

УДК 81'1

DOI: 10.17223/19986645/43/6

З.И. Резанова, А.А. Миклашевский

МОДЕЛИРОВАНИЕ ОБРАЗНО-ПЕРЦЕПТИВНОГО КОМПОНЕНТА ЯЗЫКОВОЙ СЕМАНТИКИ ПРИ ПОМОЩИ ПСИХОЛИНГВИСТИЧЕСКОЙ БАЗЫ ДАННЫХ¹

В статье проведен анализ существующих подходов к описанию образно-перцептивного компонента языковой семантики, предлагается и обосновывается использование психолингвистических баз данных, позволяющих применять статистические методы анализа слабоструктурированной семантической информации. Рассматривается проект Томского государственного университета по разработке такой базы: представлены образцы инструкций, стимульный материал, описание процедуры сбора и анализа данных, предварительные выводы. Обсуждаются возможности применения разработанной базы данных в исследовательской практике.

Ключевые слова: базы данных, когнитивная семантика, перцептивная семантика.

1. Перцептивная информация в структуре концепта

В отечественной когнитивной лингвистике широко распространено представление о концепте как сложной неоднородной ментальной структуре, которая чаще всего описывается по полевой модели (ядро, ближняя и дальняя периферия). Несмотря на то, что набор конкретных элементов этой структуры и их иерархия могут определяться по-разному, большинство исследователей разделяет точку зрения, согласно которой именно конкретно-образные характеристики составляют ядро концепта или входят в число его ключевых компонентов: образная составляющая (С.Г. Воркачев [1. С. 7]), образ (М.В. Никитин [2. С. 59–60], З.Д. Попова и И.А. Стернин [3. С. 75]), конкретно-образные характеристики (Н.Н. Болдырев [4. С. 29]).

В.И. Карасик напрямую связывает образный компонент концепта с перцептивными ощущениями, полученными человеком из телесного опыта: «Уточним важнейшие измерения концепта – образное, понятийное и ценностное. Образная сторона концепта – это зрительные, слуховые, тактильные, вкусовые, воспринимаемые обонянием характеристики предметов, явлений, событий, отраженных в нашей памяти, это релевантные признаки практического знания» [5. С. 107]; на ту же связь указывают З.Д. Попова и И.А. Стернин, отсылая к понятию универсального предметного кода (УПК) Н.И. Жинкина: «Наличие в концепте образного компонента определяется самим нейролингвистическим характером универсального предметного кода: чувственный образ кодирует концепт, формируя единицу универсального предметного кода» [3. С. 75]. Значимость перцептивной информации в структуре кон-

¹ Научное исследование (№8.1.37.2015) выполнено при поддержке Программы «Научный фонд им. Д.И. Менделеева Томского государственного университета» в 2015–2016 гг.

This study (research grant No. 8.1.37.2015) was supported by The Tomsk State University Academic D.I. Mendeleev Fund Program in 2015–2016.

цепта подчеркивается А.А. Залевской в том числе и для абстрактных концептов: «Значение любого слова как единицы идиолексикона в принципе сводимо к некоторому исходному чувственному образу объекта (зрительному, слуховому, двигательному и т. д., актуализуемому прямо или через опосредование вербальными «переходами»), что должно находить проявление в констатации носителями языка наличия у идентифицируемых ими слов (даже с наиболее абстрактным значением) определенной степени конкретности и образности» [6. С. 181].

Иногда исследователи идут еще дальше и выделяют структуру внутри самого образного компонента. Так, например, З.Д. Попова и И.А. Стернин говорят о перцептивном образе (совокупности признаков, сформированных в процессе непосредственного восприятия фрагмента внешнего мира) и метафорическом или когнитивном образе (образных признаках, сформированных в результате метафорического осмысления предмета или явления) [3. С. 76]. В основании данной классификации лежит признак происхождения образа (был ли он получен при помощи органов чувств или же появился как результат формирования концептуальной метафоры); в обоих случаях, однако, образ состоит из одного и того же «психического материала», имеет одну природу, поэтому далее в настоящей работе различение перцептивного и образного компонентов не делается.

2. Существующие методы исследования перцептивного компонента языковой семантики

Несмотря на то, что значимость перцептивного компонента в структуре концепта широко признана, в настоящий момент нет окончательно сложившегося метода выявления перцептивной семантики языковых единиц. Исследовательское моделирование информационно-понятийного ядра концепта осуществляется при помощи традиционных лингвистических методов (таких как компонентный анализ, анализ контекстов, анализ словарных дефиниций, работа с фразеологическим материалом и т.д.), и результатом такого моделирования является набор смысловых компонентов, образующих соответствующий слой концепта и выраженных на метаязыке. Данные средства, однако, не так эффективны при моделировании образного компонента концепта. Причина этого видится в том, что сама природа перцептивной информации другая, эта информация слабо поддается семиотизации средствами естественного языка и потому с трудом может быть выявлена лингвистическими методами. Хотя язык тесно включен в общую когнитивную систему и может рассматриваться в двух качествах – как подверженный влиянию общих психофизиологических закономерностей, в том числе перцептивных, так и оказывающий влияние на ряд психических процессов, в том числе восприятие [7. С. 198–199], – очевидно, что не всякого рода информация, содержащаяся в когнитивной системе человека, одинаково легко доступна для анализа с использованием собственно лингвистических методов, что заставляет исследователей обращаться к поиску новых приемов выявления перцептивной информации как составной части концепта.

В работе И.А. Стернина и М.Я. Розенфельд «Слово и образ» предлагаются два метода извлечения образного компонента концептуального ядра: ана-

лиз словарных дефиниций и направленный психолингвистический эксперимент, при котором участникам предлагается следующая инструкция: «*Опишите всё, что вы видите, слышите, чувствуете, когда звучит каждое из слов экспериментального списка*» [8. С. 45]. Ниже кратко рассмотрены возможности и ограничения каждой из процедур.

1. При анализе словарных дефиниций авторы обнаруживают образный компонент собственно в толковании значения, в примерах, сопровождающих толкование, а также выявляют «лакуны» в толковании, которые, по их мнению, должны быть заполнены описаниями чувственных образов (например, указательные и неопределенные местоимения – *где-либо, в чем-либо*). Авторы утверждают, что образный компонент значения выражается инфинитивами (например, *выйти*), которые указывают на действующее лицо, не называя его, а также некоторыми существительными абстрактной семантики (*масса, множество, поток, вереница*). В случае с конкретными концептами исследователь часто вынужден признать определение целиком и полностью «образным». Например, приводя толкование для слова *рука*, авторы выделяют образный компонент в 7 словах дефиниции из 12: «каждая из двух верхних конечностей человека *от плечевого сустава до кончиков пальцев, а также от запястья до кончиков пальцев*» [9. С. 1132]. В зависимости от исследовательской позиции слова *два, верхний, конечность* и *человек* также легко могут быть включены в «образную» часть толкования. Авторы указывают, что в дефиниции могут встречаться отсылки к предметам и явлениям, воспринимаемым при помощи разных органов чувств (зрения, слуха, вкуса, обоняния, осязания), эмоционально-оценочные образы и образы разной степени детализации.

2. Направленный ассоциативный эксперимент, предлагаемый авторами, позволяет извлекать богатый материал для содержательного анализа; задание («*Опишите всё, что вы видите, слышите, чувствуете, когда звучит каждое из слов экспериментального списка*») требует от респондентов обращения к собственному перцептивному опыту, связанному со словом, и не накладывает ограничений на перцептивную модальность порождаемых образов (например, визуальная, аудиальная модальность и т.д.).

Очевидно, что метод анализа словарных дефиниций имеет ряд ограничений: (1) он основывается на материале, который не предназначен непосредственно для решения поставленной задачи (описание перцептивно-образного компонента семантики слова); (2) он не содержит однозначных критериев выделения перцептивного компонента (как его наличия, так и границ в тексте) и его классификации, так что в значительной мере данный вопрос решается на основе исследовательской интуиции; сами исследователи отмечают, что «если чувственное представление овнешляется в связном тексте, или даже в одном предложении текста, то в описании чувственного образа могут участвовать все слова предложения» [8. С. 56]; (3) поскольку в словарных толкованиях нет намеренного и систематического отражения информации о перцептивном компоненте семантики, анализ дефиниций не позволяет выявить соотношение между образами различных типов внутри одного концепта, степень их относительной значимости (удельный вес) в общей структуре семантики данного слова.

Предложенный И.А. Стерниным и М.Я. Розенфельд вариант ассоциативного эксперимента предоставляет, напротив, материал, собранный специально для решения поставленной задачи, и потому позволяет сфокусировать усилия исследователя на решении главной проблемы – анализе перцептивного компонента семантики. Кроме того, абстрагируясь от дискурсивной специфичности вхождения разных перцептивных компонентов в текст, данный метод позволяет анализировать относительную значимость образов разных модальностей (визуальной, аудиальной и т.д.) в структуре семантики слова, дает материал для содержательного анализа перцептивных образов, отвечает на вопрос «Какие именно образы возникают при восприятии слова?».

Ограничения обоих методов, рассмотренных выше, заключаются в том, что эти методы (1) приводят к включению в перцептивное ядро каждого концепта большого количества слабо упорядоченных компонентов с неопределенными отношениями между ними и (2) являются трудоемкими, очень сложно собрать ручным методом достаточно представительную выборку для того, чтобы делать выводы относительно языковой и когнитивной системы в целом.

Далее представлен метод создания психолингвистической базы данных, позволяющий преодолеть эти ограничения за счет перевода перцептивного компонента семантики в числовую форму.

3. Использование психолингвистических баз данных в современной лингвистике и когнитивной науке

Технологии баз данных активно применяются в современной лингвистике в связи с ростом объемов информации, необходимостью эффективно и быстро управлять информационными потоками (анализировать, структурировать данные), необходимостью систематизации и интеграции сведений, накопленных в разных дисциплинах и при помощи разных методов в рамках междисциплинарных исследований [10]. Базы данных применяются для работы с лингвистической информацией самого разного рода – фонетической, лексической, информации о морфемно-словообразовательном уровне языка, о типологических особенностях разных языков. Особое место занимают полнотекстовые базы данных, корпуса текстов, методологическая и практическая значимость которых как для лингвистики в целом, так и для когнитивной лингвистики в частности в настоящий момент интенсивно рефлексивируется (см., например, [11]); активно также используются материалы параллельных корпусов, включающих тексты на исходном языке и переводные тексты (см., например, [12]).

Примером базы данных может служить и частотный словарь, например «Новый частотный словарь русской лексики» О.Н. Ляшевской и С.А. Шарова [13], созданный на основе Национального корпуса русского языка: каждому слову присваивается значение (частотность), отражающее его позицию на некой шкале (от наименее частотных слов к наиболее частотным).

При этом в рамках психологии и психолингвистики стало уже традиционным использование шкал, предназначенных для количественного измерения семантических явлений, т.е. перевода нечисловой информации в числовую – семантический дифференциал Ч. Осгуда, шкала Лайкерта, суть кото-

рых заключается в построении семантического пространства (например, заданного координатами от 1 до 7) и размещении семантического объекта на заданной шкале, т.е. в определенной точке семантического пространства. Так, например, респонденту предлагается оценить степень проявления признака по шкале от 1 до 7, где 1 будет означать самую слабую степень проявления, а 7 – наибольшую. Аналогичным образом могут быть переведены в числовую форму другие слабоструктурированные виды информации.

В рамках когнитивной лингвистики этот метод был дополнен статистическим анализом данных значительной выборки респондентов для того, чтобы иметь возможность судить не об индивидуальных особенностях восприятия и оценки семантических компонентов, а об общих закономерностях, справедливых для среднего носителя языка. Как правило, в такой базе данных представлено два базовых показателя для совокупности значений – среднее (mean), отражающее усредненную оценку семантического объекта носителями языка, и стандартное отклонение (standard deviation), отражающее степень разброса оценок. Одной из первых работ по созданию подобной психолингвистической базы данных стала работа A. Paivio, J. Yuille и S. Madigan «Concreteness, imagery, and meaningfulness values for 925 nouns» [14], опубликованная в 1968 г., в которой собраны средние оценки по таким показателям, как субъективная конкретность/абстрактность, образность слова (способность слова вызывать мысленные образы различных модальностей), субъективная оценка многозначности слова. В работе Keuleers и Balota «Megastudies, crowdsourcing, and large datasets in psycholinguistics: An overview of recent developments» [15] представлен обзор психо- и собственно лингвистических баз данных, которые используются сегодня в психолингвистике, когнитивной лингвистике (в том числе корпусов текстов). Авторы приходят к выводу, что в эпоху больших данных эти методы изменяют представления о самом исследовательском процессе («This has opened new ways of *doing* psycholinguistics») [15. С. 1458].

4. Психолингвистические шкалы, использующиеся для создания баз данных

Для оценки могут быть выбраны самые разные параметры, и в настоящее время на материале английского и ряда других европейских языков создано значительное количество баз данных, содержащих оценки по различным психолингвистическим шкалам. Обобщая информацию из разных источников, назовем наиболее важные и активно используемые в психолингвистической исследовательской практике типы шкал.

I. Отдельную группу характеристик составляют параметры, отражающие психофизиологический и эмоциональный опыт человека, связанный с теми или иными единицами языка:

- 1) образность слова (imageability) [14, 16–18];
- 2) субъективная конкретность/абстрактность [14, 17, 19];
- 3) физическое взаимодействие человеческого тела с объектом (BOI, body-object interaction) [18, 20];
- 4) возможность манипулировать объектом при помощи рук (manipulability) [21];

5) связь слов с отдельными модальностями восприятия (modality rating) [22, 23];

6) расположение референта в пространстве по шкале верх-низ;

7) субъективная оценка слова (pleasantness);

8) эмоциональность слова (emotionality).

II. В другую группу объединяются характеристики, описывающие функционирование лексических единиц:

1) субъективный возраст понимания (subjective age of acquisition) [24];

2) субъективная частотность слова (subjective frequency);

3) оценка того, насколько слово знакомо (familiarity) [16];

4) доступность контекста¹ (context availability) [17, 24];

5) оценка того, насколько знакома идея, выражаемая словом (concept familiarity) [21, 25];

6) количество ассоциаций (number of associations) [24];

7) количество значений, многозначность (number of meanings).

III. Многие характеристики не укладываются в какую-либо группу и возникают ситуативно в связи с частными исследовательскими задачами. Ниже представлены некоторые из них:

1) семантическая прозрачность (semantic transparency) [16];

2) легкость определения (ease of definition) [24];

3) типичность (typicality, semantic typicality) [21, 25];

4) легкость произношения слова (pronounceability) [24].

IV. Помимо субъективных характеристик, исследователи нередко включают в БД и объективные, например, такие, как объективный возраст понимания², объективная частотность (примером может служить уже упомянутый словарь О.Н. Ляшевской и С.А. Шарова [13], основанный на материале Национального корпуса русского языка [26]). К объективным, хотя и полученным в ходе психолингвистического эксперимента характеристикам можно отнести *meaningfulness* [14], показатель, отражающий количество ассоциаций, данных участниками исследования на стимульное слово за единицу времени (например, за 0,5 мин или за 1 мин). Данный показатель отражает способность слова легко и быстро вызывать ассоциации³.

Нередко в исследованиях используется целая группа объективных параметров, описывающих форму слова (*word surface form*) – количество букв, количество звуков, количество слогов, номер ударного слога (для языков с подвижным ударением), количество слов со сходным звучанием, написанием

¹ В данном случае слово «контекст» употребляется в несколько непривычном для лингвистики смысле, отсылая не к собственно лингвистическому контексту, а непосредственно к внеязыковым ситуациям.

² Определяется при помощи *picture naming task*, эксперимента, в ходе которого экспериментатор показывает респонденту серии изображений и называет слова, а задача респондента – указать на изображение, соответствующее названному слову.

³ В данном случае результатом эксперимента является не содержание ассоциаций, а их число, причем участники исследования не знают о том, что именно будет измеряться в ходе эксперимента, а исследователь рассчитывает число ассоциаций самостоятельно, поэтому данный показатель отнесен к объективным.

(word neighbours), количество слов с той же начальной буквой, тем же начальным звуком, число рифм к слову и другие показатели¹.

Перевод семантических и других показателей в числовое выражение делает возможным последующее применение статистических методов к полученным данным – нахождение среднего арифметического и стандартного отклонения, анализ распределения оценок, применение дисперсионного анализа (ANOVA), при использовании нескольких шкал одновременно может применяться корреляционный и регрессионный анализ, факторный анализ, кластерный анализ. Статистический анализ больших массивов данных позволяет снимать исследовательский субъективизм и нивелировать уникальные особенности индивидуального сознания, традиционно бывшего предметом исследования психолингвистики, находя, таким образом, усредненные значения для языковых единиц. Следует особо подчеркнуть, что именно языковые элементы и их группы являются единицами такого анализа, следовательно, выводы делаются о языковой системе и ее семантике, а не только о самих психологических или психолингвистических переменных.

Одной из первых подобных работ на материале русского языка является словарь «Органы чувств, эмоции и прилагательные русского языка» [27], в котором собраны сведения об оценке связи русских прилагательных с различными органами чувств. Релевантными, хотя и несколько отличающимися по целям создания, являются электронная система «Существительное и объект: библиотека стимулов и нормативы для экспериментальных исследований» [28] и «Библиотека стимулов: глаголы и существительные» [29], созданные на материале исследований, представленных в работе [30].

5. Проект базы данных, моделирующей перцептивный компонент языковой семантики

В Томском государственном университете в настоящее время ведется работа по созданию психолингвистической базы данных, включающей информацию о связи русских существительных с модальностями восприятия, т.е. зрением, слухом, вкусовыми, обонятельными и осязательными ощущениями, а также дополнительную информацию, такую, например, как субъективный возраст усвоения слова, субъективная частотность слова и т.д. На данный момент в исследовании приняли участие свыше 500 информантов.

5.1. Отбор стимульного материала и параметров оценивания

Для создания базы данных были отобраны существительные различных семантических категорий, представляющие собой наименования:

- 1) действий и процессов (например, *атака, прыжок, вдох, подъем, сжати*);
- 2) животных (например, *барсук, медведь, жук, птица, кошка, хамелеон*);
- 3) частей тела (например, *макушка, ухо, бровь, лицо, стопа, череп*);

¹ Значения для данных показателей рассчитываются вручную либо определяются по специальным словарям и информационным системам. Они также могут быть субъективными, например, как в работе [24].

4) строений и их частей (например, *подвал, погреб, шахта, навес, небо-скреб*);

5) предметов одежды и аксессуаров (например, *пуговица, сапог, серьга, наушники, носок, кольцо*);

6) продуктов питания (например, *гриб, брусника, яблоко, крыжовник, тыква, кокос*);

7) элементов ландшафта, типов поверхности (например, *гора, углубление, трава, пруд, газон, асфальт*);

8) ментальных процессов и объектов (например, *воображение, теория, замысел, знание, сон, фантазия*);

9) чувств и эмоций (например, *обида, любовь, ревность, счастье, веселье, смущение*);

10) физических ощущений (например, *озноб, мягкость, боль, тошнота, сухость*);

11) звуков (например, *гудок, звон, писк, хрип, стук, кашель, шум*);

12) небесных объектов и явлений (например, *звезда, луна, метеорит, молния, радуга, рассвет, туча*);

13) инструментов и орудий, приводимых в действие с помощью мелкой моторики (например, *клавиша, гвоздь, игла, гайка, скрепка, бокал, точилка*);

14) инструментов и орудий, приводимых в действие с помощью крупной моторики (например, *пила, сковорода, метла, стакан, мяч, лопата, топор*);

15) транспортных средств (например, *электричка, машина, карета, мопед, повозка, самолет, скейт, танк*);

16) отдельных артефактов, включенных в предыдущие исследования, и не отнесенных к какой-либо категории в настоящем исследовании (например, *антенна, гардина, капкан, половик, розетка, флаг*); закодированы в базе данных как категория object.

Общее количество существительных – 506. В базу данных попали как конкретные, так и абстрактные существительные, что не является характерным для аналогичных зарубежных баз, однако такой подход позволяет, как представляется, сравнивать конкретные и абстрактные существительные между собой по заданным параметрам. Все слова оценивались респондентами по следующим шкалам:

1. Образность слова (imageability) – легкость, с которой слово вызывает у респондента образы объекта, качества, действия и т.д. При оценке образности слова использовалась следующая инструкция: *Слова различаются по своей способности вызывать мысленные образы вещей или явлений. Например, когда Вы слышите или читаете слово «яблоко», то, вероятно, Вы легко и быстро можете представить себе образ этого предмета. С другой стороны, не так легко представить, например, «факт». Пожалуйста, оцените слова в списке по шкале от 1 до 7, где 1 – низкий балл образности (слово с трудом и медленно вызывает образы или не вызывает вовсе), а 7 – высокий балл образности (слово вызывает образы легко и быстро). Помните, что в данном задании не может быть правильных или неправильных ответов, ориентируйтесь на свои собственные ощущения. Пожалуйста, не делайте перерывов и оцените все слова за 1 сеанс.*

2. Связь слов с отдельными модальностями восприятия (modality rating) – зрением, осязанием, слухом, вкусом, обонянием.

При оценке данного параметра использовалась следующая инструкция: *Разные слова могут быть связаны с разными модальностями восприятия – зрением, слухом, вкусом, запахом или тактильными ощущениями (осязанием). Пожалуйста, оцените слова в списке по их связи со зрением: 7 обозначает максимально сильную связь слова со зрением, 1 – самую слабую связь слова со зрением. Работайте по возможности быстро, не задумываясь подолгу над каждым словом. В то же время старайтесь быть объективными и сосредоточенными. Если у Вас возникает такая необходимость, Вы можете вернуться к инструкции и прочитать ее еще раз, а после продолжить оценивать существительные. Пожалуйста, не делайте перерывов во время работы над данной анкетой, оцените все слова за 1 сеанс.* Для других модальностей восприятия была использована аналогичная инструкция, в которой слово зрение заменялось на вкус/обоняние/осязание/слух соответственно.

После сбора первичной информации по данному показателю слова могут быть распределены по тем модальностям, с которыми они статистически значимо связаны (т.е. значения для слов будут переведены в номинальную шкалу), могут быть выделены унимодальные (связанные только с одной модальностью) и полимодальные (связанные с разными модальностями) слова.

3. Возможность манипулировать объектом при помощи рук (manipulability). При оценке данного параметра использовалась следующая инструкция: *Для того, чтобы некоторые предметы выполняли свои функции, человек должен использовать руки. Например, чтобы сигарета выполняла свою функцию («была выкурена»), использование руки необходимо. С другой стороны, человек никогда не взаимодействует при помощи рук с вулканом, чтобы тот выполнял функцию («извергался»). Пожалуйста, оцените слова в списке, в зависимости от того, насколько необходимо использование человеческих рук, чтобы предметы выполняли свои типичные функции: 7 – использование рук необходимо всегда, 1 – руки никогда не используются для взаимодействия с этим предметом. Некоторые предметы могут выступать в разных качествах: устрица может быть ингредиентом блюда (и тогда люди используют руки, чтобы устрица «была приготовлена»); с другой стороны, это живое существо, и для того, чтобы устрица «жила» людям не нужно взаимодействовать с ней при помощи рук. В таких случаях ориентируйтесь на тот вариант, который первым пришел Вам в голову.*

4. Расположение референта в пространстве по шкале верх–низ – оценка того, насколько высоко или низко относительно наблюдателя расположен референт слова. Данная характеристика была применена в настоящем проекте не только для оценки конкретных существительных, референты которых могут быть восприняты при помощи органов чувств, но и для абстрактной лексики. Была использована следующая инструкция: *Некоторые предметы и явления действительности обычно расположены выше или ниже в пространстве. Пожалуйста, оцените, насколько высоко или низко расположен предмет, используя шкалу от 1 до 7, где 1 – «очень низко», а 7 – «очень высоко». Работайте по возможности быстро, не задумываясь подолгу над каждым словом. В то же время старайтесь быть объективными и сосредото-*

ченными. Если у Вас возникает такая необходимость, Вы можете вернуться к инструкции и прочитать ее еще раз, а после продолжить оценивать существительные. Пожалуйста, не делайте перерывов во время работы над данной анкетой, оцените все слова за 1 сеанс¹.

5. Субъективный возраст понимания (subjective age of acquisition) – показатель, отражающий тот возраст, в котором респонденты, по их мнению, выучили какое-либо слово [24]. При оценке данного параметра использовалась следующая инструкция: *Разные слова мы выучиваем в разном возрасте. Пожалуйста, поставьте возле каждого слова в списке число, которое соответствует возрасту (в годах), когда Вы выучили это слово. Пожалуйста, используйте числа от 0 до 15. Поставьте прочерк, если слово Вам незнакомо. Работайте по возможности быстро, не задумываясь подолгу над каждым словом. В то же время старайтесь быть объективными и сосредоточенными. Если у Вас возникает такая необходимость, Вы можете вернуться к инструкции и прочитать ее еще раз, а после продолжить оценивать существительные. Пожалуйста, не делайте перерывов во время работы над данной анкетой, оцените все слова за 1 сеанс.*

6. Субъективная частотность слова (subjective frequency) – степень того, насколько часто респонденты, по их мнению, встречаются с тем или иным словом или употребляют его. Для данного параметра была использована следующая инструкция: *Разные слова употребляются в речи с разной частотой. Пожалуйста, оцените, насколько часто Вы встречаетесь с каждым словом, используя шкалу от 1 до 7, где 1 – «очень редко», а 7 – «очень часто». Работайте по возможности быстро, не задумываясь подолгу над каждым словом. В то же время старайтесь быть объективными и сосредоточенными. Если у Вас возникает такая необходимость, Вы можете вернуться к инструкции и прочитать ее еще раз, а после продолжить оценивать существительные. Пожалуйста, не делайте перерывов во время работы над данной анкетой, оцените все слова за 1 сеанс.*

Отбор данных параметров оценивания и семантических групп проведен в соответствии с последними нейрокогнитивными исследованиями. Так, в частности, было показано, что обработка слов, отсылающих к разным перцептивным модальностям, приводит к активации разных нейронных паттернов в головном мозге (см. обзор в работе [31]), таким же образом различается и обработка слов, принадлежащих к разным семантическим группам, вплоть до тонких различий наименований действий, выполняемых при помощи руки или ноги [31. С. 437–442]; пациенты с нейрофизиологическими расстройствами (например, болезнью Альцгеймера) иногда демонстрируют утрату знания одних категорий (растений, животных) при сохранении других (инструментов, предметов мебели) [32], что, вероятнее всего, связано с различными формами взаимодействия человека с представителями разных категорий и, как следствие, с различным перцептивным опытом, хранящимся в сознании [32. С. 806]. Значимость перцептивной модальности как компонента се-

¹ Для половины респондентов использовалась инструкция с противоположными значениями (1 – «верх», 7 – «низ»); все полученные ответы затем были приведены к единой системе измерения.

мантической структуры, ее психологическая реальность продемонстрированы в поведенческих экспериментах с измерением времени реакции [33, 34], причем даже для идиоматических выражений, где моторно-перцептивные компоненты используются в метафорическом смысле, как, например, в выражении *протянуть руку помощи* [35]. В целом данные исследования согласуются с фундаментальным положением об антропоцентризме языка на всех его уровнях, которое принято и в отечественной когнитивной лингвистике (например, [36. С. 16–17]), развивая эту идею в русле теории воплощенного познания (*embodied cognition*, см., например, [37]).

5.2. Сбор данных

Респонденты заполняли анкеты в бумажном виде, затем переводились в электронный вид (использовалось ПО Microsoft Office Excel). Каждая анкета была распечатана на листе А4 и содержала инструкцию и около 60 случайно отобранных слов из общего списка. Каждый участник эксперимента получал на руки по 5 таких анкет, все анкеты – с разными параметрами оценивания; ни одно слово не повторялось дважды для одного участника. Об участниках эксперимента собирались следующие демографические сведения: пол, возраст, уровень образования (школьник vs студент vs высшее образование), направление образования (гуманитарное vs естественно-научное vs техническое). Также собирались сведения о родном языке респондента; результаты тех респондентов, которые указали в качестве родного не русский язык, были исключены из дальнейшего анализа. При желании участники могли пройти эксперимент повторно, получив новые слова и/или шкалы для оценивания; в таком случае они регистрировались как новые участники. По возможности соблюдалось равномерное распределение анкет среди респондентов мужского и женского пола.

Для достижения статистической надежности результатов на каждое слово по каждой шкале было собрано около 30 оценок.

5.3. Обработка собранных данных, возможности применения результатов

Собранные в ходе психолингвистического эксперимента оценки участников исследования были усреднены, найдено стандартное отклонение (использовалось ПО STATISTICA).

На рис. 1 представлена дескриптивная статистика по 9 параметрам оценивания (без субъективной частотности) – средние, стандартные отклонения, максимальные и минимальные значения (для всей выборки из 506 существительных).

Минимальное и максимальное значения – наибольшее и наименьшее значения в выборке, полученные после усреднения 30 оценок, данных участниками на каждое слово. Так, например, по силе связи с тактильными ощущениями максимальное значение (5,93 из 7) получило слово *кошка*, т.е. в данной выборке это слово сильнее всего связано для респондентов с тактильными ощущениями. По этому же параметру минимальное значение (1,14 из 7) получило слово *манера*, т.е. оно слабее всего связано с тактильными ощущениями. Еще одним примером могут служить слова *малина* и *морковь*, полу-

чившие максимальные значения по степени образности (7 из 7), т.е. эти слова быстрее и легче всего вызывают образы в сознании. Минимальное значение (1,93 из 7) по степени образности получило слово *допущение*, т.е. респонденты с трудом формируют образы при восприятии семантики данного слова.

Параметр	Mean	SD	Min	Max
Vis	3,98	1,03	1,38	6,56
Aud	2,71	1,19	1,10	6,48
Olf	2,32	1,15	1,00	6,41
Gus	2,17	1,39	1,00	6,59
Hap	3,39	1,18	1,14	5,93
Img	5,60	1,40	1,93	7,00
Man	3,91	1,62	1,03	6,80
Space	3,45	1,23	1,10	6,82
AoA	5,70	1,87	2,54	11,78

Рис. 1. Дескриптивная статистика: Vis – сила связи слова с визуальной модальностью (от 1 до 7); Aud – сила связи слова с аудиальной модальностью (от 1 до 7); Olf – сила связи слова с ольфакторной модальностью (запахом) (от 1 до 7); Gus – сила связи слова со вкусовой модальностью (от 1 до 7); Hap – сила связи слова с тактильной модальностью (от 1 до 7); Img – степень образности слова (от 1 до 7); Man – возможность взаимодействовать с референтом при помощи рук (от 1 до 7); Space – локализация референта в пространстве (1 – низ, 7 – верх); AoA – субъективный возраст усвоения слова (в годах); Mean – среднее значение, SD – стандартное отклонение, Min – минимальное значение, Max – максимальное значение (статистики найдены на материале 506 значений)

Стандартное отклонение (standard deviation, SD) показывает, насколько сильно значения в выборке рассеяны вокруг среднего значения. Так, например, можно заметить, что стандартное отклонение для связи слов с визуальной модальностью меньше, чем со вкусовой (1,03 и 1,39 из 7 соответственно). Это означает, что значения по шкале связи с визуальной модальностью более «плотно» сосредоточены вокруг среднего значения (3,98 в данном случае), т.е. большая часть элементов выборки схожа по данному параметру (рис. 2), в то время как значения по силе связи со вкусовой модальностью более сильно рассеяны вокруг среднего (2,17), и, как можно видеть на рис. 3, здесь даже выделяются две четко различимые группы слов – одна (небольшая по количеству) с выраженной связью со вкусовыми ощущениями (в верхней части графика) и вторая, не связанная вовсе или слабо связанная (в нижней части графика, основная масса единиц).

Среднее значение получается в результате простого арифметического усреднения всех значений выборки (в данном случае 506 значений) и отражает общий «центр тяжести» выборки по данному параметру. Как можно судить по приведенным на рис. 1 данным, среднее значение по шкале связи слов выборки со зрением является наибольшим среди пяти модальностей (3,98), на втором месте находится осязание (3,39), затем – слух (2,32); самая слабая сила связи слов выборки со вкусом (2,71) и запахом (2,32). Данные закономерности соответствуют обнаруженным ранее в рамках психологии восприятия: человек получает наибольшее количество информации через зрение, затем следуют слух и осязание, наименьшую роль играют ольфакторные и вкусовые ощущения [38, С. 91, 177, 199, 222].

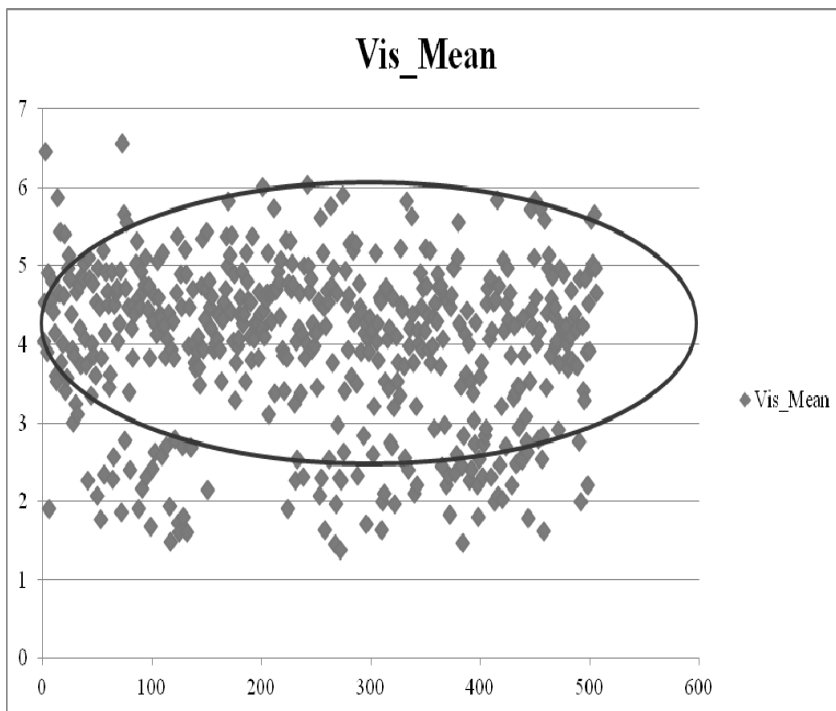


Рис. 2. Значения слов по силе связи со зрительной модальностью (от 1 до 7)

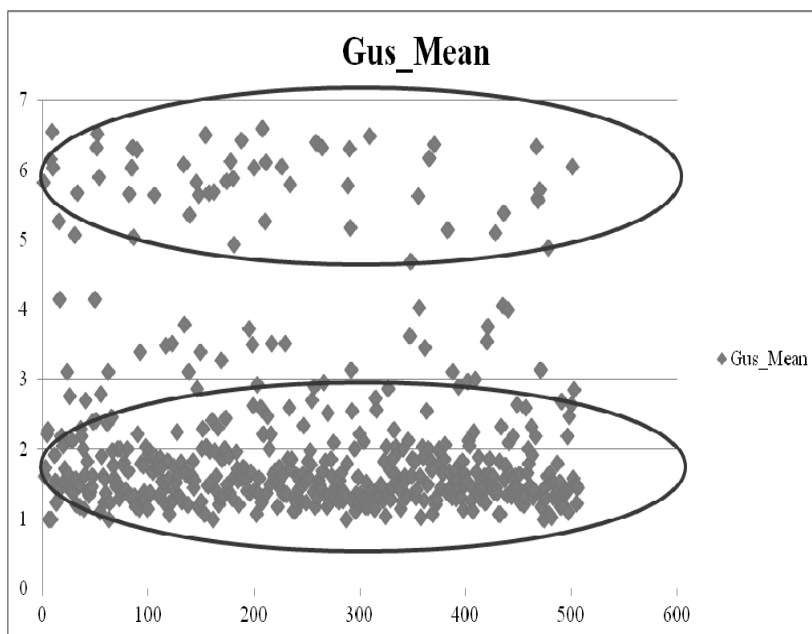


Рис. 3. Значения слов по силе связи со вкусовой модальностью (от 1 до 7)

Проведенный анализ демонстрирует, что семантика естественного языка изоморфна структуре человеческого восприятия. Также обнаружено, что среднее значение по шкале образности (5,6) существенно выше середины шкалы (3,5) и приближается к верхней границе (7), что соотносимо с результатами зарубежных исследований на материале других языков и подтверждает, что понимание семантики естественного языка в значительной степени опирается на образы различной природы, перцептивно-образный компонент.

Аналогичным образом может быть рассчитана дескриптивная статистика по отдельным группам слов, выделенным по различным основаниям (формальным, семантическим). Например, на рис. 4 представлена дескриптивная статистика по шкале образности для каждой из семантических категорий слов, входящих в базу данных. Анализ показывает, что Слова разных семантических категорий существенно различаются по среднему значению образности, что более наглядно представлено на рис. 5.

Group	N	Img_Mean	Img_SD
Action	30	4,03	1,05
Animal	76	6,58	0,27
Body_Part	14	6,29	0,44
Building	8	6,1	0,51
Clothes	28	6,43	0,47
Food	51	6,59	0,45
Ground	25	5,84	0,9
Intelligence	29	3,07	0,67
Object	15	6,23	0,42
Sense_Emotion	45	3,6	0,59
Sense_Phys	21	3,74	0,7
Sound	17	3,74	0,67
Space	18	6,09	0,98
Tool_power_grip	34	6,29	0,8
Tool_precise_grip	48	6,38	0,49
Transport	47	6,3	0,62
All groups	506	5,6	1,4

Рис. 4. Дескриптивная статистика для образности (imageability) по семантическим категориям: Group – семантическая категория; N – количество единиц в категории; Img_Mean – среднее значение по шкале образности; Img_SD – стандартное отклонение по шкале образности; Action – наименования действий и процессов; Animal – наименования животных, Body_Part – наименования частей тела; Building – наименования строений и их частей; Clothes – наименования предметов одежды и аксессуаров; Food – наименования продуктов питания; Ground – наименования элементов ландшафта, типов поверхности; Intelligence – наименования ментальных процессов и объектов; Object – наименования отдельных артефактов (не вошедших в другие категории); Sense_Emotion – наименования чувств и эмоций; Sense_Phys – наименования физических ощущений; Sound – наименования звуков; Space – наименования небесных объектов и явлений; Tool_power_grip – наименования инструментов и орудий, приводимых в действие с помощью крупной моторики; Tool_precise_grip – наименования инструментов и орудий, приводимых в действие с помощью мелкой моторики; Transport – наименования транспортных средств; All_Groups – общие показатели по всей выборке

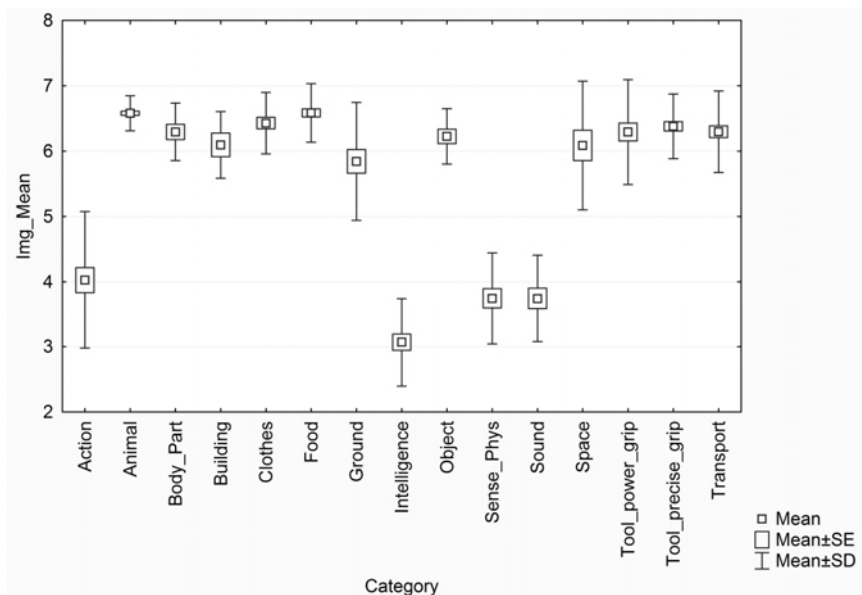


Рис. 5. Графическое представление дескриптивной статистики для образности (imageability) по семантическим категориям. Средние значения по каждой категории изображены в виде квадратов, вертикальные линии обозначают стандартное отклонение, отложенное в большую и меньшую сторону. Обозначения категорий см. в пояснении к рис. 4

Так, например, можно сделать вывод, что наименования действий, ментальных состояний и объектов, физических ощущений и звуков в среднем обладают значительно меньшей образностью, чем существительные других семантических категорий (следует, однако, отметить, что это предварительный вывод, сделанный на основании субъективной оценки графика; чтобы утверждать наличие значимой разницы между категориями по какому-либо признаку следует использовать соответствующие статистические критерии, например критерий Крускала – Уоллиса).

WORD	Category	Vis	Aud	Olf	Gus	Hap	Img	Man	Space	AoA
абрикос	Food	4,03	1,63	5	5,82	4,97	5,93	6	3,92	4,13
автобус	Transport	4,53	4,77	2,97	1,61	5,14	6,67	5,66	3,22	4,03
аист	Animal	3,9	1,79	2,21	2,21	2,66	6,55	1,9	5,24	4,33
акцент	Sound	1,91	4,64	1,38	1	1,68	3,82	1,8	3,64	9,4
асфальт	Ground	4,14	1,79	4	1,91	3,97	6,66	3,86	2,15	4,57
атака	Action	3,6	3,71	1,81	1,54	3,11	5	5,14	4,22	5,62

Рис. 6. Фрагмент базы данных: WORD – слово; Category – семантическая группа (обозначения категорий см. в пояснении к рис. 4); Vis – среднее значение силы связи слова с визуальной модальностью (от 1 до 7); Aud – среднее значение силы связи слова с аудиальной модальностью (от 1 до 7); Olf – среднее значение силы связи слова с ольфакторной модальностью (запахом) (от 1 до 7); Gus – среднее значение силы связи слова со вкусовой модальностью (от 1 до 7); Hap – среднее значение силы связи слова с тактильной модальностью (от 1 до 7); Img – среднее значение степени образности слова (от 1 до 7); Man – среднее значение по шкале manipulability (возможность взаимодействовать с референтом при помощи рук) (от 1 до 7); Space – локализация референта в пространстве (1 – низ, 7 – верх); AoA – субъективный возраст усвоения слова (в годах)

На рис. 6 представлен фрагмент базы данных, позволяющий судить о ее структуре.

Представленная база данных позволяет применять к семантическим переменным статистические методы, как напрямую сравнивая значения групп слов (выделенных по тому или иному основанию, например семантическому, грамматическому и т.д.), используя анализ дисперсии ANOVA, так и сопоставляя значения по разным шкалам между собой (т.е. находить корреляции между шкалами). Приведем далее ряд установленных на собранной базе данных корреляций психолингвистических признаков слов.

Так, например, проведенный корреляционный анализ собранных в базе данных позволил установить сильную корреляцию между такими показателями, как связь слова со зрением и образность слова ($r=0,73$), что можно интерпретировать как доказательство того, что именно зрительные ощущения вносят наибольший вклад в образность слова; так, например, слова *фейерверк*, *тигр*, *яблоко*, получившие высокие оценки по шкале образности, также получили высокие оценки по силе связи со зрением, тогда как слова *допущение*, *гипотеза*, *смысл* получили низкие оценки по обеим шкалам. С другой стороны, хотя эти параметры связаны между собой, они не являются тождественными: можно найти слова с высокой степенью образности, но слабой силой связи со зрительными ощущениями (*вдох*, *стук*, *гудок*).

Также сильно коррелируют между собой вкусовые и обонятельные ощущения, отраженные в семантической структуре слова ($r=0,74$), что соответствует особенностям телесного взаимодействия человека со внешним миром: как правило, человек знает запах тех объектов, которые пробует на вкус (продукты питания), следовательно, в его сознании ольфакторная и вкусовая информация часто оказываются связаны друг с другом. Например, установлена высокая степень связи одновременно с ольфакторными и вкусовыми ощущениями у таких слов, как *малина*, *апельсин*, *вишня*; в то же время есть слова со значительной силой связи с ольфакторными ощущениями, не связанные при этом со вкусовыми, – *сырость*, *собака*, *газон*. При этом обратное соотношение (существительные, связанные со вкусовыми и не связанные с ольфакторными ощущениями) наблюдается в меньшем количестве примеров: такие слова, как *утка* или *мел*, хотя и обладают большими показателями по шкале связи со вкусовыми ощущениями, чем с ольфакторными, однако данная тенденция выражена слабее.

Кроме того, обнаружено множество других значимых, хотя и менее сильных корреляций между отдельными показателями. Так, например, выявлена отрицательная корреляция между возрастом усвоения и образностью слова ($-0,56$), т.е. чем раньше человек выучивает слово, тем большей образностью оно обладает, что может объясняться как собственно когнитивными механизмами (последовательным развитием мышления от наглядно-образного к абстрактно-логическому в ходе онтогенеза), так и дискурсивными средами, в которых оказывается человек в ходе своего обучения и взросления (освоение терминологии, абстрактной лексики в процессе получения образования); оба аспекта, разумеется, тесно связаны между собой.

На основе базы данных возможно и изучение соотношения образности с различными модальностями восприятия: если между образностью и связью с

визуальной модальностью есть сильная корреляция, как уже было сказано выше, то между образностью и аудиальной модальностью значимой корреляции не выявлено. Можно сделать вывод, что слуховые ощущения «мысленно представить» сложнее по сравнению с ощущениями других модальностей; умеренно коррелируют с образностью значения по шкале тактильных ощущений, слабые корреляции наблюдаются между образностью и вкусовыми и ольфакторными ощущениями.

Базы данных такого типа важны для современной отечественной психо- и когнитивной лингвистики и в контексте обращения последней к аппаратным экспериментальным исследованиям – измерению времени реакции (см., например, [39]), исследованиям при помощи технологии ай-трекинга (см., например, [40]). Для такого рода экспериментальных исследований, как правило, требуется стимульный материал, проконтролированный по различным психолингвистическим параметрам, и первым этапом исследования является отбор такого материала, что значительно упрощается при наличии готовых баз данных, в которых содержится информация о наиболее типичных и востребованных психолингвистических характеристиках стимулов разного рода (лексических, фразеологических единиц, метафор и т.д.).

6. Обсуждение возможных ограничений метода

Наиболее очевидным ограничением с точки зрения привычного лингвистического описания представляется числовой формат представления семантики, который не позволяет говорить о содержательных различиях между точками на отдельно взятой шкале. Данное ограничение является закономерным продолжением достоинства метода: снимая семантическую неоднородность в процессе математического моделирования, он дает возможность массово сравнивать между собой языковые единицы по выбранному критерию, применяя статистические процедуры к семантическим переменным. Данный метод может рассматриваться и как дополнительный по отношению к направленному ассоциативному эксперименту, представленному в работе И.А. Стернина и М.Я. Розенфельд «Слово и образ» [11] (см. выше), где ассоциативный эксперимент выявляет содержательное наполнение перцептивного компонента, а психолингвистическая база данных предлагает количественное выражение различных элементов перцептивного ядра концепта.

Другим возможным замечанием является то, что в исследованиях семантики центральной единицей должен быть лексико-семантический вариант, отдельное значение слова, в то время как в анкете участники исследования работают со словом. Данное ограничение является общим для всех психолингвистических методик, в которых слова предъявляются вне контекста (в том числе для ассоциативного эксперимента, рассмотренного выше): либо исследователь однозначно определяет слово при помощи контекста и тогда не может устранить влияние этого контекста, либо слово предъявляется изолированно, и тогда нельзя с полной уверенностью сказать, с каким именно ЛСВ работают участники исследования. Это ограничение частично преодолевается за счет того, что разные ЛСВ обладают различной частотностью, и в условиях ограниченного по времени эксперимента участники с большей вероятностью реагируют именно на наиболее частотное значение, так как оно

ближе всего к актуальному слою сознания. Кроме того, включение в базу данных такого показателя, как стандартное отклонение, позволяет выявить степень различия между оценками отдельных участников и определить, таким образом, какие слова вызвали наибольшие расхождения при интерпретации (возможно, вследствие актуализации разных ЛСВ).

7. Заключение

Проблема описания перцептивного компонента семантики является частью фундаментальной проблемы установления соотношения языка и восприятия, традиционной для лингвистики (см. обзор в работе [41. С. 76–95]), и потому вряд ли может быть решена одним методом; продуктивным представляется использование разных методик для достижения частных исследовательских целей. В настоящей работе представлен метод создания психолингвистических баз данных, позволяющий перевести слабоструктурированную и/или труднодоступную для описания информацию о перцептивном компоненте семантики в числовую форму, что делает возможным применение статистических процедур при работе с семантикой. Перцептивная информация в семантической структуре слова представлена в виде средних значений по шкале связи слова с пятью модальностями восприятия (зрением, слухом, вкусом, обонянием, осязанием); кроме того, найдены средние значения по таким показателям, как образность слова (способность быстро и легко вызывать образы в сознании), возможность манипулировать объектом-референтом при помощи рук, локализация референта в пространстве по вертикальной оси, субъективный возраст усвоения слова. Собранный баз данных может быть использован как в качестве самостоятельного инструмента анализа семантики, так и в качестве источника стимульного материала для психолингвистических и когнитивных экспериментальных исследований.

Литература

1. Воркачев С.Г. Счастье как лингвокультурный концепт. М.: Гнозис, 2004. 192 с.
2. Никитин М.В. Развернутые тезисы о концептах // Вопросы когнитивной лингвистики. 2004. № 1. С. 53–64.
3. Попова З.Д., Стернин И.А. Когнитивная лингвистика. М.: АСТ: Восток – Запад, 2007. 314 с.
4. Болдырев Н.Н. Когнитивная семантика: курс лекций по английской филологии. 2-е изд., стер. Тамбов: Изд-во Тамбов. ун-та, 2001. 123 [110] с.
5. Карасик В. Языковой круг: личность, концепты, дискурс. Волгоград: Перемена, 2002. 477 с.
6. Залевская А.А. Слово в лексиконе человека: Психолингвистическое исследование. Воронеж: Изд-во Воронеж. ун-та, 1990. 208 с.
7. Резанова З.И. Когнитивная лингвистика в парадигмах лингвистического функционализма и интегральных концепций сознания // Вестн. Том. гос. ун-та, 2010. № 334. С. 195–199.
8. Стернин И.А., Розенфельд М.Я. Слово и образ / под ред. И.А.Стернина. Воронеж: Истоки, 2008. 243 с.
9. Большой толковый словарь русского языка / гл. ред. С.А. Кузнецов. СПб.: Норинт, 1998.
10. Мишанкина Н.А. Базы данных в лингвистических исследованиях // Вопросы лексикографии. 2013. № 1 (3). С. 25–34.
11. Методы когнитивного анализа семантики слова: компьютерно-корпусный подход / под ред. В.И. Заботкиной. М.: Языки славянской культуры, 2015 [Электронный ресурс]. URL: http://www.rfbr.ru/rffi/ru/books/o_1924029#1 (дата обращения: 16.07.2016).

12. Nagel O. Investigating Russian Derivational Suffix – yaka: Russian Parallel Corpus Study. // *Procedia-Social and Behavioral Sciences*, 2014, 154. P. 122–129.
13. Ляшевская О.Н., Шаров С.А. Частотный словарь современного русского языка (на материалах Национального корпуса русского языка). М.: Азбуковник, 2009 [Электронный ресурс]. URL: <http://dict.ruslang.ru/freq.php> (дата обращения: 16.07.2016).
14. Paivio A., Yuille J.C., Madigan S.A. Concreteness, imagery, and meaningfulness values for 925 nouns. // *Journal of Experimental Psychology*, 1968, Monograph Supplement, Vol. 76 (1, Pt. 2). P. 1–25.
15. Keuleers E., Balota D.A. Megastudies, crowdsourcing, and large datasets in psycholinguistics: An overview of recent developments. // *The Quarterly Journal of Experimental Psychology*, 2015. Vol. 68, No. 8. P. 1457–1468.
16. Juhasz B.J., Lai Y.-H., Woodcock M.L. A database of 629 English compound words: ratings of familiarity, lexeme meaning dominance, semantic transparency, age of acquisition, imageability, and sensory experience // *Behavior Research Methods*, 2015, 47(4). P. 1004–1019. doi: 10.3758/s13428-014-0523-6.
17. Altarriba J., Bauer L.M., Benvenuto C. Concreteness, context availability, and imageability ratings and word associations for abstract, concrete, and emotion words // *Behavior Research Methods, Instruments, and Computers*, 1999, 31(4). P. 578–602.
18. Bennett S.D.R., Burnett A.N., Siakaluk P.D., Pexman P.M. Imageability and body-object interaction ratings for 599 multisyllabic nouns // *Behavior Research Methods*, 2011, 43. P. 1100–1109. doi: 10.3758/s13428-011-0117-5.
19. Brysbaert M., Warriner A.B., Kuperman V. Concreteness ratings for 40 thousand generally known English word lemmas // *Behavior Research Methods*, 2014, 46. P. 904–911.
20. Tillotson S.M., Siakaluk P.D., Pexman P.M. Body-object interaction ratings for 1 618 monosyllabic nouns // *Behavior Research Methods*, 2008, 40 (4). P. 1075–1078. doi: 10.3758/BRM.40.4.1075.
21. Moreno-Martínez F.J., Montoro P.R., Rodríguez-Rojo I.C. Spanish norms for age of acquisition, concept familiarity, lexical frequency, manipulability, typicality, and other variables for 820 words from 14 living/nonliving concepts // *Behavior Research Methods*, 2014, 46(4). P. 1088–1097. doi: 10.3758/s13428-013-0435-x.
22. Lynott D., Connell L. Modality exclusivity norms for 423 object properties // *Behavior Research Methods*, 2009, 41. P. 558–564. doi: 10.3758/BRM.41.2.558.
23. Lynott D., Connell L. Modality exclusivity norms for 400 nouns: The relationship between perceptual experience and surface word form // *Behavior Research Methods*, 2013, 45. P. 516–526. doi: 10.3758/s13428-012-0267-0.
24. Clark J.M., Paivio A. Extensions of the Paivio, Yuille, and Madigan (1968) norms // *Behavior Research Methods, Instruments, & Computers*, 2004, 36 (3). P. 371–383.
25. Schröder A., Gemballa T., Ruppin S., Wartenburger I. German norms for semantic typicality, age of acquisition, and concept familiarity // *Behavior Research Methods*, 2012, 44. P. 380–394. doi: 10.3758/s13428-011-0164-y.
26. [Электронный ресурс]. URL: <http://www.ruscorgo.ru> (дата обращения: 16.07.2016).
27. Колбенева М.Г., Александров Ю.И. Органы чувств, эмоции и прилагательные русского языка: Лингво-психологический словарь. М.: Языки славянских культур, 2010. 368 с.
28. [Электронный ресурс]. URL: <http://nounobject.ru> (дата обращения: 16.07.2016).
29. [Электронный ресурс]. URL: <http://stimdb.ru> (дата обращения: 16.07.2016).
30. Akinina Y., Malyutina S., Ivanova M., Iskra E., Mannova E., & Dragoy O. Russian normative data for 375 action pictures and verbs // *Behavior Research Methods*, 2015, 47(3). P. 691–707. doi:10.3758/s13428-014-0492-9.
31. Pulvermüller F. Meaning and the brain: The neurosemantics of referential, interactive, and combinatorial knowledge // *Journal of Neurolinguistics*, 2012, Vol. 25, 5. P. 423–459, doi:10.1016/j.jneuroling.2011.03.004.
32. Kiefer M., Pulvermüller F. Conceptual representations in mind and brain: Theoretical developments, current evidence and future directions // *Cortex*, 2012, 48(7). P. 805–825. doi:10.1016/j.cortex.2011.04.006.
33. Connell L., Lynott D., Dreyer F. A Functional Role for Modality-Specific Perceptual Systems in Conceptual Representations // *PLoS ONE*, 2012; 7(3):e33321. doi: 10.1371/journal.pone.0033321. Epub 2012 Mar 13.

34. *Pecher D., Zeelenberg R., Barsalou L.* Verifying Different-Modality Properties for Concepts Produces Switching Costs // *Psychological Science*, 2003, Vol. 14, No. 2. P. 119–124. doi: 10.1111/1467-9280.t01-1-01429.

35. *Gradinarova G., Kanchev P., Janyan A.* Give Me Time to Picture That: Effects of Time and Imageability on Effector-Specific Motor Activation in Idiom Processing // *The Russian Journal of Cognitive Science*, 2014, Vol. 1 (1–2). P. 22–30.

36. *Рахилина Е.В.* Когнитивный анализ предметных имен: семантика и сочетаемость. М.: Русские словари, 2008. 416 с.

37. *Barsalou L.W.* Perceptual symbol systems // *Behavioral and Brain Sciences*, 1999, 22(4). P. 577–609; discussion. P. 610–660.

38. *Величковский Б.М., Зинченко В.П., Лурия А.П.* Психология восприятия. М.: Изд. МГУ, 1973 247 с. [Электронный ресурс]. URL: <http://psychlib.ru/mgppu/vpv/VPV-001-.HTM#> (дата обращения: 16.07.2016).

39. *Rezanova Z., Nekrasova E., Shilyaev K.* Gender-marked metaphors: influence of grammatical gender and animateness on referential choice of metaphorical name of the person in the Russian language // *Procedia – Social and Behavioral Sciences*, 2014, Vol. 215. P. 273–278.

40. *Novikova E.G., Janyan A., Tsaregorodtseva O.V.* Metaphorical Salience in Artistic Text Processing: Evidence From Eye Movement // *Perception*, 2015, 44(8-9). P. 1098-1102. doi: 10.1177/0301006615596692. Epub 2015.

41. *Кубрякова Е.С.* Язык и знание: На пути получения знаний о языке: Части речи с когнитивной точки зрения. Роль языка в познании мира. М.: Языки славянской культуры, 2004. 560 с.

MODELING OF THE PERCEPTUAL-BASED COMPONENT OF LANGUAGE SEMANTICS USING A PSYCHOLINGUISTIC DATABASE

Vestnik Tomskogo gosudarstvennogo universiteta. Filologiya – Tomsk State University Journal of Philology. 2016. 5 (43). 71–92. DOI: 10.17223/19986645/43/6

Zoya I. Rezanova, Tomsk State University, Tomsk Polytechnic University (Tomsk, Russian Federation). E-mail: rezanovazi@mail.ru

Aleksey A. Miklashevsky, Tomsk State University (Tomsk, Russian Federation). E-mail: armanster31@gmail.com

Keywords: databases, cognitive semantics, perceptual-based knowledge.

In Russian cognitive linguistics, a point of view is widely accepted according to which the core information of a concept is concrete-imagery (perceptual-based), acquired first of all through bodily experience and interaction with the external world. This is true for concrete concepts, as well as for abstract ones.

At present there is no prevalent linguistic method to extract imagery information a concept contains. In the article, methods proposed by I.A. Sternin and M.Ya. Rosenfeld are considered: the analysis of the dictionary definition and the directed associative experiment. Advantages and disadvantages of the methods are discussed.

The usage of psycholinguistic databases is proposed, including the scaling method (semantic differential, Likert scale), in order to describe the concrete-imagery (perceptual-based) semantic component. The classification of the psycholinguistic scales used in similar projects (e.g., imageability, modality rating, manipulability, age of acquisition etc.) is presented. The description of the project of the psycholinguistic database which is currently being developed at Tomsk State University is given: 506 Russian nouns of different semantic categories are rated by participants by strength of their relation to five perceptual modalities (modality rating), imageability, spatial localization on the vertical axis, ability to interact with the object with participants' hands (manipulability), and other scales. Preliminary conclusions have been made: it is revealed that relation of the words of the sample to visual modality is the strongest one; relation to haptic modality is less strong, followed by relation to auditory modality. The relation to gustatory and olfactory modalities is the weakest one. It is discovered that, in general, the words of the sample have a relatively high level of imageability, which is in line with findings of studies with other languages. It is shown that words of different semantic categories differ in their level of imageability. The strongest positive correlations were found between relation to visual modality and word imageability, as well as between olfactory and gustatory components of word semantics; negative correlation between word imageability and its age of acquisition was revealed.

Advantages and restrictions of the proposed method are discussed: numeric format of semantic information allows applying statistical methods to solving semantic problems, though one cannot judge on qualitative parameters of conceptual structures. The method of psycholinguistic databases creating is considered as a complementary one to the directed associative experiment, because it gives an opportunity to present quantitative parameters of concrete-imagery (perceptual-based) information in the word semantic structure.

References

1. Vorkachev, S.G. (2004) *Schast' e kak lingvokul'turnyy kontsept* [Happiness as a linguocultural concept]. Moscow: Gnozis.
2. Nikitin, M.V. (2004) *Razvernutyte tezisy o kontseptakh* [Expanded theses on concepts]. *Voprosy kognitivnoy lingvistiki*. 1. pp. 53–64.
3. Popova, Z.D. & Sternin, I.A. (2007) *Kognitivnaya lingvistika* [Cognitive linguistics]. Moscow: AST: Vostok-Zapad.
4. Boldyrev, N.N. (2001) *Kognitivnaya semantika: Kurs lektsiy po angliyskoy filologii* [Cognitive semantics: A course of lectures on English philology]. 2nd ed. Tambov: Tambov State University.
5. Karasik, V. (2002) *Yazykovoy krug: lichnost', kontsepty, diskurs* [Language circle: Person, concepts, discourse]. Volgograd: Peremena.
6. Zalevskaya, A.A. (1990) *Slovo v leksikone cheloveka: Psikholingvisticheskoe issledovanie* [The word in the lexicon of the person: A psycholinguistic research]. Voronezh: Voronezh State University.
7. Rezanova, Z.I. (2010) Cognitive linguistics in the paradigms of linguistic functionalism and integral concepts of cognition. *Vestnik Tomskogo gosudarstvennogo universiteta – Tomsk State University Journal*. 334. pp. 195–199. (In Russian).
8. Sternin, I.A. & Rozenfel'd, M.Ya. (2008) *Slovo i obraz* [Word and image]. Voronezh: Istoki.
9. Kuznetsov, S.A. (ed.) (1998) *Bol'shoy tolkovyy slovar' russkogo yazyka* [Great Dictionary of the Russian Language]. St. Petersburg: Norint.
10. Mishankina, N.A. (2013) Databases in linguistic research. *Voprosy leksikografii – Russian Journal of Lexicology*. 1 (3). pp. 25–34. (In Russian).
11. Zabotkina, V.I. (ed.) (2015) *Metody kognitivnogo analiza semantiki slova: komp'yuterno-korpusnyy podkhod* [The methods of cognitive analysis of the semantics of the words: computer-corpus approach]. Moscow: Yazyki slavyanskoy kul'tury. [Online] Available from: http://www.rfbr.ru/rffi/ru/books/o_1924029#1. (Accessed: 16th July 2016).
12. Nagel, O. (2014) Investigating Russian Derivational Suffix – yaka: Russian Parallel Corpus Study. *Procedia - Social and Behavioral Sciences*, 154. pp. 122–129. DOI: 10.1016/j.sbspro.2014.10.123
13. Lyashevskaya, O.N. & Sharov, S.A. (2009) *Chastotnyy slovar' sovremennogo russkogo yazyka (na materialakh Natsional'nogo korpusa russkogo yazyka)* [Frequency Dictionary of Modern Russian language (on materials of the Russian National Corpus)]. Moscow: Azbukovnik. [Online] Available from: <http://dict.ruslang.ru/freq.php>. (Accessed: 16th July 2016).
14. Paivio, A., Yuille, J.C. & Madigan, S.A. (1968) Concreteness, imagery, and meaningfulness values for 925 nouns. *Journal of Experimental Psychology: Monograph Supplement*. 76 (1:2). pp. 1–25.
15. Keuleers, E. & Balota, D.A. (2015) Megastudies, crowdsourcing, and large datasets in psycholinguistics: An overview of recent developments. *The Quarterly Journal of Experimental Psychology*. 68:8. pp. 1457–1468. DOI: 10.1080/17470218.2015.1051065
16. Juhasz, B.J., Lai, Y.-H. & Woodcock, M.L. (2015) A database of 629 English compound words: ratings of familiarity, lexeme meaning dominance, semantic transparency, age of acquisition, imageability, and sensory experience. *Behavior Research Methods*. 47(4). pp. 1004–1019. DOI: 10.3758/s13428-014-0523-6
17. Altarriba, J., Bauer, L.M. & Benvenuto, C. (1999) Concreteness, context availability, and imageability ratings and word associations for abstract, concrete, and emotion words. *Behavior Research Methods, Instruments, and Computers*. 31(4). pp. 578–602.
18. Bennett, S.D.R. et al. (2011) Imageability and body-object interaction ratings for 599 multisyllabic nouns. *Behavior Research Methods*. 43. pp. 1100–1109. DOI: 10.3758/s13428-011-0117-5
19. Brysbaert, M., Warriner, A.B. & Kuperman, V. (2014) Concreteness ratings for 40 thousand generally known English word lemmas. *Behavior Research Methods*. 46. pp. 904–911.
20. Tillotson, S.M., Siakaluk, P.D. & Pexman, P.M. (2008) Body-object interaction ratings for 1

618 monosyllabic nouns. *Behavior Research Methods*. 40 (4). pp. 1075–1078. DOI: 10.3758/BRM.40.4.1075

21. Moreno-Martínez, F.J., Montoro, P.R. & Rodríguez-Rojo, I.C. (2014) Spanish norms for age of acquisition, concept familiarity, lexical frequency, manipulability, typicality, and other variables for 820 words from 14 living/nonliving concepts. *Behavior Research Methods*. 46(4). pp. 1088–1097. DOI: 10.3758/s13428-013-0435-x

22. Lynott, D. & Connell, L. (2009) Modality exclusivity norms for 423 object properties. *Behavior Research Methods*. 41. pp. 558–564. DOI: 10.3758/BRM.41.2.558

23. Lynott, D. & Connell, L. (2013) Modality exclusivity norms for 400 nouns: The relationship between perceptual experience and surface word form. *Behavior Research Methods*. 45. pp. 516–526. DOI: 10.3758/s13428-012-0267-0

24. Clark, J.M. & Paivio, A. (2004) Extensions of the Paivio, Yuille, and Madigan (1968) norms. *Behavior Research Methods, Instruments, & Computers*. 36 (3). pp. 371–383. DOI: 10.3758/BF03195584

25. Schröder, A. Et al. (2012) German norms for semantic typicality, age of acquisition, and concept familiarity. *Behavior Research Methods*. 44. pp. 380–394. DOI: 10.3758/s13428-011-0164-y

26. Russian National Corpus. [Online] Available from: <http://www.ruscorpora.ru>. (Accessed: 16th July 2016). (In Russian).

27. Kolbeneva, M.G. & Aleksandrov, Yu.I. (2010) *Organy chuvstv, emotsii i prilagatel'nye russkogo yazyka: Lingvo-psikhologicheskii slovar'* [Senses, emotions and adjectives of the Russian language: Linguo-psychological dictionary]. Moscow: Yazyki slavyanskikh kul'tur.

28. The Noun and the Object: Library of Stimuli and Regulations for Experimental Research. [Online] Available from: <http://nounobject.ru>. (Accessed: 16th July 2016). (In Russian).

29. The Library of Stimuli. [Online] Available from: <http://stimdb.ru>. (Accessed: 16th July 2016). (In Russian).

30. Akinina, Y. et al. (2015) Russian normative data for 375 action pictures and verbs. *Behavior Research Methods*. 47(3). pp. 691–707. DOI:10.3758/s13428-014-0492-9

31. Pulvermüller, F. (2012) Meaning and the brain: The neurosemantics of referential, interactive, and combinatorial knowledge. *Journal of Neurolinguistics*. 25:5. pp. 423–459. DOI:10.1016/j.jneuroling.2011.03.004

32. Kiefer, M. & Pulvermüller, F. (2012) Conceptual representations in mind and brain: Theoretical developments, current evidence and future directions. *Cortex*. 48(7). pp. 805–825. DOI:10.1016/j.cortex.2011.04.006

33. Connell, L., Lynott, D. & Dreyer, F. (2012) A Functional Role for Modality-Specific Perceptual Systems in Conceptual Representations. *PLoS ONE*. 7(3):e33321. DOI: 10.1371/journal.pone.0033321

34. Pecher, D., Zeelenberg, R. & Barsalou, L. (2003) Verifying Different-Modality Properties for Concepts Produces Switching Costs. *Psychological Science*. 14:2. pp. 119–124. DOI: 10.1111/1467-9280.t01-1-01429

35. Gradinarova, G., Kanchev, P. & Janyan, A. (2014) Give Me Time to Picture That: Effects of Time and Imageability on Effector-Specific Motor Activation in Idiom Processing. *The Russian Journal of Cognitive Science*. 1 (1–2). pp. 22–30.

36. Rakhilina, E.V. (2008) *Kognitivnyy analiz predmetnykh imen: semantika i sochetaemost'* [Cognitive analysis of subject names: semantics and compatibility]. Moscow: Russkie slovari.

37. Barsalou, L.W. (1999) Perceptual symbol systems. *Behavioral and Brain Sciences*. 22(4). pp. 577–609; discussion: pp. 610–660. DOI: 10.1017/S0140525X99002149

38. Velichkovskiy, B.M., Zinchenko, V.P. & Luriya, A.R. (1973) *Psikhologiya vospriyatiya* [The psychology of perception]. Moscow: Moscow State University. [Online] Available from: <http://psychlib.ru/mgppu/vpv/VPV-001-.HTM#>. (Accessed: 16th July 2016).

39. Rezanova, Z., Nekrasova, E. & Shilyaev, K. (2014) Gender-marked metaphors: influence of grammatical gender and animateness on referential choice of metaphorical name of the person in the Russian language. *Procedia – Social and Behavioral Sciences*. 215. pp. 273–278.

40. Novikova, E.G., Janyan, A. & Tsaregorodtseva, O.V. (2015) Metaphorical Salience in Artistic Text Processing: Evidence From Eye Movement. *Perception*. 44(8-9). pp. 1098–1102. DOI: 10.1177/0301006615596692

41. Kubryakova, E.S. (2004) *Yazyk i znanie: Na puti polucheniya znaniy o yazyke: Chasti rechi s kognitivnoy tochki zreniya. Rol' yazyka v poznanii mira* [Language and knowledge: On the way of learning the language: Parts of speech with a cognitive point of view. The role of language in the knowledge of the world]. Moscow: Yazyki slavyanskoy kul'tury.