



УДК 811.161.1

**КОНСТРУИРОВАНИЕ ГИПЕРТЕКСТОВОГО ТЕЗАУРУСА МЕТАЯЗЫКА НАУКИ****С. В. Лесников***Сыктывкарский  
государственный  
университет**e-mail:  
serg@lsw.ru*

В настоящей статье даётся описание конструирования гипертекстового тезауруса метаязыка науки при финансовой поддержке Российского фонда фундаментальных исследований по инициативному, исследовательскому проекту (грант) № 11-07-00733 (2011–2013) «Гипертекстовый информационно-поисковый тезаурус “Метаязык науки” (структура; математическое, лингвистическое и программное обеспечения; разделы: лингвистика, математика, экономика)» (научный руководитель – С.В. Лесников).

Ключевые слова: алгоритм, гипертекст, гизаурус, глоссарий, глотология, Интернет, компьютер, лингвистика, метаязык, лексикография, поиск, программа, русский язык, словарь, справочник, тезаурус, энциклопедия, языковедение, языкознание.

При конструировании гипертекстового тезауруса метаязыка науки принимается во внимание опыт ведущих языковедов и, главным образом, уже изданные тезаурусы на русском языке: информационно-поисковый, по библиотечному делу и научной информации, по информатике, в автоматизированных системах управления и обработки информации, по геологии; по литературе, литературоведению и фольклористике; органических реакций, по философии, по автомобильному транспорту, по атомной науке и технике, лингвистической терминологии, по языкознанию, по чёрной металлургии, по теоретической и прикладной лингвистике, семантических множителей русского языка, семантический, средств каталогизации, по образованию, по порошковой металлургии, органических соединений, менеджера, ассоциативный, физики, финансовой англomorphicологии, экономики и демографии, русской версии тезауруса EUROVOC, правоведения, по социологии, общеобразовательный, по литературоведению, синонимов, балто-славянский, гендерных исследований, идиоматики, пословиц, по политологии, по религиоведению и др.

Учитывались также принципы лингвистического конструирования и приёмы разработки тезауруса на лингвистическом материале, рассмотренных в работах следующих учёных: А. С. Герд, Ю. Н. Караулов, Р. Г. Пиотровский, А. Я. Шайкевич, С. Д. Шелов, А. И. Михайлов, А. И. Черный, Р. С. Гиляревский, А. Ф. Паркер-Роудс и С. Уордли, Д. Варга, А. А. Витухновской, А. В. Соколов, С. Е. Никитина, В. Д. Сидорченко, М. П. Данилов; Э. Я. Лесохина, В. Б. Смиренский; Е. В. Шингарева; В. А. Чижаковский, Л. Н. Беляева; О. А. Лавренова; И. В. Азарова, О. А. Митрофанова, А. А. Синопальникова, А. А. Ушакова, М. В. Яворская; С. В. Жмайло и др.

В гипертекстовом информационно-поисковом тезаурусе метаязыка лингвистики (в формате компьютерного, гипертекстового тезауруса /ГИЗАУРУСА/ - базе данных с перекрёстными гипертекстовыми ссылками) даётся определение (дефиниция, иллюстрация, интерпретация, объяснение, понятие, пояснение, разъяснение, толкование, трактовка, формулировка, цитата, экскурс и/или эксцерпция) ключевых понятий-терминов, относящихся к сфере определённой науки и ряду смежных дисциплин, которые в той или иной мере связаны с проблемами данной науки. Конструирование тезауруса обусловлено необходимостью единства в терминологии - нередко учёные одно и то же явление называют по-разному, с другой стороны, один и тот же термин бывает полисемичен даже в рамках одной научной области. Кроме этого, при анализе уже опубликованных терминологических словарей определённой науки бросается в глаза лакуарность состава словников и некая тенденциозность подачи научных материалов авторами-составителями. Фактически речь идёт о составлении метаязыка науки – осо-





бого языка, объектом которого является содержание и выражение другого языка, т.е. в нашем случае метаязык – это язык второго порядка как специальная семиологическая система, употребляемая тогда, когда надо говорить о естественном (или искусственном) языке же, выступающем в качестве «языка-объекта». При этом можно отдельно выделить в метаязыке следующее: 1) собственно термины (слова, которые или совсем не употребляются в языке-объекте, либо приобретают, после заимствования из языка-объекта, особое значение), 2) сочетания слов, которые характеризуют собственно метаязык определённой науки в его полном виде (напр., составные термины), 3) определённый социально-предметный аспект (напр., когда метаречь отражает специфику эпохи, того или иного научного направления, школы).

Тезаурус метаязыка науки конструируется в виде следующего гипертекстового фрейма и пяти базовых слотов:

	верхний	
левый	центральный	правый
	нижний	

Верхний – название тезауруса и базовые кнопки дружественного интерфейса для быстрого, оперативного переключения слотов (в свою очередь др. фреймов), напр., переход на стартовую, главную страницу, вызов помощи, возврат назад и переход по некоторым базовым конструкциям (ключевым концептам: знак, слово, текст, язык, метаязык, приложения). Ссылки из верхнего слота активируют левый слот, т. е. использование гипертекстовых ссылок в верхнем слоте приводит к изменению содержимого левого слота посредством загрузки в этом левый слот соответствующего содержания ассоциированного фрейма.

Левый слот играет роль глобального или частного оглавления (содержания). Активирование ссылки в левом слоте приводит к изменению центрального слота.

В правом слоте помещён латинский (английский) алфавит, для вызова соответствующих индексов словарных статей тезауруса побуквенно. Нижний – соответственно русский алфавит. Из правого и левого слотов активизируется центральный слот.

В центральном слоте в основном и происходит выдача контента.

Разумеется, имеется возможность открыть по любой ссылке новый фрейм (окно) посредством активизации гиперссылки правой клавишей манипулятора «мышь».

При этом необходимо подчеркнуть, что лишь эти пять слотов размечаются вручную. Все словарные статьи тезауруса хранятся пофайлово, а как алфавитные (и русский и английский) пофайловые, так и внутренние перекрёстные (по содержимому собственно словарных статей) гиперссылки генерируются при помощи программного комплекса на основе предварительно препарированных, размеченных исходных словарных материалов.

Процесс использования гипертекстового тезауруса метаязыка науки в целом напоминает обычную работу в системе Интернет, но в локальном варианте на персональном компьютере. В перспективе предполагается, что после необходимой апробации в локальном варианте в интерактивном варианте сконструированный тезаурус будет доступен и в Интернете, напр., по адресам <http://lsw.ru>, <http://ЛСВ.РФ>, <http://МФРЯ.РФ>.

Немного предыстории о гипертексте и некоторые определения:

Первая попытка сконструировать, точнее, описать "гипертекстовую машину" была предпринята в 1945 г. [1]. Рассматривая историю развития счётных машин, начиная с Лейбница и Бэббиджа, В. Буш предложил проект электронно-механического устройства, названного "memex" (от английского MEMory EXtension), которое должно дать человеку инструмент для усиления мыслительных способностей. Memex позволяет человеку определённым способом организовать свои записи, статьи, книги. С помощью механических средств (рычагов, кнопок и т. п.) тексты, записанные на микрофильмы, могут быть связаны между собой так, что переход от одного к другому осуществляется непосредственно, даже если физически они расположены далеко друг от





друга. При чтении таким способом возможны пропуски страниц, возврат назад или просмотр каждых десяти страниц. Одновременно на экран могут вызываться несколько микрофильмов. Дополнительно к существующим материалам могут записываться новые – статьи, заметки, а также комментарий, ссылки и т. п.

Следующее слово в истории гипертекста сказал Теодор Нельсон – программист, придумавший термин «гипертекст», дав следующее определение гипертекста: текст, не имеющий линейной последовательности («нелинейное письмо»). Нельсон стремился реализовать идеи удачного, по его мнению, использования компьютера в виде всеобщей системы электронных публикаций и архива. А в основе подобной системы должно лежать понятие гипертекста. По мнению Т. Нельсона, компьютер призван сделать человека свободным, а поэтому нужно разработать систему, максимально облегчающую работу с информацией [2]. Литература (понимаемая в широком смысле как информация, предназначенная для сохранения) представляет собой непрерывную систему взаимосвязанных произведений, и это надо рассматривать не как определение, а как обнаруженную данность, т. е. «просто не-последовательное письмо (writing), текст, который разветвляется» и предоставляет читателю выбор [3].

В 1980-е гг. создаётся множество коммерческих гипертекстовых систем. Согласно Conklin J., гипертекстовые системы классифицируются так: 1) библиотечные макросистемы; 2) средства исследования проблем («инструментальные средства для поддержки неструктурированного мышления на ранней стадии исследования проблем, когда человек имеет дело с многочисленными, плохо связанными одна с другой идеями»); 3) системы для просмотра («browsing») базы данных (то есть системы, в которых добавление данных пользователями не предусматривается); 4) системы широкого назначения (для экспериментирования в различных областях применения гипертекста). Механизм, заключающийся в возможности связать отрывки текста, переходить от одного к другому, называется гипертекстом, или нелинейным текстом: а) «надтекст» – некая единица информации, частями которой являются тексты, и/или б) текст, части которого имеют «сверхсвязи», то есть соединены друг с другом не линейным отношением в одномерном пространстве (отношением следования, как в обычном тексте естественного языка), а множеством различных отношений, представляемых в многомерном пространстве [4].

На русском языке о гипертексте одним из первых заявил публично М. М. Субботин: гипертекст – это соединение смысловой структуры, структуры внутренних связей некоего содержания, и технической среды, технических средств, дающих возможность человеку осваивать структуру смысловых связей, осуществлять переходы между взаимосвязанными элементами [5]. Под руководством М. М. Субботина группа исследователей, занимаясь вопросами организации электронных текстов, предложила метод логико-смыслового моделирования является как бы особой версией гипертекстовой технологии. Авторы этого метода утверждают, что помимо чтения гипертекста, существует ещё одна качественно новая форма интеллектуальной деятельности – «создание линейного текста путём линеаризации нелинейного» [6].

Британская энциклопедия упоминает о гипертексте в разделе «информатика» (information science), при обсуждении методов семантического представления информации. Гипертекст определяется как метод, при котором документы объединяются посредством «связей» в сеть. Условием объединения может служить сходство идей, выводов, одинаковое расположение частей и так далее [7].

В энциклопедии языкознания гипертексту посвящена целая статья, смысл которой приблизительно соответствует смыслу статьи в британской энциклопедии с тем добавлением, что гипертекст – это структурированный текст [8].

В толковом словаре по информатике гипертекст трактуется как информационный массив, на котором заданы и автоматически поддерживаются ассоциативные и смысловые связи между выделенными элементами, понятиями, терминами или разделами; наряду с этим приводится дополнительное пояснение, что гипертекст – это



текст, в котором отдельные термины на экране дисплея выделены подсветкой и их определение сразу может быть выдано на экран [9].

Л. В. Шуткин структуры гипертекстов представляет моделью в виде регулярной конфигурации теории паттернов, при этом структуры гипертекстовых тем рассматриваются как образующие темы [10].

Безусловно, суть гипертекста можно уяснить лишь непосредственно в компьютерном варианте, однако для примера можно привести опубликованную книжную версию словаря-гипертекста, где авторы определяют: «Гипертекст – это метод построения и организации электронной компьютерной документации, позволяющей отразить ассоциативные связи между её отдельными разделами, понятиями и терминами. Читая такую документацию, как электронную книгу на экране ЭВМ, и встретив новое, неясное или забытое понятие, можно немедленно заставить машину провести Вас по всем ассоциативным связям этого понятия и показать любые участки текста, поясняющие его смысл или раскрывающие взаимодействие с др. объектами, категориями, терминами» [11, с. 3]. Элементы гипертекста и использование соответствующих гипертекстовых технологий в лингвистике можно наблюдать также в виде паралингвистических реалий и средств, которые могут быть представлены в виде поля паралингвистических средств, определяющих внешнюю организацию письменного текста и создающих его (текста) «оптический образ». Собственно паралингвистическое поле может детерминироваться как графическим представлением текста (напр., сегментацией и расположением), так и использованием разнообразных типографских знаков, шрифтов, средств иконоческого и пиктографического языка, схем, рисунков, таблиц, фотографий и др. [12].

Для полноты представления понятия «гипертекст» приведём цитату из «Словаря культуры XX в.: ключевые понятия и тексты» В. П. Руднева: «ГИПЕРТЕКСТ – текст, устроенный таким образом, что он превращается в систему, иерархию текстов, одновременно составляя единство и множество текстов. Простейший пример Г. – это любой словарь или энциклопедия, где каждая статья имеет отсылки к другим статьям этого же словаря. В результате читать такой текст можно по-разному: от одной статьи к другой, по мере надобности, игнорируя гипертекстовые отсылки; читать статьи подряд, справляясь с отсылками; наконец, пуститься в гипертекстовое плавание, то есть от одной отсылки переходить к другой. Настоящий словарь является подчёркнуто гипертекстовым, то есть построен так, чтобы постоянно провоцировать именно третий путь чтения» [13, с. 69 – 72]. Словарь культуры XX в. Руднева представляет собой уникальный словарь-гипертекст и содержит 140 словарных статей, посвящённых наиболее актуальным понятиям и текстам культуры XX века. Для нас актуальны словарные статьи: *автокоммуникация, виртуальные реальности, генеративная лингвистика, гипотеза лингвистической относительности, знак, метаязык, многозначные логики, принцип дополнительности, структурная лингвистика, текст в тексте, формальная школа* и др. Текст словаря Руднева в полном объёме доступен в Интернете, напр., по адресу <http://rudnevslovar.narod.ru/>.

"Гипертекст обеспечивает нелинейную организацию содержания со стороны продуциента и возможность нелинейного восприятия со стороны реципиента. Формально гипертекст – это вся совокупность текстов (...), связанных с воспринимаемым текстом посредством ссылочного аппарата. Гипертексты, «технически» соединённые друг с другом ссылочным аппаратом, следует отличать от тематически и организационно (но не «технически») объединённой последовательности текстов, напр. совокупности реплик всех участников какой-нибудь дискуссионной группы. В гипертексте на место двумерного сочетания горизонтального и вертикального контекстов приходит контекстуальная трёх- и даже  $n$ -мерность ( $n > 3$ ). Каждая ссылка представляет собой вектор-указатель на другой текст, который, в свою очередь, либо задаёт  $(n-1)$ -мерность, имея выход на другие тексты (т. е. соединяется с  $n$  текстов минус один первичный текст), либо является двумерным, обладая только горизонтальным





и вертикальным контекстами (имеет вход, но не имеет выхода на другие тексты). В связи с этим культурно-речевой аспект гипертекстуальности может включать среди прочего исследование способов оптимизации ссылок и установление определённых ссылочных норм. Идея формального и семантического взаимоотношения между различными текстами, идея интер- и гипертекстуальности не нова. В применении к компьютерным документам её возводят к 1965 г. Типичным же гипертекстом досетевой эпохи является совокупность текстов Священного Писания. Их можно читать не только линейно, но и нелинейно, следуя эксплицитно указанному во многих изданиях книг Ветхого и Нового Завета порядку ассоциативных ссылок и параллельных мест (о Священном Писании как гипертексте ср.: [14]. «Гипертекст позволяет связывать текст, аудио, фотографии, чертежи, карты, движущиеся картинки и другие формы информации в осмысленное целое, к которому может осуществляться доступ при помощи системы индексации, ориентированной на конкретные идеи, а не на конкретные слова в тексте» [15].

Таким образом, под гипертекстом может пониматься как:

- метод объединения документов;
- текст, организованный по-особому, при этом, важную роль в этой организации играют «связи» между текстами, устанавливающиеся произвольно;
- механизм, позволяющий эти тексты связывать гиперссылками, однако связанные тексты здесь рассматриваются как единое целое, большее, чем текст, т. е. «гипер-текст»;
- форма организации материала (электронная, компьютерная);
- средство организации текстов;
- документация, которая ветвится. В разных контекстах трактовка гипертекста может приобретать различные оттенки и рассматриваться как метод, средство, форма, механизм актуализации информации и т. д.

Практическим итогом гипертекстовых разработок стала компьютерная сеть Интернет.

Для компьютерных технологий в Интернете принято следующее определение. Гипертекст – hypertext – метод подготовки и распространения текстов, который идеально подходит для компьютера, на котором читатель может выбирать свой путь прохождения по материалу. Для подготовки гипертекста вы сначала «нарезаете» информацию в виде небольших частей, которыми легко манипулировать, таких, как одна страница текста. Затем вы встраиваете в текст гиперсвязи. Когда читатель щелкает мышью на гиперсвязи, программное обеспечение поддержки гипертекста отображает информацию, предоставляемую другим узлом. Процесс управления прохождением среди узлов, связанных в этот путь, называется просмотром. Набор узлов, взаимосвязанных посредством гиперсвязи, называется паутиной WWW – World Wide Web и представляет собой гипертекстовую систему мирового масштаба.

Отметим ещё несколько базовых определений, используемых в информационных технологиях в связи с понятием «гипертекст»:

1) http – Hypertext transfer protocol – протокол передачи гипертекстов – способ передачи страниц WWW в Интернете;

2) HTML – Hypertext markup language – язык ссылок гипертекста – позволяет включать в текст коды (тэги, признаки, параметры, разметку) для определения шрифтов, стилей, графики, аудио и видео данных как непосредственно, так и посредством гипертекстовых ссылок. Процесс конструирования HTML-текстов называется авторизацией. Любая HTML страничка может быть создана и соответственно отредактирована практически в любом текстовом редакторе, который позволяет обрабатывать файлы в коде ASCII = American Standard Code for Information Interchange – американский стандартный код для обмена информацией (версия ISO).

Основные понятия из предметной области «гипертекст»:





- браузер (программа просмотра гипертекста);
- связь (отношение между фрагментами текста, выраженное некоторым образом в данных информационной системы);
- ссылка (указатель при ... на другой фрагмент);
- текстура (текст, оформленный некоторым образом посредством подбора шрифтов, цветов, размещения компонентов, фрагмент (фрейм – часть текста, имеющая смысловую законченность, ссылаемая как целое) – язык разметки гипертекста.

Далее, на этой основе и осуществляется конструирование гипертекстового тезауруса метаязыка науки.

Кратко изложим основные этапы конструирования тезауруса:

1) разработка структуры тезауруса. Определение тематического охвата тезауруса. Выбор источников и кластеризация предметной области (прежде всего, лингвистика, математика, экономика). При выборе источников учитывается:

- наиболее точное соответствие лексического материала предметной области;
- технологическую насыщенность и важность источников;

2) конструирование программного обеспечения; обеспечение интуитивно-понятного и дружественного интерфейса тезауруса, который является не только самостоятельным и самодостаточным компьютерным ресурсом, но и инструментом для поиска, классификации, систематизации и индексации ресурсов. Пользователь тезауруса имеет возможность осуществлять:

- просмотр тезауруса;
- поиск по ключевым словам и ассоциированным терминам (дескрипторам, классам, темам);
- навигацию по тезаурусу (поиск искомого понятия в тезаурусе с последующим запросом ресурсов, соответствующих этому понятию);

3) составление генерального словника (аскрипторы, дескрипторы, лексические единицы, одиночные слова (существительные, прилагательные, глаголы, наречия); именные словосочетания; лексически значимые компоненты сложных слов; сокращения слов и словосочетаний; реляторы) тезауруса. В генеральный словник включаются вокабулы (заголовки словарных статей) с указанием, в каких словарях можно найти информацию о том или ином слове. В генеральном словнике актуализируется информация не столько по электронным версиям (которые загружены в компьютерную словарную базу данных, т. к. такие вокабулы и так найдутся по запросу а) «целиком слово»), а, прежде всего, словники книжных (бумажных) словарей и словники с других словарных порталов. По числовому коду словаря выдаётся библиография к исходному словарю и гиперссылки соответствующих источников;

4) выявление семантических полей посредством определения кортежей условной эквивалентности /КУЭ/. В КУЭ объединяются лексические единицы, между объёмами понятий которых существуют следующие отношения:

- равнозначность – объёмы понятий совпадают полностью;
- пересечение – объёмы понятий совпадают частично;
- подчинённость – объём одного понятия строго составляет часть другого понятия, но не наоборот;
- внеположенность – объёмы понятий полностью исключают друг друга и при этом не исчерпывают области предметов;

5) установление парадигматических (способы выражения: аналитический, графический, лексикографический, табличный), иерархических (транзитивность и антисимметричность), ассоциативных, грамматических (сюда же можно отнести транслитерацию и варианты записи слов), межъязыковых (основа построения машинного переводчика), семантических (антонимы, гиперонимы, гипонимы, синонимы, гендерные синонимы; уменьшительно-ласкательные, усилительные-увеличительные), семантико-грамматических (связи действия, свойства, инструмента, местопребывания);





6) создание словарных статей и лексико-систематических (тематические, категориальные, смешанные), пермутационных, иерархических и др. указателей и списков специальных категорий лексических единиц (общие категории: названия дисциплин и отраслей деятельности; предметы, материалы; методы, процессы, операции, явления; свойства, величины, параметры, характеристики; отношения, структуры, модели, законы, правила, абстрактные понятия);

7) автоматизация процессов построения тезауруса: частотный, статистический анализ, коррекция статей, сортировка словников, проверка взаимности и непротиворечивости ссылок, составление указателей, конвертация в требуемых форматах и кодировках;

8) ввод в научный оборот и технологическое внедрение.

Для экономии объёма настоящей работы именно здесь подробную библиографию не приводим, однако при необходимости вы можете её получить, напр., по электронной почте [serg@lsw.ru](mailto:serg@lsw.ru) или на сайте <http://lsw.ru>, <http://ЛСВ.РФ> [16].

### Список литературы

1. Bush V. As we may think // *Atlantic Monthly*. – 1945. Vol. 176. – № 1. – P. 101 – 108.
2. Nelson T. *Computer Lib // Dream machines*. Sausalito, CA: Mindful Press, 1974.
3. Nelson T. *Literary machines*. Sausalito, CA: Mindful Press, 1993.
4. Conklin J. Hypertext: an introduction and survey // *Computer*. – 1987. Vol. 20. № 9. – P.20.
5. Субботин М.М. Новая информационная технология: создание и обработка гипертекста // *Научно-техническая информация*. Сер.2. – 1988. – № 5. – С. 2.
6. Субботин М.М. Теория и практика нелинейного письма (взгляд сквозь призму «грамматологии» Ж. Деррида) // *Вопросы философии*. – 1993. – №3. – С.36 – 45.
7. *The New Encyclopedia Britannica*. 15-th edition, Chicago, 1994.
8. *The Encyclopedia of Language and Linguistics*. Vol. 3. England, Pergamon Press, 1994.
9. Першиков В. И., Савинков В. М. *Толковый словарь по информатике*. – М.: Финансы и статистика, 1995. – 544 с.
10. Шуткин Л. В. Паттерновое моделирование гипертекстов // *Научно-техническая информация*. Сер.2. – 1995. – № 9. – С.20 – 25.
11. Евсеев О.В., Савицкая Е.В. *Словарь-гипертекст экономики и бизнеса*. – М.: Изд-во МГОУ; А/О «Росвузнаука», 1993. – 163 с.
12. Анисимова Е. Е. Паралингвистика и текст (к проблеме кролизированных и гибридных текстов) // *Вопросы языкознания*. – 1992. – № 1. – С.71 – 78.
13. Руднев В. П. *Словарь культуры XX века. Ключевые понятия и тексты*. – М.: Аграф, 1997. – 384 с.
14. Иванов Л. Ю. Воздействие новых информационных технологий на русский язык: системно-языковая и культурно-речевая проблематика // *Словарь и культура русской речи. К 100-летию со дня рождения С.И. Ожегова*. – М.: Индрик, 2001. – С.145 – 146.
15. Эпштейн В. Л. *Гипертекст и гипертекстовые системы*. – М.: Ин-т проблем упр. РАН (препринт), 1998. – 39 с.; Эпштейн В. Л. *Введение в гипертекст и гипертекстовые системы*. 02.07.1998 <http://www.ipu.ru/publ/epstn.htm>
16. Лесников С. В. *Словарь русских словарей: более 3500 источников / Предисловие проф. В.В. Дубичинского*. – М.: Азбуковник, 2002. – 334 с.



## CONSTRUCTION OF HYPERTEXT THESAURUS OF META LANGUAGE OF SCIENCE

**S. V. Lesnikov**

*Syktuykar  
State  
University*

*e-mail: serg@lsw.ru*

This article describes the design of hypertext meta-language thesaurus of science supported by the Russian Foundation for Basic Research on proactive, research project (grant) N 11-07-00733 (2011-2013) "Hypertext information-search thesaurus "The meta-language of science" (the structure, mathematical, linguistic, and software; sections of linguistics, mathematics, economics)" (scientific leader - S.V. Foresters).

Key words: algorithm, hypertext, gizaurs, glossary, glottologiya, Internet, computer linguistics, meta-language, lexicography, search, program, Russian language, dictionary, thesaurus, encyclopedia, linguistics, linguistics.