РАС, 2002) показало, что практически все слова, которыми передают-

ся семантические признаки понятия, присутствуют среди ассоциативных вербальных реакций на соответствующее слово. Принимая во внимание отличие задания ассоциативного эксперимента от толкования (где выбор языковых средств является имплицитной задачей), полученные лексемы-признаки могут рассматриваться как вербальные ассоциации, обусловленные семантикой понятия (то есть семантическими ассоциациями, отграничение которых от прочих типов вербальных ассоциаций является отдельной проблемой).

Корпусное исследование

Корпусное исследование на материале расшифровок устной речи Национального корпуса русского языка (<http://www.ruscorpora.ru>) заключалось в анализе ближайших контекстных окружений ($\approx/\pm 20$ слов) ЦС для проверки факта наличия в них тех лексем, с помощью которых носители языка передают значение этих слов.

Результаты. Было установлено, что в ближайших контекстах ЦС (за исключением низкочастотных и встречающихся в однообразных контекстах в анализируемом подкорпусе ЦС *купальник* и *ножницы*) обнаруживаются все их лексемы-признаки. Таким образом, наблюдаемая взаимосвязь между лексемами, используемыми для вербализации семантики понятий, и их представленностью в контекстах употребления соответствующих слов может служить средством доступа к значению этих понятий, способствующим успешному прогнозу их появления в речи.

Методика заполнения пропуска во фразе

Для проверки этого предположения была применена методика заполнения пропуска во фразе: были использованы контексты (из устного подкорпуса НКРЯ), в которых было пропущено одно из ЦС. Информантам (190 женщин и 50 мужчин) предлагалось заполнить пропуск наиболее подходящим словом. Во избежание влияния на ответы дополнительной информации, которая доступна при чтении текста, контексты предъявлялись в записи без знаков пунктуации. Контексты были модифицированы так, чтобы выполнялось шесть разных условий: лексемы-признаки присутствуют как в правом, так и в левом контексте (уу); только в левом контексте (уп); только в правом контексте (пу); в левом контексте при отсутствии правого контекста (у); отсутствуют в контексте (пп); отсутствуют в контексте при отсутствии правого контекста (п). Каждый участник получал анкету, содержащую 20 фраз с пропущенными словами: 10 – с целевыми и 10 контекстов-филлеров. Для статистической обработки был использован критерий хи-квадрат Пирсона.

Результаты. Выявлено значимое различие ($p < .001$) распределений верных и неверных ответов в условиях наличия лексем-признаков (Y = условия уу, уп, пу, у) и их отсутствия (N = пп, п). Сравнение по этому же параметру ответов для контекстов с отсутствующей правой частью (условия у и п) показало также значимое различие ($p < .001$). Положение лексем-признаков (до/после ЦС, условия уп и пу) не влияло на успешность восстановления пропущенного слова.

Выводы и обсуждение

Были выявлены определенные элементы ментального лексикона, которые могут быть рассмотрены как составляющие конвенциональную основу семантических репрезентаций использованных в исследовании целевых слов в словаре говорящего и слушающего. Важно подчеркнуть, что сами выявленные лексемы (а точнее, словообразовательные гнезда) не являются семантическими признаками понятий в привычном понимании, но, видимо, являются вербальными признаками таких семантических признаков. Когда эмпирически обнаружено, что *хвост* является существенным признаком понятия *собака*, какие семы или какие слоты фрейма «собака» и «хвост» были актуализированы? Признаком какого типа он может быть проинтерпретирован? Является ли это указанием на форму собаки, описанием зрительно воспринимаемого образа, обозначением части целого, результатом энциклопедических знаний об анатомии собаки или функциональной деталью, являющейся индикатором ее настроения? Вариантом интерпретации может быть множество, и яркий тому пример — разнообразие классификаций типов семантических признаков, компонентов, самих наборов сем и т. п., предлагаемых разными исследователями. Ситуативная зависимость категоризации одних и тех же понятий также была продемонстрирована экспериментально (напр., Barsalou, 1983). Это обстоятельство указывает на то, что носители языка, вероятно, в процессе коммуникации интерпретируют то или иное понятие различными способами. Смыслы говорящего никогда не могут быть абсолютно равны смыслам слушающего; технически это означает, что то в значении/-ях, что приводит к употреблению в речи того или иного слова, не обязательно является тем же, что формирует понимание или способствует прогнозированию этого же слова при восприятии. Но и то, и другое реализуется (при порождении) и обеспечивается (при восприятии) посредством конкретных лексических единиц, выполняющих таким образом функцию точек опоры репрезентации семантики понятий в ментальном лексиконе.

Литература

- Алефиренко Н. Ф.* Спорные проблемы семантики. М: Гнозис, 2005.
- Апресян Ю. Д.* Избранные труды: в 2 т. Т. 1. Лексическая семантика (синонимические средства языка) М.: Школа «Языки русской культуры», 1995.
- Вежицкая А.* Язык. Культура. Познание. М: Русские словари, 1996.
- Венцов А. В., Нигматулина Ю. О., Раева О. В., Риехакайнен Е. И., Слепокурова Н. А.* Корпус русских спонтанных текстов: структура и единицы // Труды международной конференции «Корпусная лингвистика 2013». СПб: С.-Петербург. гос. ун-т, 2013. С. 223 – 231.
- Залевская А. А.* Значение слова через призму эксперимента. Тверь: Твер. гос. ун-т, 2011.
- Ирисханова О. К.* Игры фокуса в языке. Семантика, синтаксис и прагматика дефокусирования. М: Школа «Языки русской культуры», 2014.
- Касевич В. Б., Ягунова Е. В.* Корпуса письменных текстов и моделирование восприятия речи // Вестник СПбГУ. 2006. Серия 2. № 3. С. 20 – 32.
- Кронгауз М. А.* Семантика. М: Рос. гос. гуманитар. ун-т, 2001.

Лич Д. Н. К теории и практике семантического эксперимента // Новое в зарубежной лингвистике. Вып. XIV. Проблемы и методы лексикографии. М: Прогресс, 1983. С. 108 – 133.

Методы когнитивного анализа семантики слова. Компьютерно-корпусный подход // Под ред. В. И. Заботкиной. М: Языки славянской культуры, 2015.

Найда Ю. А. Процедуры анализа компонентной структуры референционного значения // Новое в зарубежной лингвистике. Вып. XIV. Проблемы и методы лексикографии. М: Прогресс, 1983. С. 61 – 74.

Риехакайнен Е. И. Восприятие русской устной речи: контекст + частотность. СПб.: С.-Петербург. гос. ун-т, 2016.

Barsalou L. W. Ad hoc categories // *Memory & Cognition*. 1983. Vol. 11. No. 3. P. 211 – 227. doi:10.3758/bf03196968

Brouwer S., Mitterer H., Huettig F. Discourse context and the recognition of reduced and canonical spoken words // *Applied Psycholinguistics*. 2013. Vol. 34. No. 3. P. 519 – 539. doi:10.1017/s0142716411000853

Lucas M. Semantic priming without association: A meta-analytic review // *Psychonomic Bulletin & Review*. 2000. Vol. 7. No. 4. P. 618 – 630. doi:10.3758/bf03212999

McRae K., Cree G. S., Seidenberg M. S., McNorgan C. Semantic feature production norms for a large set of living and nonliving things // *Behavior Research Methods*. 2005. Vol. 37. No. 4. P. 547 – 559. doi:10.3758/bf03192726

Turney P. D., Pantel P. From frequency to meaning: Vector space models of semantics // *Journal of Artificial Intelligence Research*. 2010. Vol. 37. P. 141 – 188. doi:10.1613/jair.2934

Источники

Корпус русского литературного языка [Электронный ресурс]. <http://narusco.ru/>
Национальный корпус русского языка (НКРЯ) [Электронный ресурс]. <http://www.ruscorpora.ru/>

Русский ассоциативный словарь (РАС): В 2 т. М: АСТ Астрель, 2002. <http://www.tesaurus.ru/dict/dict.php>

VERBAL MEANS OF SEMANTIC REPRESENTATIONS OF CONCEPTS AS THE ANCHOR POINTS OF A SPEAKER'S AND A LISTENER'S MEANINGS IN THE MENTAL LEXICON

O.I. Pilatova

oiPilatova@yandex.ru

Saint Petersburg State University, Saint Petersburg

Abstract. How are semantics accessed by a listener to facilitate speech processing? To answer this question, we carried out an empirical study on eliciting semantic features of 10 concepts (expressed by 10 Russian nouns: *a book, a door, a bed*, etc.), a corpus-based study and a study on the predictiveness of context. It was observed that native speakers tend to use the same words (or ones which belong the same family of words) for semantic descriptions of each concept, even if they mention different features. Such semantically relevant words were grouped together and ranked by frequency. Virtually all of the words proposed by the participants were observed in their nearest verbal contexts. The presence in context of such feature-words leads to more successful prediction of the target words ($p < .\backslash001$). The results of the study show that the semantics of a concept are most likely presented in the mental

lexicon with the particular words that can serve as a conventional framework of a speaker's and a listener's vocabularies. They can be considered as lexical boosts for providing access to semantics and the prediction of a certain word quickly, precisely and at the same time flexibly enough as the same word can refer to various contextually relevant semantic features. Thus, such verbal means can be considered as the anchor points of a speaker's and a listener's meanings during communication.

Keywords: lexical semantics, mental lexicon, speech perception, word prediction, verbal associations, Russian nouns, family of words, semantic features