

ББК 81.001.4

С.А. Осокина

Лексикографические истоки лингвистической концепции тезауруса

S.A. Osokina

Lexicographic Sources of the Linguistic Conception of Thesaurus

В связи со сложившейся в современной лингвистике необходимостью создания общелингвистической концепции тезауруса предлагается анализ наиболее известных тезаурусных словарей с целью выявления комплекса идей, служащих лексикографическими истоками концепции.

Ключевые слова: тезаурус, слово, языковой знак, значение.

Возникновение идеи формирования лингвистической концепции тезауруса подготовлено развитием научной мысли на протяжении XX–XXI вв. В настоящее время концепция тезауруса развивается в ряде гуманитарных и естественных наук, в рамках которых можно выделить, по крайней мере, три основных толкования ключевого понятия – «тезаурус». Во-первых, под тезаурусом понимается словарь, в котором максимально полно представлены все слова языка и показаны семантические отношения между ними, или идеографический словарь. Во-вторых, тезаурусом именуется семантическая система определенного национального языка или формализованного языка для автоматизированной системы управления. Такое представление мотивировало создание информационно-поисковых тезаурусов и проникновение понятия «тезаурус» в сферу естественных и точных наук. В настоящее время формируется третье, более общее представление, согласно которому тезаурус – это сложнейшая система, содержащая информацию о действительности, а также «метаинформацию» (сведения об информации), обеспечивающую возможность приема новых сообщений.

Поскольку в основе всех трех представлений лежит система языка, это позволило сформулировать предположение, что тезаурус – это объективно существующая языковая данность, функционирующая как информационная система в человеческой среде, определяющая структуры знания отдельного индивида и общества в целом на определенном этапе его развития, приближительной моделью которой является словарь. Данное представление можно считать отправной точкой формирующейся лингвистической концепции

The paper provides an analysis of several published thesauruses in English and Russian languages in order to summarize the ideas which can support the developing linguistic conception of thesaurus.

Key words: thesaurus, word, language sign, meaning.

тезауруса. Однако для окончательного оформления данной концепции в виде самостоятельной теории необходимо тщательно проанализировать имеющиеся предпосылки ее возникновения. Важнейшим источником концепции является комплекс идей, выработанных в ходе создания тезаурусных словарей в таком направлении языкознания, как лексикография.

В истории создания идеографических тезаурусных словарей можно условно выделить три этапа: 1) древние идеографические словари; 2) словари эпохи Нового времени; 3) современные словари, изданные в XX–XXI вв.

Первые попытки построения подобных словарей относятся к II в. н.э. Как считает В.В. Морковкин, «На заре цивилизации, когда люди могли выразить свои мысли на письме при помощи идеограмм и символов, единственно возможным словарем был, вероятно, такой, в котором слова располагались по тематическим группам» [1, с. 10].

Из дошедших до нас известны словари Ю. Поллукса «Ономастикон» (II в.), состоящий из десяти книг, в каждой из которых содержатся слова, относящиеся к той или иной теме, и словарь санскритского языка «Амарокша» (II–III вв.), содержащий 10 тыс. слов, в трех книгах, каждая из которых разбивается на общие и частные темы. Данные словари представляют собой «первые попытки системного представления лексики, базирующиеся на особенностях мировоззрения того времени» [2, с. 66].

В эпоху Нового времени возникает идея создания единого философского языка, который отражал бы общие мыслительные категории человека. Наиболее известным тезаурусом Нового времени, который не

потерял свою актуальность до сих пор, регулярно переиздается и переведен на несколько языков, является «Тезаурус английских слов и выражений» Питера Марка Роже (первое издание – в 1852 г.). Слова в нем разделены на шесть классов, которые, в свою очередь, делятся на логические группы и подгруппы с цифровыми указателями на соответствующие рубрики. Роже не ставил себе целью дать дефиницию слов или описание смысла слов, но считал необходимым привести синонимы и антонимы, так как это помогает найти слово в словаре и служит подсказкой памяти, если невозможно вспомнить, какие еще слова выражают определенное понятие.

Еще один словарь, оказавший воздействие на современные представления об идеографических словарях, – «Аналогический словарь французского языка» П. Буасьера. Принципиальным отличием данного словаря является отказ от логической схемы представления словесного материала и использование алфавитного принципа для облегчения поиска слов – этим составитель стремился преодолеть недостатки тезауруса Роже.

Среди отечественных идеографических словарей, изданных до XX в., известны «Номенклатор, на русском, латинском и немецком языке» И.Ф. Копиевского (1700 г.) и «Краткое собрание имен» Ф.П. Поликарпова (1701 г.). Оба словаря представляют собой тематические переводные тезаурусы с русского на латинский и немецкий (первый словарь) и с русского на латинский и греческий (второй словарь). Лексика в них разделена на тематические группы. Сопоставительный анализ этих словарей между собой с и другими тезаурусами того времени показывает, что нет тождества в количестве выделяемых рубрик, в расположении рубрик в словаре, в наполнении их лексическим материалом. Так, даже применительно к одному языку обнаруживаются значительные расхождения в словесном наполнении рубрик, например, в словаре И.Ф. Копиевского гораздо богаче представлена и более детально классифицирована бытовая лексика русского языка, а в словаре Ф.П. Поликарпова в большем объеме представлена книжная лексика, значительное место занимают классы, описывающие науку, церковь и судопроизводство. Таким образом, по данным соответствующих словарей можно судить, скорее, об уровне образованности и круге интересов авторов-составителей, нежели о специфических системных особенностях лексики русского языка XVII–XVIII вв. [3].

Подводя итог краткому обзору идеографических словарей прошлого, выделим основные идеи, повлиявшие на современные представления о сущности тезауруса.

1. Тезаурусы представляют собой первые попытки осмысления и систематического представления лексики языка. Лексическая система предстает в них как

базирующаяся на строгих логических отношениях, преимущественно родо-видовых, позволяющих выстраивать слова языка в иерархию соподчиненных классов. Иерархия родо-видовых отношений трактуется как основная структура человеческого мышления. Показательно, во-первых, что первыми словарями человеческих языков являются именно словари со строгой иерархией классов слов. Во-вторых, родо-видовые отношения впоследствии были положены в основу дефиниций, предлагаемых в толковых словарях. Итак, система лексики языка – это иерархическая система понятий и соответствующих слов.

2. В ходе истории создания тезаурусных словарей сформировались три основных типа тезаурусов: 1) тематические словари, в которых представлены слова, объединенные некоей общей темой; 2) собственно тезаурусы, или идеографические словари, отражающие ход мысли от общих мыслительных категорий и понятий к словам языка; 3) аналогические словари, совмещающие алфавитный принцип изложения слов и разбиение слов по смысловым группам. Данные типы, по сути, отражают ход исследовательской мысли по осознанию устройства лексики вообще: сначала от частного к общему (от объединения конкретных слов в темы и дальнейшего осмысления стоящих за темами философских категорий), а затем, вновь, к отдельному слову как центральной единице, от которой следует отталкиваться при поиске других слов.

3. Способы представления слов в тезаурусах (разбиение на классы с соответствующим буквенным или цифровым индексированием) предопределило развитие идеи изучения языка в формальных категориях логики, идеи записывания выражений языка на формальном языке логики и идеи создания искусственных формальных языков, которые впоследствии привели к появлению компьютерных языков и попыткам конструирования систем искусственного интеллекта.

Однако история создания тезаурусных словарей позволила выявить и ряд существенных проблем методологического характера.

1. Проблема соответствия смысла и слова, которая может быть сформулирована в виде вопроса: действительно ли в номинациях рубрик идеографических словарей дается идея, или «концепт», в соответствии с современной номенклатурой? Во-первых, концепт – сущность идеальная, и не может присутствовать на печатных страницах в принципе. Допустим, заглавия рубрик – это «имена» идей, тогда чем оправдан тот факт, что одни и те же идеи называются разными именами? Данная проблема имеет глубокие философские корни и к началу XX в. стала интерпретироваться как семиотическая проблема соотношения знака и значения.

2. Проблема членимости объективного мира, вытекающая из предыдущей, встает в связи с несопадением рубрик и подрубрик в разных тезаурусах.

Если цель тезауруса – рассредоточить слова языка в соответствии с наиболее общими категориями мышления, то на основе сопоставления словарей возникает мысль об отсутствии объективных универсальных общечеловеческих мыслительных категорий. На каком основании в таком случае производится выделение заглавных категорий в словарях: мотивировано ли их выделение лексикой самого языка или субъективными доводами составителей словарей? За данными вопросами фактически встает одна из главных проблем философской эпистемологии – проблема объективности знания, которая отрицательно решается в философии сенсуализма (Беркли, Юм, Кант), приведшей к появлению феноменологии, а также в философских воззрениях В. фон Гумбольдта, оказавших колоссальное влияние на лингвистику XX в. В трудах последнего объективное знание невозможно не в силу различности субъективного чувственного восприятия мира, как в сенсуализме, а в силу того, что между человеком и реальностью стоит третий мир – мир языка, фактически определяющий то, каким реальность дана человеку. Вопрос об объективности знания затрагивает фундаментальные проблемы соотношения сознания и реальности, с одной стороны, и языкового знака и реальности – с другой.

3. Проблема направления логического поиска встает в связи с несовпадением словесного наполнения рубрик (и также вытекает из предыдущих). Одной из целей создания тезаурусных словарей является упрощение поиска нужного слова в соответствии с заданным понятием и, наоборот, возведение конкретного слова к абстрактному философскому понятию. Такая задача декларировалась, к примеру, в словаре П. Роже, однако его первым изданием чрезвычайно трудно пользоваться в силу чрезмерной формальности отнесения слов в ту или иную категорию. Так, например, чтобы найти слова, связанные со словом *музыка*, необходимо сначала догадаться, что само это слово надо искать в категории «материя», а не, предположим, в «чувства», хотя семантика слова *музыка* вполне допускает его включение в последнюю категорию. Данная проблема коренится в более общем методологическом вопросе о соответствии логических формальных связей и связей языковых, логического и естественного языков.

4. Системное представление лексики – это одна из основных методологических проблем языкознания. В явном виде в тезаурусных словарях, изданных до XX в., представлена только иерархическая организация лексического состава языка. Это обусловлено тем, что данные словари имеют своей основной целью упорядоченное представление мыслительных категорий, а не собственно слов языка, однако история создания тезаурусов, несомненно, подвела к постановке вопроса о системном характере лексики. Дальнейшее развитие лексикографии

и лексикологии в XX в. в целом сопряжено с изучением именно этой проблемы.

К началу XX в. обозначилось два подхода к системному изучению лексики и ее представлению в словарях. Основанием послужила актуальная в то время проблема определения слова как центральной единицы языка.

В рамках первого подхода слово трактуется как самостоятельная единица языка, соответственно, лексическая система представляет собой систему слов, находящихся в определенных отношениях, но акцент ставится на самом слове, нежели на отношениях между словами. Этот подход связан в лингвистике с деятельностью Пражского лингвистического кружка; в отечественном языкознании этот подход имеет место в работах А.А. Потебни (психологизм его концепции), В.В. Виноградова и их последователей. Поскольку в центре внимания оказывается слово, основное внимание уделяется исследованию его внутренней семантической структуры, его полисемии. Внешние связи слова, контекст рассматриваются как средство уточнения значений слова.

Представители второго подхода, напротив, считают самостоятельность слова умозрительной, даже иллюзорной, слово будто «растворено» в своих парадигматических и синтагматических связях, следовательно, чтобы понять, что представляет собой слово, необходимо изучать то, что дано в непосредственном наблюдении, – его связи с другими словами. Данный взгляд в лингвистике связан с трудами Ф. де Соссюра (социологизм его концепции) и его учеников. Основное внимание уделяется изучению внешних связей слова, от которых зависят его семантические границы. В рамках данного подхода в трудах Йоста Трира стала развиваться идея существования семантических полей слов. В отечественном языкознании полевой подход долгое время игнорировался, но позднее прочно занял свою нишу, благодаря работам Г.С. Щура.

Развитие концепции тезауруса как словаря, репрезентирующего лексическую систему языка, как убедительно показывает Ю.Н. Караулов, позволяет совместить оба подхода. По его мнению, тезаурус представляет собой систему взаимосвязанных полей слов, но в их взаимосвязи прослеживается и внутренняя структура слова, заложенная в дефиниции, которая является основным объектом изучения и конструирования у представителей словоцентристского подхода. Дефиниция, во-первых, предоставляет материал для создания полей, а во-вторых (и это главное), дефиниция воспроизводится в самой структуре тезаурусного словаря: тезаурус представляет собой иерархию семантических полей, а сама иерархия от родового понятия к видовым – не что иное, как сущность дефиниции [4, с. 64].

Идеи указанных двух подходов к осмыслению лексической системы языка отражены в концепциях

тезаурусных словарей XX в. Для того чтобы детально охарактеризовать, какой является лексическая система языка, представленная в виде тезауруса, обратимся к анализу новейших тезаурусных словарей. Поскольку образцом тезаурусного словаря (на который составители практически всех современных тезаурусов не только ссылаются, но и ставят обязательной задачей преодоление его недостатков) является «Тезаурус английских слов и выражений» Питера Роже, правомерно начать исследование тезаурусного способа представления лексики с анализа словарей английского языка как имеющих более богатую традицию тезаурусного осмысления лексики.

Из трех типов тезаурусных словарей, сложившихся в ходе истории их создания, большинство современных тезаурусов английского языка являются словарями аналогического типа, т.е. иерархическая система понятий отражена в них внутри группировок слов, изложенных в алфавитном порядке, кроме «Нового тезауруса Роже», который является последним переизданием оригинального словаря XIX в.

«**New Roget's Thesaurus**» издания 1993 г. представляет собой классический идеографический словарь, в котором лексика языка распределена в соответствии с логическими категориями, каждая из которых делится на различное число зависимых категорий. Найти нужное слово в таком словаре чрезвычайно трудно, если не воспользоваться алфавитным указателем слов, приведенным в конце. Так, слово *light* фигурирует в следующих рубриках с указанием индекса рубрик *luminosity* (свечение) 420, *speed* (скорость) 274, *levity* (ветреность, легкомысленность) 320, *to kindle* (зажигать) 384, *unencumbered* (необремененный, свободный) 705, *small* (маленький) 32, *trifling* (несерьезный) 645, *gay* (веселый) 836, *idea* (идея) 453, *knowledge* (знание) 490, *to arrive* (прибывать) 292, *loose* (широкий) 961.

На первый взгляд, это группа практически не связанных слов, которую можно проинтерпретировать как поле, но этот набор слов отражает скорее периферию лексического поля «light», нежели непосредственно связанные со словом *light* слова. Более того, данный перечень связей слова *light* соответствует двум омонимам английского языка: *light*¹ – «освещение, дневной свет» и *light*² – «легкий на вес». Таким образом, в предложенном выше перечне слов обнаруживается наложение, как минимум, двух крупных омонимичных полей, и, только вникая в сложную иерархию логических категорий, можно определить, откуда взялось каждое слово в данном списке.

В словаре Роже настолько много внимания уделяется организации понятийной стороны, что словесная сторона оказывается представленной очень раздробленно; анализируя ее, можно сделать лишь приблизительный набросок семантических потенциалов отдельного слова, многие промежуточные связи

между словами приходится домысливать. Искомое слово может входить во все шесть исходных логических категорий и встречаться в словаре в разных местах; вместе с тем под одним подзаголовком могут оказаться слова, никак не связанные семантически, только имеющие одинаковую материальную оболочку. Напрашивается вывод, что при подходе «от понятий к словам» лексемы предстают как «голые» знаки, потенциально могущие выражать бесконечный смысл.

Преодолеть этот недостаток попытались составители тезаурусов английского языка, строящихся по принципу «от слов к словам». К ним относятся, например, «The Oxford Paperback Thesaurus», «The Oxford Thesaurus», «American Edition», «The Cambridge Thesaurus of American English» и «Merriam-Webster's Collegiate Thesaurus». Эти словари построены по синонимическому принципу: приводится ряд синонимов к заглавным словам, которые помещаются в алфавитном порядке. Для определения точности синонимов предлагается отталкиваться от контекста: дается пример употребления слова в контексте, а затем следует список синонимов, которые могут заменять заглавное слово в данном контексте.

Идея использовать в качестве подсказки контекст очень плодотворна, она позволяет включить парадигму синонимов в синтагматическую последовательность слов. Однако замена заглавного слова на синоним в предлагаемых контекстах, как правило, предполагает модификацию контекстов, что вызывает вопрос, можно ли считать предлагаемые синонимы абсолютно взаимозаменяемыми, и поднимает проблему сущности синонимии вообще.

Еще одно важное замечание: в словарях наблюдается различие в наборе синонимов, приводимых к конкретному слову. Причина несовпадения набора синонимов, на наш взгляд, кроется в том, что авторы словарей имеют различные представления о семантических границах отдельных слов. Так, в словаре «The Oxford Paperback Thesaurus» в качестве синонимов к слову *light* приводятся *sunlight* и *moonlight*, но нет слова *starlight*, имеется *lamplight*, но нет *candlelight* и *torchlight*, хотя эти слова фигурируют как синонимы слова *light* в других тезаурусных словарях. Это свидетельствует о том, что выбор исследователя субъективен и основывается на внутреннем опыте и чувстве языка в тех случаях, когда нет очевидных объективных причин для включения или невключения того или иного слова в словарь.

Поскольку все синонимические тезаурусы идут от слов к словам, в них репрезентируется замкнутая внутри себя лексическая система, в которой нет выхода на концепты, или идеи. Если в чем-то в этих словарях и прослеживается продолжение традиций Роже, так это в том, что слово в них тоже предстает как знак, не соотношенный с определенным значением, а обнаруживающий это значение в своих связях.

Следовательно, и при подходе «от понятий к словам», и при подходе «от слов к словам» значение слова ускользает. В первом случае значение исчезает, так как система понятий не соответствует системе слов: нет такого, чтобы определенное понятие соотносилось только с одним конкретным словом и больше ни с каким, каждому понятию соответствует целый ряд слов; вместе с тем одно и то же слово (материальный знак) может соответствовать нескольким понятиям. В втором случае значение исчезает потому, что, если отталкиваться от значений, то к каждому конкретному слову практически невозможно будет подобрать синоним, так как абсолютные синонимы, если они вообще возможны, — это исключение, а не правило языка.

Парадокс в том, что при первом подходе выход на понятийный уровень заявлен в словаре (точнее, с него надо спускаться на уровень слов), но значение как феномен обнаруживает себя еще в более рассеянном виде, чем в синонимических словарях, в которых выход на понятийный уровень даже не заявлен, но можно более явно представить себе значение как феномен в силу большей компактности представления словесных связей (благодаря сконцентрированности связей под заглавным словом, а не понятием).

Возникает вопрос о связи значения с понятийным уровнем и вообще о необходимости и целесообразности вводить значение в поле лингвистического рассмотрения. По крайней мере, при том подходе к представлению лексической системы, который присутствует в тезаурусных словарях, значение как лингвистическая величина оказывается ненужным. Этот вывод согласуется с семиотическими воззрениями Р. Карнапа, рассматривающего языковой знак как элемент системы языка, значением которого может быть только его место в этой системе.

Итак, тезаурус в рассмотренных словарях английского языка предстает как система словесных связей. Репрезентируются связи парадигматические (преимущественно синонимические) и синтагматические (контекстуальные, идиоматические). В этой системе не важно, какое значение имеет слово, а важно только, с какими другими словами его можно поместить в один ряд. Эта концепция строилась, казалось бы, с учетом словоцентристского подхода для преодоления недостатков тезауруса Роже, но на деле вылилась в представление о взаимосвязанных словесных знаках, растворяющих отдельное слово в сложной системе собственных отношений с другими словами.

В отечественной лексикографии в настоящее время существует множество словарей, которые можно было бы отнести к тезаурусам, однако далеко не все из них имеют такое название. Правомнее было бы говорить о множестве словарей, альтернативных по отношению к толковым: словари синонимов, антонимов, сочетаемости, частотные, семантические словари, тематические словари и т.д. Их главное от-

личие от толковых состоит в том, что они не ставят самоцелью формулирование дефиниций, хотя могут их использовать.

Основываясь на опыте создания тезаурусов на материале английского языка, можно сказать, что указанные типы словарей представляют собой отдельные части тезауруса русского языка. Возможно, отечественные лексикографы отказываются от создания комплексного тезаурусного словаря, потому что зарубежные издания не дают четкого представления о том, какие словесные связи необходимо раскрывать в тезаурусных словарях, европейской лексикографией и лексикологией не выработан их общий перечень.

Единственным словарем русского языка, созданным по типу тезауруса Роже, является «Идеографический словарь» О.С. Баранова, изданный в 1995 г. «Целью создания Идеографического словаря русского языка, издаваемого впервые, — пишет О.С. Баранов, — является представление лексики русского языка в систематизированном виде, в соответствии со смыслом слов для уяснения места каждого понятия среди других, выделения основных тем и закономерностей» [5, с. 30].

Словарь делится на 17 больших отделов («Отношение», «Количество», «Зависимость», «Пространство», «Процесс», «Организованные системы», «Материя», «Органический мир» и т.д.), которые делятся на разделы и подразделы. Однако, как вытекает из поставленной цели, автор не задавался идеей поиска слова исходя из системы понятий; напротив, он стремился изложить систему лексики так, чтобы за ней встала система понятий — смыслов, тем и закономерностей отношений между ними. Следовательно, если П. Роже отталкивался от всеобщего философского языка, то Баранов выводит систему понятий из лексики русского языка.

Если это так, то вполне понятно, почему в тезаурусе Роже и в словаре Баранова не совпадают количество и наполнение отделов, почему как равноправные отделы фигурируют такие, например, как «Отношения» и «Зависимость» (с точки зрения логики было бы правомее включить «Зависимость» в отдел «Отношения» как подраздел), или «Материя» и «Органический мир» — очевидно, лексика русского языка располагает таким набором слов, который позволяет рассматривать указанные пары как равноправные отделы. Аналогично в словаре имеется отдел «Техника», но нет отдела «Наука» — тоже, может быть, в связи с тем, что в русском языке много слов для обозначения техники, так что они заслуживают их рассмотрения в качестве самостоятельного отдела, а слов с общей семантикой «наука» меньше, или они редки в употреблении, или еще по каким-то причинам не выделены в самостоятельный отдел. Иначе говоря, по поводу разбиения лексики русского языка на отделы, представленные в словаре, возникает много вопросов,

и если выделение отделов не мотивировано самим содержанием лексики, то никакой логики в этом выделении не прослеживается.

В 2002 г. словарь переиздан со значительными изменениями. Теперь в нем шесть отделов – «Порядок», «Природа», «Человек», «Общество», «Деятельность», «Культура» – распадающихся на разделы и подразделы, расположение которых внешне выглядит гораздо более логично, но вопрос о мотивации выделения разделов остается открытым.

Понять содержание словарной статьи можно только после длительного изучения правил построения словаря. Найти конкретное слово можно только по алфавитному указателю в конце. Статья с заголовком «Свет» имеет такое содержание (с указанием индексов и символических обозначений): 47.1.1. *виды света, Δ электромагнитные волны 179, свечение, огонь, пламя, каление (белое #), искра (искры летят, сыпятся), озарять, -ся, озаренный, подсвечивать, просвечивать, сквозить, -ся, фото... (фотосинтез), оптика, оптический, ↓ геометрическая оптика, 216 физическая оптика, 47.1.2. сияние, ↑ кайма 104, ореол – светлая кайма вокруг светящегося, раскаленного предмета, заря, зарево, зарница, сполохи* и т.д.

Связь с лексическим значением слова *свет* или его употреблением в контекстах практически не обнаруживается. Системное представление лексики подчинено в первую очередь раскрытию соответствующего научного понятия. Многие слова, приведенные в разделе, даны в других словарях с пометой «*спец.*» и не входят в активный словарный состав русского языка.

«Тематический словарь русского языка» Л.Г. Саяховой, Д.М. Хасановой и В.В. Морковкина соответствует традиционной структуре тезауруса и соотносится со структурой «Идеографического словаря» О. Баранова (номинации разделов, конечно, не совпадают). Распределение слов по темам свидетельствует, что в основе действительно лежит лексика языка, а не абстрактные научные понятия. Содержание статьи актуализирует синтагматические связи слова, например: *свет, -а, предл. в свете, на свету, только ед., м.; яркий ~, солнечный ~, дневной ~, лунный ~, ~ солнца, ~ луны, ~ зари, ~ звезды, источник -а, луч -а, испускать ~* и т.д.

По сути, статья представляет собой фрагмент словаря сочетаемости, и это принципиально важно, так как сочетательные возможности гораздо более объективно раскрывают семантику конкретного слова, чем синонимы или антонимы. Если в тезаурусах английского языка устойчивые сочетания слов использовались как дополнительный способ представления словесных связей, то в тематическом словаре русского языка устойчивые синтагматические связи рассматриваются как ведущие для характеристики слова в составе конкретной темы. Это показывает, что в основе тематической иерархии слов лежит син-

тагматика словесных выражений, и вовсе не факт, что иерархический принцип необходимо рассматривать как основополагающий.

Стоит упомянуть и «Частотный словарь семантических множителей русского языка» Ю.Н. Караулова (1980 г.), так как содержание этого словаря было положено в основу первого тезауруса русского языка, сделанного при помощи ЭВМ – «Русского семантического словаря» (1982 г.). Под семантическими множителями понимаются «единицы содержательного плана, которые, соотносясь друг с другом в различных количествах и комбинациях, задают значения любого слова в языке» [6, с. 4]. В качестве множителей предстают единицы, внешне напоминающие морфемы, но не всегда тождественные им, например «взаим», «помощь», «поним», «дух», «близ» и т.д. Создание данного словаря явилось результатом работы специалистов в области информатики и лингвистов, вдохновленных идеями Ю.Н. Караулова, который развивал концепцию тезауруса русского языка на протяжении нескольких лет.

В середине прошлого века языковые тезаурусы стали использоваться как инструмент для создания информационно-поисковых тезаурусов в различных сферах деятельности. Стали появляться словари, которые были не просто перечнем слов, а источником сведений по определенному вопросу. Развивая идею, что языковой тезаурус представляет собой лингвистический конструкт, воссоздающий лексическую систему языка в виде семантических связей слов, что позволяет использовать его как систему поиска нужной информации, Ю.Н. Караулов приходит к мысли, что словом «тезаурус» можно назвать и когнитивную систему поиска информации. Так, в работе «Русский язык и языковая личность» 1987 г. тезаурусом называется когнитивный уровень языковой личности, или «личностный тезаурус» [7, с. 52–53]. И в этом смысле тезаурус также является конструктом (языковая личность в целом – это модель), не дающим непосредственного выхода на концептуальную систему, но, используя который, можно строить предположения о сущности последней. Эта идея послужила основой создания тезаурусного словаря принципиально нового типа, аналогов которому нет в европейской лексикографии, – «Русского ассоциативного словаря» с подзаголовком «Ассоциативный тезаурус русского языка».

Отличие данного словаря от всех ранее рассмотренных состоит в том, что в нем система словесных связей выводится не путем анализа лексики языка, а путем массового психолингвистического эксперимента. Содержание словаря представляет собой словесные реакции испытуемых на предлагаемые словестимулы. Зафиксированные таким образом связи слов, по мысли составителей словаря, являются моделью «речевых знаний носителей русского языка, представленных в виде ассоциативно-вербальной сети,

позволяющей объяснить феномен владения языком и служащей – наряду с текстовым и системным – способом представления русского языка» [8, с. 6].

Такой подход предполагает раскрытие словесных связей не напрямую от слова к слову, а от слова к мысли, и только затем ко второму слову. Это как бы развитие подхода «от слова к слову», но с экспликацией того, что было в промежутке, в то время как подход от «слова к слову» выдает словесные связи как уже готовый результат системного существования лексики.

При таком способе репрезентации языка можно проследить все значения полисемантических слов, синонимические и антонимические ряды, синтаксическую сочетаемость, словоизменительную и словообразовательную вариативность. В словаре отражены наиболее частотные семантические связи слов (указан индекс частотности) и, прежде всего, «наиболее привычные для носителей русского языка “модели двух слов” (Н.Н. Жинкин) – стандартные повторы, клише, устойчивые словосочетания, идиомы» [8, с. 8].

В ходе анализа словаря установлено, что наиболее частотные словесные реакции испытуемых представляют собой не синонимы, антонимы или слова, более общие по значению (как в дефинициях), а реакции синтагматического типа, например, наиболее частотные реакции на слово *магазин* – «*пустой*» и «*закрыт*». Из этого следует, что при рассмотрении словесных связей, представленных в виде ассоциативного тезауруса, более важными оказываются синтагматические связи. Поскольку ассоциативный тезаурус именуется «моделью речевых знаний», то напрашивается вывод, что синтагматические связи претендуют на роль основных эпистемологических связей, а построенный на их основе тезаурус – на роль эпистемологической модели языка.

Анализ тезаурусных словарей XX в. позволяет выделить ключевые идеи, которые подводят к ответу на вопрос, что такое тезаурус, и могут быть положены в основание формирующейся лингвистической концепции тезауруса.

Тезаурус – это прежде всего система. Система чего? Возможны варианты: система лексики, эксплицирующая только словесные связи, система понятий, отражающая состояние знания (в том числе научного) на определенном этапе; система знания, имеющаяся в голове человека и обнаруживаемая путем психолингвистического эксперимента, – так или иначе, это словесная информационная система.

Тезаурус – это семиотическая система, элементы которой – языковые знаки. Соответственно, по аналогии с фундаментальными вопросами семиотики о соотношении знака, реальности и ментальности, встают вопросы: как существует реальность в тезаурусе (или тезаурус в реальности?) и как существует тезаурус в голове человека?

Тезаурусы, созданные по типу словаря Роже, наводят на мысль, что объективное существование системы понятий – само собой разумеющееся. Однако несовпадение выводимых в таких словарях понятийных систем заставляет усомниться в правильности такого решения. Тезаурусы, идущие от слова к слову, предполагают феноменологическое основание: невозможно восстановить объективную систему понятий, поэтому восстанавливается только то, что реально дано наблюдателю – словесные связи. Данный подход более обоснован, однако несовпадение перечня связей для отдельного слова в разных словарях свидетельствует о значимости фигуры исследователя-составителя и о важности его субъективного восприятия релевантности связей. Наконец, тезаурус, раскрывающий словесные реакции испытуемых в ходе психолингвистического эксперимента, еще более абсолютизирует субъективность видения словесных связей, так как предлагаются словесные реакции конкретных людей. Создатели словаря пытаются преодолеть проблему субъективизма полученных данных частотным анализом словесных реакций, но вопрос о соотношении реальности и тезауруса остается открытым.

Вопрос о том, как существует тезаурус в голове человека, не имеет четкого ответа, если опираться только на данные проанализированных словарей, однако можно выделить несколько «подсказок», в каком направлении стоит двигаться, изучая данную проблему.

Во-первых, все тезаурусные словари (как и словари вообще) служат инструментом поиска информации. Однако словарь сам ничего не ищет – ищет пользователь словаря по заданному алгоритму. Если поиск затрудняется (а это один из главных недостатков идеографических словарей), значит, алгоритм поиска, предложенный в словаре, не совпадает с алгоритмом поиска, из которого исходит пользователь. Следовательно, можно предположить, что тезаурус в голове человека представляет собой информационную систему с определенным алгоритмом поиска. Каков этот алгоритм – пока неясно. Алгоритм поиска по алфавиту – изобретение лексикографов – возможно, соотносится с внутренним алгоритмом поиска, поскольку аналогичными тезаурусами действительно пользоваться легче, чем идеографическими, однако вряд ли стоит думать, что слова укомплектованы в голове человека в алфавитном порядке (на это указывает тот факт, что трудно вспомнить все слова, например, на букву «а»). Однако в обращении к поиску по алфавиту есть и рациональное зерно: располагая слова по алфавиту, составитель словаря исходит только из материальной оболочки слова – можно предположить, что и алгоритм поиска слов в голове человека связан с их материальной оболочкой, а не только со значением, как принято думать.

Во-вторых, поиск информации всегда связан с категорией памяти, не случайно авторы тезаурусных словарей пишут, что предлагаемые издания необходимы для извлечения нужного слова из памяти. Что собой представляет «память на слова», пока сказать трудно, но, очевидно, она базируется на связях слов, поскольку тезаурусные словари предлагают помощь в извлечении слов из памяти путем представления их связей с другими словами. Авторы словарей предлагают две метафоры для определения совокупности словесных связей: «сеть» и «процессор». Думается, эти метафоры весьма эффективны для дальнейшего осмысления идеи памяти.

Возвращаясь к представлению о тезаурусе как о системе, затронем еще один важный вопрос – об элементах и связях этой системы. Что касается элементов, то предварительно можно ответить, что это слова, но слова, растворенные в своих связях. Понятые так слова не существуют как самоценности, их ценность определяется количеством и качеством их связей. Здесь составители словарей обращаются к частотности как формальному показателю стабильности связей, на который можно опереться и при отборе и заглавных слов, и слов, составляющих содержание статьи.

Вопрос о качестве тезаурусных связей наиболее трудный. Анализ словарей показывает, что это могут быть связи собственно словесные и понятийные. По поводу того, какие из них главные, а какие вспомогательные, можно ответить двояко. Если изначально встать на позицию признания существования объективных понятийных связей, то словесные связи – только средство их выражения. Однако структура представления понятийных связей в идеографических словарях наводит на мысль об отсутствии четкой логики этих связей, которая неизбежно предполагается самим фактом их объективности. Соответственно, возникает второй вариант ответа: возможно, иерархия понятийных связей не

существует объективно, а мотивирована словесными связями языка, преломленными через их субъективную оценку составителем словаря (главным образом на основе критерия частотности).

Что касается собственно словесных связей, то, исходя из содержания проанализированных тезаурусных словарей, можно сказать, что в сети тезауруса слова могут быть объединены любыми известными лексическими связями (парадигматическими и синтагматическими), а также ассоциативными связями. Причем трудно сказать, какие из них более важны в системе тезауруса, поскольку в одних словарях раскрываются преимущественно синонимические связи, в других – преимущественно контекстуальные, в третьих – ассоциативные и т.д. Логика презентации связей в словарных статьях либо воссоздает структуру дефиниции (в широком смысле как иерархию значений или мыслительных категорий), либо соответствует полевому принципу. Ю.Н. Караулов полагает, что связи в идеографическом словаре – это семантические связи, однако, определяя их сущность, он пишет: «Говоря о “семантической связи”, мы опираемся на интуитивно очевидное представление носителей языка о том, что это такое» [4, с. 76]. Главное назначение таких связей – обеспечивать «непрерывность» словаря, особенно «словаря в голове человека» [4, с. 78]. Вопрос о тезаурусных связях требует отдельного глубокого осмысления.

Высказанные мысли – пока предварительные соображения о том, что такое тезаурус, основанные на анализе ряда лексикографических источников. Лингвистическая концепция тезауруса требует дальнейшего развития, так как обладает мощным экспланаторным потенциалом, с которым связываются современные устремления объяснить сущность человеческого знания и предложить качественно новое видение фундаментальной языковедческой проблемы соотношения языка и мышления.

Библиографический список

1. Морковкин В.В. Идеографические словари. – М., 1970.
2. Щербинин В.К. Вселенная в алфавите. Словари – вчера, сегодня, завтра. – Минск, 1987.
3. Березина О.Е. Два тематических лексикона начала XVIII века (Сравнительная характеристика) // Словари и словарное дело в России XVIII века. – Л., 1980.
4. Караулов Ю.Н. Общая и русская идеография. – М., 1976.
5. Баранов О.С. Идеографический словарь русского языка. – М., 1995.
6. Караулов Ю.Н. Частотный словарь семантических множителей русского языка. – М., 1980.
7. Караулов Ю.Н. Русский язык и языковая личность. – М., 2007.
8. Русский ассоциативный словарь. Кн. 1: Прямой словарь: от стимула к реакции. Ассоциативный тезаурус современного русского языка. – Ч. I. – М., 1994.