

**Заключение.** В ходе проведенных исследований были предложены принципы построения алгоритма для создания интеллектуальной обучающей системы для изучения иностранного языка с когнитивным подходом. На основании вышеизложенной концепции будет написано техническое задание для программной реализации описанной интеллектуальной обучающей системы.

Открытыми все еще остаются вопросы методики оценивания знаний обучаемого, а также совершенствование обратной связи система-обучаемый, что является темой для дальнейших исследований.

1. *Бабенко А. Е.* E-tutoring в обучении иностранному языку. // Сб. науч. тр. XI Международной науч. конф. им.Т.А.Таран «Интеллектуальный анализ информации» (ИАИ-2011) КПИ, Киев, 17-19 мая 2011 года, издательство «Просвита», стр. 203-209
2. *Валькман Ю. Р., Бабенко А. Е.* E-TUTORING в обучении иностранному языку русскоязычной аудитории // Зб. наук. пр. ПМЕ ім. Г.Є.Пухова, - Моделювання та інформаційні технології, Вип. 61. - 2011, с. 12-18
3. *Валькман Ю. Р., Бабенко А. Е.* Когнитивный поход в обучении иностранному языку. // Зб. наук. пр. ПМЕ ім. Г.Є.Пухова, Вип. 64. – 2012, с.121-129
4. *Валькман Ю. Р., Бабенко А. Е.* Модель процесса интеллектуального обучения с использованием объектно-ориентированного языка JAVA. // Зб. наук. пр. ПМЕ ім. Г.Є.Пухова, Вип. 68. – 2012, с. 83-90.
5. *Дональд Кнут.* Искусство программирования, том 1. Основные алгоритмы . - The Art of Computer Programming, vol.1. Fundamental Algorithms. — 3-е изд. — М.: «Вильямс», 2006. — 720. с.

*Поступила 14.03.2013р.*

УДК 801+008(04)

Е.А.Хала, г.Киев

## **ИНСТРУМЕНТАРИЙ И МЕТОДЫ ПРИКЛАДНОЙ КОМПЬЮТЕРНОЙ ЛИНГВИСТИКИ**

Рассматриваются задачи компьютерной лингвистики. Проводится сравнение особенностей естественного и искусственного языка, методических инструментариев описательной, теоретической и прикладной лингвистики.

*Ключевые слова:* Компьютерная лингвистика, естественный язык, искусственный язык, системы знаков, прикладная лингвистика.

**Введение.** Компьютерная лингвистика — научное направление в области математического и компьютерного моделирования

интеллектуальных процессов у человека и животных при создании систем искусственного интеллекта, которое ставит своей целью использование математических моделей для описания естественных языков. Термин компьютерная лингвистика в последние годы все чаще встречается в связи с разработкой различных прикладных программных систем, в том числе – коммерческих программных продуктов. Связано это бурным ростом в обществе текстовой информации, в том числе в сети Интернет, и необходимостью автоматической обработки текстов на естественном языке. Указанное обстоятельство стимулирует развитие компьютерной лингвистики как области науки и разработку новых информационных и лингвистических технологий [1,11].

В рамках компьютерной лингвистики, существующей уже более 50 лет предложено много перспективных методов и идей, но далеко не все они еще нашли свое выражение в программных продуктах, используемых на практике.

### **1. Задачи компьютерной лингвистики**

Истоки компьютерной лингвистики восходят к исследованиям известного американского ученого Н. Хомского в области формализации структуры естественного языка [2]; ее развитие опирается на результаты в области общей лингвистики [3,6]. Языкознание изучает общие законы естественного языка – его структуру и функционирование, и включает такие области: фонология, морфология, синтаксис, семантика и прагматика, лексикография.

Результаты Н. Хомского, полученные на стыке лингвистики и математики, заложили основу для теории формальных языков и грамматик. Эта теория относится ныне к математической лингвистике и применяется для обработки не столько ЕЯ, но искусственных языков, в первую очередь – языков программирования. По своему характеру это вполне математическая дисциплина.

Таким образом, перед компьютерной лингвистикой стоят, прежде всего, задачи лингвистического обеспечения процессов сбора, накопления, обработки и поиска информации. Наиболее важными из них являются:

1. Автоматизация составления и лингвистической обработки машинных словарей;
2. Автоматизация процессов обнаружения и исправления ошибок при вводе текстов в ЭВМ;
3. Автоматическое индексирование документов и информационных запросов;
4. Автоматическая классификация и реферирование документов;
5. Лингвистическое обеспечение процессов поиска информации в одноязычных и многоязычных базах данных
6. Машинный перевод текстов с одних естественных языков на другие;
7. Построение лингвистических процессоров, обеспечивающих общение пользователей с автоматизированными интеллектуальными

информационными системами (в частности, с экспертными системами) на естественном языке, или на языке, близком к естественному;

8. Извлечение фактографической информации из неформализованных текстов [3].

Машинные словари являются неотъемлемой частью любой системы автоматической обработки текстовой информации. Они могут представлять собой словари слов и/или словари словосочетаний, выражающих устойчивые научно-технические понятия. При составлении словарей необходимо стремиться к тому, чтобы они в максимальной степени отражали лексический состав текстов. Поэтому их нужно составлять по текстам достаточно большого объема (как минимум, по текстам объемом в несколько десятков миллионов лексических единиц). А такая работа может быть выполнена в разумные сроки только на основе широкого применения средств автоматизации.

Задача автоматизированного составления словарей словосочетаний, выражающих научно-технические понятия, сложнее задачи составления словарей слов, поскольку словосочетания в тексте формально не выделены и их границы "отмечены" лишь в сознании человека. Тем не менее, как показывают эксперименты, границы именных словосочетаний могут быть определены с удовлетворительной точностью на основе простых процедур синтаксического анализа, а возникающие при этом ошибки могут быть устранены с помощью статистических методов и путем последующего редактирования составленного словаря.

В системах автоматической обработки текстовой информации важной проблемой является проблема установления парадигматических отношений между понятиями, выраженными отдельными словами или словосочетаниями (отношений типа "синонимия", "род-вид", "целое-часть", "причина-следствие" и других устойчивых в неконтекстных ассоциативных отношений). Такие отношения может устанавливать человек, опираясь на свой интеллект и различного рода словарные пособия. Однако это очень трудная и трудоемкая задача. Поэтому при ее выполнении желательно использовать средства автоматизации и новые программы.

## **2. Особенности языка естественного и искусственного**

Язык представляет собой, так называемую систему кодов, знаков представленную на рис.1. Знак – это, ничто иное, как любой воспринимаемый чувственно предмет (Визуального, Аудиального или другого), служащий представителем другого предмета и носителем информации об этом предмете (знаки-образы: фотографии, копии различных документов, отпечатки пальцев; знаки-символы – к примеру, алфавитные буквы, музыкальные ноты, знаки Морзе[4,5]).

По своему происхождению языки бывают естественные и искусственные[6].

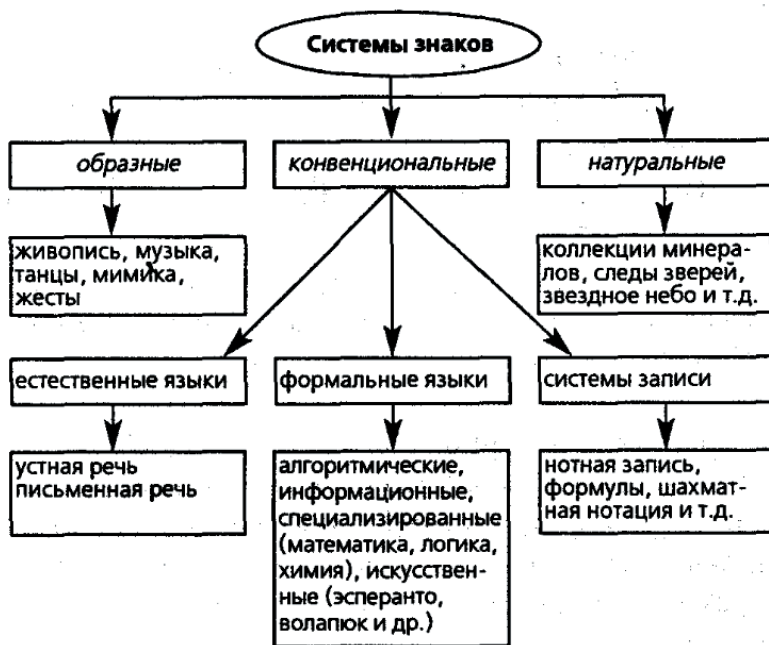


Рис.1. Системы знаков

Естественные языки – это исторически сложившиеся в обществе звуковые (речь), а затем и графические (письмо) информационные знаковые системы.

Исходя из семантического статуса [7,10] естественного языка отмечаем следующее:

- Исходя из того что, язык представляет собой совокупность правил, следовательно, существует огромное количество естественных языков. Материальный базис любого языка, естественного происхождения, является многомерным, это означает, что он разделяется на зрительные, вербальные, осязательные разновидности знаков. Все эти разновидности не зависят друг от друга, но в большом количестве существующих сегодня языков они неразрывно связаны, причем главные - вербальные символы.
- Материальный базис языка, естественного происхождения, изучается лишь в двух измерениях – вербальном и зрительном, иначе письме.

Искусственные языки – это вспомогательные знаковые системы, создаваемые на базе естественных языков для точной и экономной передачи научной и другой информации. Они конструируются с помощью естественного языка или ранее построенного искусственного языка. Язык, выступающий средством построения или изучения другого языка, называют

метаязыком, основой – языком-объектом. Метаязык, как правило, обладает более богатым по сравнению с языком-объектом выразительными возможностями.[11]

Всякий искусственный язык имеет три уровня организации:

- синтаксис – уровень структуры языка, где формируются и исследуются отношения между знаками, способы образования и преобразования знаковых систем;
- синематику, где исследуются отношения знака к его смыслу (значению, под которым понимается либо мысль, выражаемая знаком, либо объект, обозначаемый им);
- прагматику, где исследуются способы употребления знаков в данном сообществе, использующем искусственный язык.

Построение искусственного языка начинается с введения алфавита, т.е. набора символов, которыми обозначаются объект данной науки, и правил построения формул данного языка. Часть правильно построенных формул принимается за аксиомы. Таким образом, все знание, оформленное с помощью искусственного языка, приобретает аксиоматизированную форму, а вместе с нею доказательность и достоверность.[9]

Характерной особенностью искусственных языков является однозначная определенность их словаря, правил образования выражений и придания им значений. Во многих случаях эта особенность оказывается преимуществом таких языков в сравнении с естественными языками, аморфными как со стороны словаря, так и со стороны правил образования и значения.[8]

### **3. Инструментарий компьютерной лингвистики.**

Компьютерная лингвистика как особая прикладная дисциплина выделяется прежде всего по инструменту – т.е. по использованию компьютерных средств обработки языковых данных. Существуют общие принципы компьютерного моделирования мышления, которые так или иначе реализуются в любой компьютерной модели. В их основе лежит теория знаний, первоначально разрабатывавшаяся в области искусственного интеллекта, а в дальнейшем ставшая одним из разделов когнитивной науки. Важнейшими понятийными категориями компьютерная лингвистика являются такие структуры знаний, как «фреймы» (понятийные, или, как принято говорить, концептуальные структуры для декларативного представления знаний о типизированной тематически единой ситуации), «сценарии» (концептуальные структуры для процедурного представления знаний о стереотипной ситуации или стереотипном поведении), «планы» (структуры знаний, фиксирующие представления о возможных действиях, ведущих к достижению определенной цели). Тесно связано с категорией фрейма понятие «сцена». Категория сцены преимущественно используется в литературе по компьютерной лингвистике как обозначение концептуальной структуры для декларативного представления актуализованных в речевом

акте и выделенных языковыми средствами (лексемами, синтаксическими конструкциями, грамматическими категориями и пр.) ситуаций и их частей.

Определенным образом организованный набор структур знаний формирует «модель мира» когнитивной системы и ее компьютерной модели. В системах искусственного интеллекта модель мира образует особый блок, в который в зависимости от выбранной архитектуры могут входить общие знания о мире (в виде простых пропозиций типа «зимой холодно» или в виде правил продукций «если на улице идет дождь, то надо надеть плащ или взять зонтик»), некоторые специфические факты («Самая высокая вершина в мире – Эверест»), а также ценности и их иерархии, иногда выделяемые в особый «аксиологический блок».

Большинство элементов понятий инструментария компьютерной лингвистики омонимично: они одновременно обозначают некоторые реальные сущности когнитивной системы человека и способы представления этих сущностей, используемые при их теоретическом описании и моделировании. Иными словами, элементы понятийного аппарата компьютерной лингвистики имеют онтологический и инструментальный аспекты. Например, в онтологическом аспекте разделение декларативных и процедурных знаний соответствует различным типам знаний, имеющимся у человека – так называемым знаниям ЧТО (декларативным; таково, например, знание почтового адреса какого-либо NN), с одной стороны, и знаниям КАК (процедурным; таково, например, знание, позволяющее найти квартиру этого NN, даже не зная ее формального адреса) – с другой. В инструментальном аспекте знание может быть воплощено в совокупности дескрипций (описаний), в наборе данных, с одной стороны, и в алгоритме, инструкции, которую выполняет компьютерная или какая-либо другая модель когнитивной системы, с другой [7].

#### **4. Прикладная лингвистика**

Прикладная лингвистика – деятельность по приложению научных знаний об устройстве и функционировании языка в нелингвистических научных дисциплинах и в различных сферах практической деятельности человека, а также теоретическое осмысление такой деятельности [12].

С функциональной точки зрения, прикладная лингвистика может быть определена как академическая дисциплина, в которой целенаправленно изучаются и разрабатываются способы оптимизации различных сфер функционирования языковой системы. Функции языка задают точки отсчета для классификации огромной области приложения лингвистических знаний. В оптимизацию коммуникативной функции вносят вклад такие дисциплины, как теория перевода, машинный перевод, теория и практика преподавания родного и неродного языка, теория и практика информационно-поисковых систем, создание информационных и, шире, искусственных языков, теория кодирования. Для оптимизации социальной функции языка – как части коммуникативной – значимы исследования в области социолингвистики (в

частности, изучение и обоснование языковой политики), орфографии и орфоэпии, теории воздействия, политической лингвистики. Оптимизации гносеологической функции способствуют лексикография (в том числе компьютерная), исследования в области терминологии и терминографии, корпусная и полевая лингвистика. Наконец, целям оптимизации когнитивной функции служат исследования по компьютерной лингвистике, «лингвистической криминологии», психолингвистике и афизиологии, квантитативной лингвистике.

### **5. Методы прикладной лингвистики**

Разнообразие методов прикладной лингвистики вполне сопоставимо с разнообразием конкретных областей приложения научных знаний о языке: каждая конкретная прикладная дисциплина обладает своим уникальным набором методов. Тем не менее можно выделить нечто общее, характерное для методов прикладной лингвистики в целом. Эта общая часть хорошо видна при сравнении методических инструментариев описательной, теоретической и прикладной лингвистики.

Перед описательной лингвистикой стоит задача описания фактов языка. На первом плане при этом находится метод классификации, т.е. выявления той сетки параметров, которая позволяет охватить все существенные свойства языковых структур. Теоретическая лингвистика формирует само представление о том, какие свойства языка являются существенными, а какие – нет. Создаваемые в теоретической лингвистике концептуальные модели языка не просто описывают наблюдаемые факты, но и претендуют на их объяснение. При этом как описательная, так и теоретическая лингвистика исходят из познавательной установки, известной как «God's Truth» («Божественная Истина»). Иными словами, классификации языковых фактов и концептуальные модели теоретической лингвистики претендуют на описание того, как действительно устроен язык.

Прикладная лингвистика также использует и метод классификации, и метод моделирования. Однако поскольку задачи прикладной лингвистики сосредоточены в области оптимизации функций языка, а оптимизация определяется конкретной задачей, то в прикладной лингвистике широкое распространение имеет познавательная установка, известная под несколько ироническим названием «Фокус-покус» (более уважительное название – «инженерный подход»). Эта установка в качестве основной ценности выдвигает не познание того, «как все обстоит на самом деле», а решение конкретной задачи, в частном случае – удовлетворение требований «заказчика», преследующего свои собственные цели. Это, впрочем, не означает, что результаты прикладных исследований не представляют никакой ценности для теории языка: напротив, прикладные модели оказывают значительное влияние на лингвистическую теорию, способствуя обновлению концептуального аппарата современного языкознания [10,12].

В силу сказанного важнейшим общим свойством методов прикладной лингвистики является ориентация на создание такой модели данной проблемной области, которая отображает только те ее свойства, которые необходимы для решения конкретной практической задачи. Иными словами, если теоретическое исследование в идеале стремится к полному описанию проблемной области со всеми ее сложностями и т.п., то прикладное описание обязано удовлетворять только данной конкретной задаче. Оно, конечно, может обладать и большей, чем это необходимо для решения конкретной прикладной задачи, описательной адекватностью и объяснительной силой, но принципиально этого от него не требуется.

Суммируя основные частные отличия прикладных моделей от теоретических и описательных, можно сказать, что прикладные модели в целом ориентированы на конкретные подязыки, а не весь язык в целом; они часто (но не всегда) требуют большей степени формализации; прикладные модели используют знания о языке выборочно; прикладные модели не делают различий между собственно лингвистическими и экстралингвистическими аспектами семантики языковых выражений; прикладные модели в существенно большей степени огрубляют моделируемый объект, чем теоретические модели и, наконец, прикладные модели не налагают никаких существенных ограничений на инструмент моделирования

**Заключение.** Компьютерная лингвистика демонстрирует вполне осязаемые результаты в различных приложениях по автоматической обработке текстов на естественном языке. Дальнейшее ее развитие зависит как от появления новых приложений, так и независимой разработки различных моделей языка, в которых пока не решены многие проблемы.

Современная прикладная лингвистика почти столь же многообразна, как и области практической деятельности человека. К наиболее перспективным прикладным областям в первую очередь следует отнести гипертекстовые технологии, непосредственно связанные с эксплуатацией и развитием глобальной компьютерной сети Internet.

1. *Садур В.Г.* Речевое общение с электронно-вычислительными машинами и проблемы их развития. – В кн.: Речевое общение: проблемы и перспективы. М., 1983
2. *Баранов А.Н.* Категории искусственного интеллекта в лингвистической семантике. Фреймы и сценарии. М., 1987
3. *Кобозева И.М., Лауфер Н.И., Сабурова И.Г.* Моделирование общения в человеко-машинных системах. – Лингвистическое обеспечение информационных систем. М., 1987.
4. *Олкер Х.Р.* Волшебные сказки, трагедии и способы изложение мировой истории. – В кн.: Язык и моделирование социального взаимодействия. М., 1987.
5. *Городецкий Б.Ю.* Компьютерная лингвистика: моделирование языкового общения. – Новое в зарубежной лингвистике. Вып. XXIV, Компьютерная лингвистика. М., 1989.



6. Макьюин К. Дискурсивные стратегии для синтеза текста на естественном языке. – Новое в зарубежной лингвистике. Вып. XXIV, Компьютерная лингвистика. М., 1989.
7. Попов Э.В., Преображенский А.Б. Особенности реализации ЕЯ-систем. – Искусственный интеллект. Кн. 1, Системы общения и экспертные системы. М., 1990.
8. Преображенский А.Б. Состояние развития современных ЕЯ-систем. – Искусственный интеллект. Кн. 1, Системы общения и экспертные системы. М., 1990.
9. Субботин М.М. Гипертекст. Новая форма письменной коммуникации. – ВИНТИ, Сер. Информатика, 1994, т. 18.
10. Баранов А.Н. Введение в прикладную лингвистику. М., 2000.
11. Звегинцев В.А. Теоретическая и прикладная лингвистика. М., 1968.

*Поступила 28.02.2013р.*

УДК 681.61

І.О. Ляшенко, м. Київ

## **МЕТОД РОЗПОДІЛУ ФУНКЦІОНАЛЬНИХ РЕСУРСІВ ІНФОРМАЦІЙНО-УПРАВЛЯЮЧОЇ СИСТЕМИ СПЕЦІАЛЬНОГО ПРИЗНАЧЕННЯ В УМОВАХ ВІДМОВИ ФУНКЦІОНАЛЬНИХ ЕЛЕМЕНТІВ**

Запропоновано метод розподілу обмежених функціональних ресурсів за допомогою агентно-орієнтованого підходу для забезпечення живучості інформаційно-управляючих систем спеціального призначення

*Ключові слова:* живучість, відмови, функціональні елементи, інтелектуальні агенти

Предложен метод распределения функциональных ресурсов при помощи агентно-ориентированного подхода для обеспечения живучести информационно-управляющих систем специального назначения

*Ключевые слова:* живучесть, отказы, функциональные элементы, интеллектуальные агенты

The method of allocation of functional resources is offered through intellectual agents approach for providing of vitality information-control systems of the special setting

*Keywords:* vitality, refuses, functional elements, intellectual agents

### **Вступ**

Бурхливий розвиток обсягу завдань оброблення інформації, інтеграція локальних, корпоративних та регіональних мереж в глобальну мережу Інтернет, недосконалість інформаційних технологій, що застосовуються в мережах (система адресації, прикордонні шлюзові протоколи, відкритість та незашифрованість трафіку, розповсюдженість так званого мелверу), а також децентралізація