

© 1994 г. Е.В. РАХИЛИНА

О ЛЕКСИЧЕСКИХ БАЗАХ ДАННЫХ

Как ни странно, история развития лингвистики – вообще говоря, гуманитарной науки – оказалась необычайно тесным образом связана с историей информатики и вычислительной техники. Общеизвестно, что идея машинного перевода, возникшая в конце пятидесятих годов в "пик" кибернетического бума, дала совершенно новый импульс лингвистической науке (причем, как теперь очевидно, прежде всего теоретической). Конечно, некоторые области прикладной лингвистики просто не могли возникнуть до рождения первых вычислительных машин, но гораздо более неожиданно, что при этом совершенно изменилась и, так сказать, общая парадигма науки – лингвистика стала казаться **точной наукой**. Появились и новые требования к лингвистическим работам – они вторили математике: определения должны быть эффективными, описания непротиворечивыми, и т.д. Неважно, что эти требования не выполнялись, и, как стало ясно гораздо позже, не могли выполняться, – важно, что они были, оставили свой след и даже в каком-то виде сохранились до сих пор именно в качестве критериев научности.

Постепенное (а в нашей стране скорее внезапное) появление новых машин – быстрых и вместе с тем обладающих огромными возможностями памяти – опять открывает новые прикладные области для лингвистики: прежде всего это базы данных, позволяющие классифицировать языковые факты. Снова потенциал машин дает надежду (или иллюзию), что с его помощью можно бесстрашно переплыть бесконечное лексическое море, море употреблений, различного рода сочетаний и т.п.: в машине бесконечное превращается в счетное и кажется обозримым – таким образом, для каких-то задач мечта о полноте языкового материала становится реальностью. Эти задачи, на наш взгляд, связаны с описанием лексики и созданием автоматических словарей.

Обратим внимание, что здесь тоже меняется **подход** к проблеме: если традиционная лексикография ориентировалась на описание отдельного слова как самостоятельного объекта исследования, то теперь на первый план выходит исследование классов слов, объединенных по одному или нескольким признакам. Задача описания слова (или группы семантически близких слов) уже осознана как неразрешимая в силу своей неограниченности; однако с помощью базы данных можно формулировать ее, так сказать, частные "подзадачи", причем многие из них имеют исчерпывающие решения. Простым примером такого рода может служить задача сопоставления значений некоторых двух признаков, например: верно ли, что все имена обозначающие части, могут выступать в генитивной конструкции (ср. *ножка стула*, *ручка двери*, но **нос лица*). База данных, содержащая некоторую семантико-синтаксическую информацию о каждой лексеме, позволяет сравнить класс лексем, обозначающих части, и класс лексем, допускающих в конструкции "часть-целое" родительный падеж. Заметим, что ни один из существующих словарей не даст возможности получить аналогичный результат: мы не найдем там ни последовательно представленной информации о частях, ни тем более помет о возможности генитива, т.к. в свете глобального лексикографического описания такого рода сведения окажутся слишком частными.

Как кажется, уже нет необходимости защищать лингвистическую информацию, подобную этой, доказывать ее право на существование и подчеркивать степень ее важности – проблема ясна; обратим внимание лишь на следующие два ее аспекта.

Первое. Автоматическая база данных, в отличие от словаря, действительно воплощает связь между словарем и грамматикой, провозглашенную одним из главных принципов современного лексикографического описания (ср., например, идею Интегрального словаря и под.). В любом словаре, даже и в словаре типа Толково-комбинаторного, грамматика пополняется в результате описания отдельных лексических единиц, то есть по следам их описания. Между тем, лексическая база данных есть как бы одновременно и словарь, и грамматика: она сообщает нечто о частных лексемах, но при этом характеризует и целые их классы.

Второе. Мы уже говорили о том, что объемы лингвистического материала, которыми оперирует пользователь лексической базы данных, очень велики, так как база данных – это инструмент для отражения *полной* картины того или иного языкового явления: правила и исключения для нее равны. Между тем, практика показывает, что исключения есть во всяком, даже самом регулярном правиле. Примером может служить, в частности, стандартный способ выражения смысла 'во время' для имен ситуаций: 'во время S_0 ' → *при S_0* , ср.: *при болезни, при броске, при сопротивлении, при даче [свидетельских] показаний*, и т.п. Ср., однако, **при битве*, **при влиянии*, **при лестии* и др. запреты. Сплошная же классификация группы лексем по нескольким десяткам синтаксических признаков одновременно даст такую разнородную картину, что если бы мы решили разбить эти лексем на классы так, чтобы в один класс попали слова с одним и тем же набором значений признаков, классов получилось бы примерно столько же, сколько лексем. Таким образом, степень сложности поведения лексем, будучи зафиксирована машинной, становится до наглядности очевидной; по видимому, это одно из самых убедительных свидетельств того, что поверхностные (в частности, синтаксические) свойства лексем не случайны: носитель языка не в состоянии запомнить все это множество фактов, и разнообразие возникает под действием общих семантических (в широком понимании этого термина) правил.

Замечательно, что лексические базы данных не только обосновывают необходимость поиска таких правил, но и создают исследователю фундамент для этого поиска – в некотором смысле, точку опоры, чтобы перевернуть существующую лексикографическую традицию.

Простейшая база данных представляет собой таблицу с двумя входами – предназначенная для описания лексики, в качестве входов она будет иметь, с одной стороны, лексем, с другой – параметры некоторой классификации, т.е. признаки, значения которых приписываются лексемам в клетках таблицы. Минимальный набор значений признаков – два: "да" и "нет", или "+" и "-". Подобный формат оптимален для автоматических словарей сочетаемости: в таком словаре про каждую лексему из списка можно узнать, входит ли она в ту или иную конструкцию. Строго говоря, условия сочетаемости должны указываться максимально жестко, с точностью до конкретных форм лексем. Это значит, что такой словарь может отвечать на вопросы типа: сочетается ли данное существительное в винительном падеже с предлогом *в* (= входит ли в конструкцию *в + Acc*) или с прилагательным *красный*, но, вообще говоря, при совершенно аккуратной постановке задачи, не может давать ответ на вопрос, сочетается ли имя с прилагательными цвета, по той же причине, по которой некорректным окажется вопрос о сочетаемости с предлогами вообще. Другими словами, признак, по которому классифицируются лексем в такой базе, не должен был бы содержать никаких переменных. Последовательное выполнение этого требования, однако, лишило бы лексические базы данных всякой привлекательности и в большой степени научной ценности. Поэтому мы вынуждены принимать априорные решения о гомогенности некоторых лексических классов, т.е. об общности их (в данном случае, сочетаемостного) поведения – например в какой-то момент признать гомогенным класс

прилагательных цвета – считая, что все прилагательные цвета одинаковым образом сочетаются с именами артефактов. Заметим, что при исследовании сочетаемости непредметной лексики такое решение вряд ли себя оправдывает: мы "потеряем" такие сочетания, как *черная меланхолия, зеленая тоска, голубая мечта* и под.

Тем самым, уже приступая к созданию лексической базы данных, мы должны отдавать себе отчет в том, что никакие машинные возможности (объемы памяти, быстрейшее действие, единый формат представления данных и под.) не избавят нас ни от проблемы языковой правильности, ни от многозначности, ни от сочетаемостной избирательности и что всякая реальная лингвистическая задача, в особенности если она касается огромных лексических массивов и семантико-синтаксических классификаций, практически никогда не может быть сформулирована абсолютно строго – в противном случае она теряет лингвистический интерес (впрочем, совершенно так же, как и не строго сформулированная задача – маневр между Сциллой и Харибдой здесь, как и всегда, не может быть предсказан заранее).

И все-таки, какого рода задачи могут решать подобные машинные классификации? В дополнение к традиционным задачам лексикографии, которые все доступны базам данных, отметим здесь исследование фактативной сочетаемости, которая обычно не учитывается словарями. Ср., например, задачу полного описания сирконстантов глагола, которая решается в глагольной базе системы "Лексикограф-эксперт" (руководитель работ Е.В. Падучева) [1]. Сочетаемостные типы отглагольных имен в большинстве своем довольно давно и хорошо известны, т.к. попадают под классификацию так называемых лексических функций, разработанных для русского языка в рамках модели "Смысл ↔ Текст". Тем не менее, "сплошное" исследование поведения отглагольных имен в контексте лексических функций – база данных "Перифраз" (разработчики В.А. Плунгян, Е.В. Рахилина, М.В. Филипенко) – начато только в этом году. Результатами такого рода исследований, с нашей точки зрения, должны быть общие правила, объясняющие значения признаков в клетках таблицы – именно мотивированные правила контролируют лингвистическую интуицию исследователей, страхуют от случайных ошибок.

Другая задача – таблица сочетаемости конкретной лексемы – например, прилагательного *прочный* – дает материал для словарного толкования этого слова. Вообще говоря, исследование сочетаемостных особенностей всегда должно было бы быть предварительным этапом лексикографического описания: толкование должно объяснять, почему возможно *прочные замки*, но не **прочные ключи*. Тем самым, базой и одновременно стимулом для семантических обобщений по поводу слова *прочный* оказывается полный список имен с пометами о допустимости-недопустимости при них данного прилагательного.

До сих пор мы говорили о задачах, для решения которых требовалась простая классификация, т.е. классификация по одному или нескольким однородным признакам. Между тем, современные компьютеры позволяют строить и гораздо более сложные конструкции. Используются наборы разных признаков, пересечения и объединения их значений. Благодаря такому программному обеспечению можно усложнять лингвистические эксперименты, в частности, одновременно задавать несколько совершенно разных оснований классификации одной и той же группы лексем. Именно по этому принципу построена система "Лексикограф-эксперт. Предметные имена" (разработчики И.С. Красильщик, Е.В. Рахилина, при участии В.А. Плунгыана). В этой системе параллельно используется несколько стратегий классификации предметной лексики: мереологическая классификация (целое-множество; часть-элемент), таксономическая, устанавливающая иерархию родо-видовых отношений (типа *дом* → *постройка* → *сооружение* и под.). Кроме того, учитываются предикатно-аргументные отношения, "встроенные" в структуру имен: *ложка* ≈ 'приспособление (чтобы) ЕСТЬ'; *кислород* ≈ 'газ (чтобы) ДЫШАТЬ', и т.п. Пользуясь этой системой, мы можем получать самые разные "тематические" группы лексем: все имена, обозначающие части или эле-

жества, все предметы мебели, постройки, приспособления, все места обитания (связанные с предикатом ЖИТЬ), и т.п. (подробнее см. [2]).

В соответствии с общей идеей системы "Лексикограф", мы исходим из гипотезы, что все сочетаемое поведение лексемы выводится из каких-то ее глубинных свойств. В этом случае базу данных можно рассматривать как один из вариантов моделирования этих свойств. Дальнейшая работа сводится к исследованию степени адекватности предложенной модели – т.е. принятого нами формата описания предметной лексики, и проверке самой гипотезы: насколько вообще верна идея выводимости поверхностных характеристик лексемы. Тем самым, база данных служит здесь не только практической лексикографии, но и теоретической лексикологии.

Рассмотрим теперь некоторые конкретные задачи описания предметной лексики русского языка, решаемые с помощью системы "Лексикограф"¹.

Основная трудность описания такого рода лексем – в том, что в этой области пока нет никаких канонов, так сказать, лексикографических норм – в отличие, например, от предикатной лексики. В целом система ориентирована на то, чтобы отражать и предсказывать поведение слова в тексте (отсюда и ее статус *экспертной* системы по русской лексике), поэтому основным рабочим критерием необходимости для Базы данных той или иной лексикографической информации оказывается то, определяет ли эта глубинная информация какие-то поверхностные свойства лексемы, и прежде всего, ее сочетаемость.

Формат Базы данных "Предметные имена" представляет собой набор признаков, значения которых приписываются лексемам; ниже, после его краткого описания, мы приведем примеры лингвистических задач, которые могут решаться с помощью этой системы.

Формат Базы содержит 10 полей. Первые три охватывают грамматическую информацию о лексеме и отражают те сведения, которые содержатся в Грамматическом словаре А.А. Зализняка – род, одушевленность, тип склонения, акцентная парадигма. Затем следуют поля, отражающие семантико-синтаксическую информацию: мерологический статус и коррелят, класс, предикат и три его ограничителя. Мерологические поля определяют лексему как часть/множество/совокупность/элемент (значение признака "мерологический статус", или М-статус) какого-то другого объекта (имя этого объекта задает значение признака "мерологический коррелят", или М-коррелят): *ветка* – М-статус 'часть', М-коррелят 'дерево', *кузов* – М-статус 'часть', М-коррелят 'автомобиль', *оркестр* – М-статус 'множество', М-коррелят 'музыкант', *инвентарь* – М-статус 'совокупность', М-коррелят 'инструмент', и т.п.

В поле "класс" лексеме приписывает семантический класс из некоторой разрабатываемой в системе таксономической классификации; это может быть 'вещество', 'инструмент', 'емкость', 'отверстие', 'постройка', 'учреждение', 'зверь' и мн. др. Значения этого признака устанавливают родо-видовые отношения между лексемами.

Поле "предикат" соотносит предметное имя с предикатом, семантически связанным с данной лексемой – обычно это предикат функционирования описываемого объекта; значения основных трех актантов этого предиката (поля "ограничитель 1", "ограничитель 2", "ограничитель 3") описывают другие, более сложные, чем родо-видовые отношения на множестве предметных имен. Ср.: *дом* – класс 'постройка', предикат 'жить', ограничитель 1 'человек', ограничитель 3 '=', (условный знак кореферентности значения поля 'класс'); *виноградник* – класс 'участок земли', предикат 'выращивать', ограничитель 1 'человек', ограничитель 2 'виноград 2', ограничитель 3 '='; *градусник* – класс 'прибор', предикат 'измерять', ограничитель 1 'человек', ограничитель 2 'температура', ограничитель 3 '=', *клетка* – класс 'сооружение',

¹ В настоящее время в Базе данных системы "Лексикограф" представлена предметная лексика стандартного учебного словаря объемом 10–13 тыс. слов – это составляет почти 4 тыс. входов, к концу года список должен быть расширен до объема, соответствующего одному толковому словарю в 60 тыс. слов и в конечном счете приблизится к 25 тыс. входов.

предикат 'содержать', ограничитель 1 'человек', ограничитель 2 'животное', ограничитель 3 '= ', и под. Легко видеть, что в совокупности значения признаков по достаточно простым правилам "складываются" в некоторую общую схему толкования: *дом* – 'постройка, в которой живет человек'; *виноградник* – 'участок земли, на котором человек выращивает виноград 2', и т.д. Эта схема очень огрубленно представляет значение каждой конкретной лексемы – огрубленно прежде всего потому, что в любой базе данных изначально заключена идея неиндивидуализированного описания объектов за счет пренебрежения частными, несущественными для заложенной в базу системы признаков различиями. Этот "недостаток" окупается возможностью автоматического поиска списков лексем по достаточно сложным комбинациям семантико-синтаксических признаков: пользуясь Базой мы можем не только получать списки всех имен, обозначающих части, множества, имена родства, названия предметов мебели и под., но и, например, имена множеств и совокупностей, имеющиеся при этом дефектную парадигму числа (*Pluralia tantum* или *Singularia tantum*), имена частей, не являющихся при этом частями тела или, например, частями механизмов и проч. Такого рода списки необходимы для решения лингвистических задач, опирающихся на "сплошное" исследование лексического материала. Интересно, что множество задач такого рода как бы навязывается самой признаковой базой системы. В качестве иллюстрации укажем некоторые из них.

Начнем с *мереологических* признаков. В русском языке есть лексемы, которые претендуют на статус частей и есть грамматические конструкции, которые выражают отношение часть-целое, например, генитивная конструкция (ср. *ножка стула*), притяжательная конструкция (ср. *кошкин хвост*), конструкция с относительным прилагательным (ср. *березовые ветки*), предложные конструкции (ср. *пальцы на руке, платье с рукавами, туфли на каблуках* и др.). Как связаны между собой эти два обстоятельства – почему в одних случаях конструкция запрещена, а в других нет? Какими другими семантико-синтаксическими свойствами² должны обладать лексемы, входящие в отношение часть-целое, чтобы это отношение можно было выразить данной конструкцией? И – в другую сторону: каковы должны быть те "языковые свидетельства" (по терминологии А. Вежбицкой, см. [5]), которые дают право лексикографу написать в словаре, что данные две лексемы действительно входят в отношение часть-целое или элемент-множество и т.д.? Рассматривая проблему под этим углом зрения приходится усомниться, например, в такой простой истине, что остров является частью суши – не с точки зрения внеязыковой, энциклопедической реальности, а с точки зрения языкового членения мира, которое, впрочем, и призвано быть отраженным в словаре.

Таксономическая классификация нужна для решения любой сочетаемостной задачи – описание ограничений на употребление имен естественно вести именно в терминах семантических классов. Между тем есть лингвистическая задача, которая сама "контролирует" процесс создания семантической классификации предметных имен – это задача описания сочетаемости предметной лексики с отыменными прилагательными. По одной из известных теоретических гипотез, выдвинутой в свое время Е.А. Земской (последнюю по времени версию см. в [6]) и некоторыми другими лингвистами (ср., например, [7]) и с известными оговорками поддержанной Ю.Д. Апресяном (см. [8, с. 211–212]), сочетаемость имен с такими прилагательными определяется семантическими отношениями в паре из двух имен, одно из которых затем выступает в этой конструкции в виде отыменного прилагательного. Так, легко образуют эту конструкцию пары 'вещество', – 'емкость' (ср. *коньячная бутылка, молочный бидон*), 'материал изготовления', – 'артефакт' (ср. *хлорвиниловая пленка, железная труба*) и мн. др. Система "Лексикограф" позволит проверить эту гипотезу, выявить все продуктивные для данной конструкции семантические типы пар, определять словарные исключения из общих семантических

²Некоторые аспекты этой проблемы обсуждаются в [3 и 4].

правил. С другой стороны, само по себе решение этой задачи дает ключ к построению и коррекции семантической классификации предметных имен и тезаурусных отношений между ними: значение семантического класса 'емкость' или 'пространство' мы должны приписывать лексеме *море* в зависимости от того, ведет ли себя эта лексема в конструкции с отыменным прилагательным по модели емкостей или по модели пространств.

О роли, которую играет встроенный в семантику предметной лексемы ф у н к ц и о н а л ь н ы й п р е д и к а т хорошо известно, в частности, из статьи З. Вендлера [9] о прилагательном *good*: *хороший/удобный дом* – это дом, в котором удобно жить, а *хорошая балерина* – это балерина, которая хорошо танцует, – именно этот предикат попадает в сферу действия оценочных и других качественных прилагательных (подробнее см. [10]) и определяет в этих случаях сочетаемостные возможности имени и способ интерпретации полученной конструкции.

Однако если постулировать предикат в семантической структуре существительного, интересно проследить зависимость между его свойствами и свойствами существительного. Рассмотрим, например, видо-временные характеристики такого предиката. В работах Е.Я. Шмелевой (см. прежде всего [11]) показано, что при описании русских имен деятеля необходимо различать имена деятеля, выступающие в актуальном значении (*предъявитель* – 'тот, кто в данный момент предъявляет что-л.'), *перфектином* (*освободитель* – 'тот, кто освободил кого-л. '), и в значении свойства или функции (*выдумщик* – 'тот, кому свойственно выдумывать', *точильщик* 'тот, в чьи функции входит точить что-л.' – другими словами, 'тот, кто обычно делает нечто'). Понятно, что такое различие отражается на поверхностных свойствах имен деятеля, между тем лексикографически оно закрепляется именно в различных видо-временных формах встроенного в имя глагола. Система "Лексикограф" заставляет определять видо-временные характеристики встроенного предиката для каждой лексемы, в толковании которой такой предикат имеется. Таким образом выделяется, например, группа имен-результатов, таких как *зола, копоть, пар, ожог, дыра, след* и под. (диагностическим контекстом для них является конструкция "X от Y-а" в значении 'X – результат функционирования Y', ср. *зола от костра, пар от чайника, след от гвоздя*) – имена-результаты продолжают список "перфектных" имен русского языка.

Проблематика лингвистических задач, связанных с а к т а н т н о й с т р у к т у р о й встроенного предиката наследуется из опыта описания предикатной лексики, накопленного в рамках Московской семантической школы. Если рассуждать в русле принятой и глубоко разработанной в этой школе концепции семантических и синтаксических валентностей лексемы (см. ставшие классическими работы [8, 12, 13]), то слово *строитель* потому управляет лексемой *город* в словосочетании *строитель города*, что оно как бы наследует семантическую структуру исходного предиката *строить*. При этом сам предикат *строить* двухвалентен: 'кто строит что', так что если его первую валентность ('кто строит') заполняет само имя *строитель*, то вторая – объектная – остается свободной и "обеспечивает" правильность словосочетания *строитель города*. Таким образом, с одной стороны, синтаксическая модель управления имени оказывается мотивирована его семантикой и, прежде всего, семантикой "встроенного" предиката, а с другой стороны, если в семантике этого предиката оказывается свободная валентность, то естественно ожидать, что она будет как-то "реализована" в тексте. В переводе этого рассуждения на язык нашей Базы данных можно было бы сказать, что каждой заполненной строке признака "ограничитель предиката" должен быть приписан способ или способы поверхностного выражения этого семантического признака – точно так же, как это делается для модели управления глагола. Это практическая лексикографическая задача, решение которой можно в какой-то степени контролировать с помощью нашей системы. Например, можно отслеживать количество актантов встроенного предиката или устанавливать стандартные модели соответствий между типом семантической валентности и

способом ее выражения (имеются в виду такие достаточно распространенные модели, как выражение первой валентности посессивным отношением, ср. *мой дом* – 'в котором я живу') и отклонения от этих стандартных моделей.

Между тем последовательное выполнение этой программы ставит перед исследователем и неожиданные теоретические проблемы. Оказывается, что в очень большом числе случаев свободная семантическая валентность не подкрепляется никакой поверхностной реализацией. Так, и *пассажир*, и *всадник* обозначают людей, использующих некоторое средство передвижения – транспорт или лошадь. Но в первом случае мы можем выразить это средство – *пассажир трамвая*, а во втором – нет: **всадник лошади*, ср. также **всадник вороного*, **всадник крапчатой* и т.п. Точно так же *гражданин* – этот человек, который живет в некотором государстве, отсюда *гражданин Дании, Норвегии, России*. Но *иностранец* тоже называет гражданина некоторого государства (только отличного от государства, в котором живет говорящий), – тем не менее, сочетания типа **иностранец Дании, Норвегии* оказываются запрещены. *Инспектор* – это должностное лицо, посланное некоторым учреждением или организацией инспектировать какой-либо объект/учреждение и т.п. Однако при этой лексеме, как правило, поверхностно выражается только то учреждение, которое организует инспекцию, ср. *инспектор горного/главного санитарного управления*, но не **инспектор шахты/фабрики*. Воин воюет с кем-то, но мы не можем непосредственно при этой лексеме выразить, с кем: грабли сгребают что-то, но это "что-то" не входит в синтаксическую модель управления лексемы *грабли*. Семантические правила, регулирующие и запреты такого рода, могут быть сформулированы только после того, как будут выявлены все лексемы, имеющие подобные особенности. Система "Лексикограф" располагает возможностями их системного поиска и исследования.

В кратком обзоре мы рассмотрели далеко не все направления, по которым развивается и будет развиваться в ближайше время система "Лексикограф". Нам кажется, что осуществление этого проекта может дать интересные результаты и в практике, и в теории современной русистики.

СПИСОК ЛИТЕРАТУРЫ

1. Падучева Е.В., Розина Р.И. Семантический класс глаголов полного охвата: толкование и лексико-синтаксические свойства // ВЯ. 1993. № 6.
2. Красилицык И.С., Рахилина Е.В. Предметные имена в системе "Лексикограф" // НТИ. Сер. 2. 1992. № 9.
3. Рахилина Е.В. К основанию лингвистической мереологии // Семиотика и информатика. Вып. 30. 1989.
4. Воронцова М.И., Рахилина Е.В. Предметные имена и предложные конструкции // Знак (сборник статей по лингвистике и семиотике памяти А.Н. Журицкого). М., 1994.
5. Weitzbicka A. Lexicography and conceptual analysis. Ann. Arbor, 1985.
6. Земская Е.А. Относительные прилагательные как конструктивный элемент номинативной системы современного языка // Грамматические исследования. Функционально-стилистический аспект. М., 1991.
7. Павлов В.М. Понятие лексемы и проблема отношений синтаксиса и словообразования. Л., 1985.
8. Апресян Ю.Д. Лексическая семантика. М., 1974.
9. Вендлер З. О слове *good* // Новое в зарубежной лингвистике. Вып. X, М., 1981.
10. Рахилина Е.В. Прилагательные сквозь призму существительных и vice versa // НТИ. Сер. 2. 1991. № 9.
11. Шмелева Е.Я. Названия производителя действия в современном русском языке (словообразовательно-семантический анализ). Дисс. ... канд. филол. наук. М., 1984.
12. Мельчук И.А. Опыт теории лингвистических моделей "Смысл ↔ Текст". М., 1974.
13. Богуславский И.М. Исследования по синтаксической семантике. М., 1985.