

# Информационные технологии и лингвистика



# ЛИНГВИСТИКА

- лингвистика (языкознание, языковедение) - «наука о естественном человеческом языке вообще и о всех языках мира как индивидуальных его представителях»

«Лингвистический энциклопедический словарь» (1990)

# ЛИНГВИСТИКА

- Ю. С. Маслов пишет, что она исследует сущность и природу языка, проблему его происхождения и общие законы его развития и функционирования

# ЛИНГВИСТИКА

- Теоретическая лингвистика
- Прикладная лингвистика

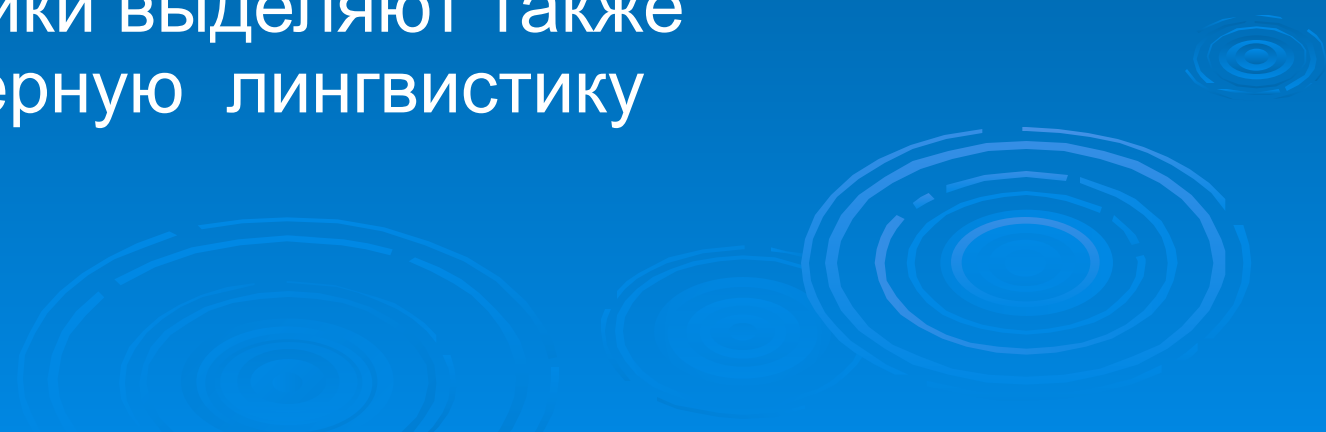
# Теоретическая и прикладная ЛИНГВИСТИКА

- С 60-70 гг. – **взаимосвязаны и дополняют друг друга**
- 70-е гг. Р.Г.Пиотровский: структурное и математическое языкознание, «новая» лингвистика-инженерная лингвистика (вычислительная лингвистика, экспериментальная фонетика, лингвистическое обеспечение систем научно технической информации )

# Теоретическая и прикладная ЛИНГВИСТИКА

80-е гг. Ленинградская школа:  
теоретическая лингвистика,  
экспериментальная лингвистика и  
прикладная лингвистика (Богданов В. В.,  
Бондарко Л. В., Буторов В. Д. и др. Прикладная лингвистика и теория  
языка ).

В последние годы в рамках прикладной  
лингвистики выделяют также  
компьютерную лингвистику



# Прикладная лингвистика

- В. А. Звегинцев. Прикладная лингвистика - новую область лингвистики, «которая осуществляет реализацию лингвистических знаний с целью решения всякого рода практических задач»

«Прикладная лингвистика представляет новый взгляд на задачи изучения языка.

Исходя из этого нового взгляда, она производит переоценку достигнутого в науке о языке, направляет по определенному руслу лингвистические исследования и, конечно, комплектует собственную тематику».

# Прикладная лингвистика

- Прикладная лингвистика - «направление в языкознании, занимающееся разработкой методов решения практических задач, связанных с использованием языка».

«Лингвистический энциклопедический словарь».



# Прикладная лингвистика

**Задачи:** традиционные, или «вечные», и «новые».

**«Вечные»:** создание и совершенствование письменности, разработка систем транскрипции устной речи, систем транслитерации иноязычных слов, унификацию и стандартизацию научно-технической терминологии, создание словарей различных типов, перевод с языка на язык, обучение языку и т.д.

**Новые:** те, появление которых обусловлено современной научно-технической революцией, характеризующейся укреплением взаимосвязи общественных, естественных и технических наук.

# Прикладная лингвистика

- **А. Е. Кибрик.** «Прикладная лингвистика — раздел языкознания, в котором разрабатываются методы решения практических задач, связанных с оптимизацией использования языка как важнейшего средства человеческой коммуникации» .
- **Четыре функции:**
  - 1) оптимизация способов фиксации и хранения речевой информации;
  - 2) оптимизация способов передачи информации;
  - 3) оптимизация интеллектуальных способностей человека, связанных с использованием языка;
  - 4) оптимизация использования языка как средства массовой коммуникации.

# ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ И ПРИЧИНЫ, СПОСОБСТВУЮЩИЕ ИХ ПОЯВЛЕНИЮ

- информационные технологии - совокупность законов, методов и средств получения, хранения, передачи, распространения, преобразования информации с помощью компьютеров.

## Достижения научно-технического прогресса, способствовавшие развитию ИТ

- 1) создание **персональных компьютеров** с большой памятью и большой скоростью выполнения операций;
- 2) разработка **звуковых плат**, дающих возможность воспроизводить и записывать речь, звуки и музыку в большом диапазоне частот;
- 3) изобретение **видеоплат**, позволяющих выводить на экран компьютеров изображение с телеэкранов и видеомагнитофонов;
- 4) разработка специальных **(мультимедийных)** компьютеров, позволяющих воспроизводить на экране дисплеев цвет, звук, музыку, движение;

## Достижения научно-технического прогресса, способствовавшие развитию ИТ

- 5) создание специальных процессоров и устройств, способных передавать информацию в сети от одного компьютера к другому;
- 6) разработка устройств электронной связи (модемов), позволяющих передавать информацию на далекие расстояния (по телефонным линиям, кабелям, радиоканалам и т.п.);
- 7) создание электронной оргтехники, связанной с персональными компьютерами и позволяющей осуществлять высокоскоростную печать документов, их копирование и размножение.

# ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ В ЛИНГВИСТИКЕ

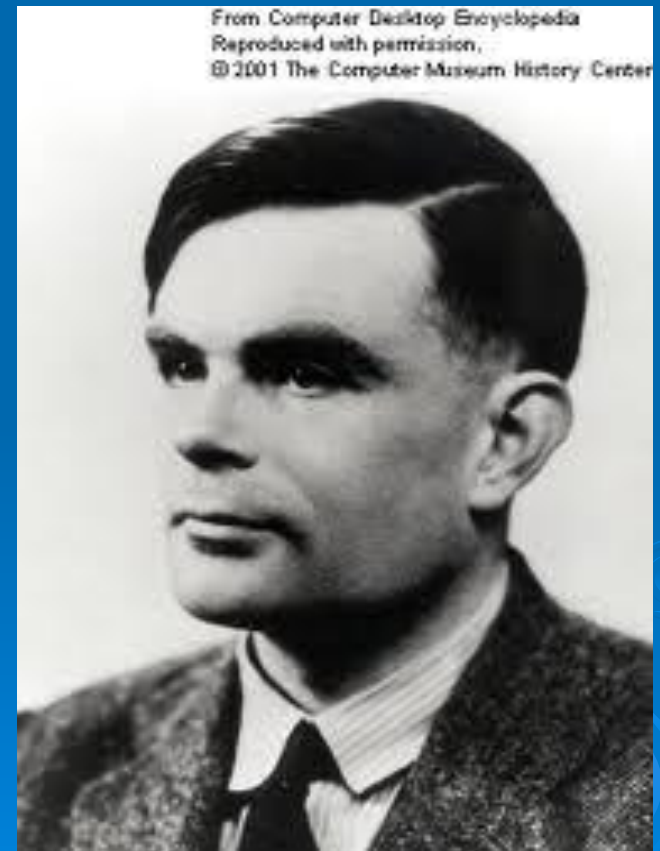
- это совокупность законов, методов и средств получения, хранения, передачи, распространения, преобразования информации о языке и законах его функционирования с помощью компьютеров.

# Информационные технологии и ЛИНГВИСТИКА

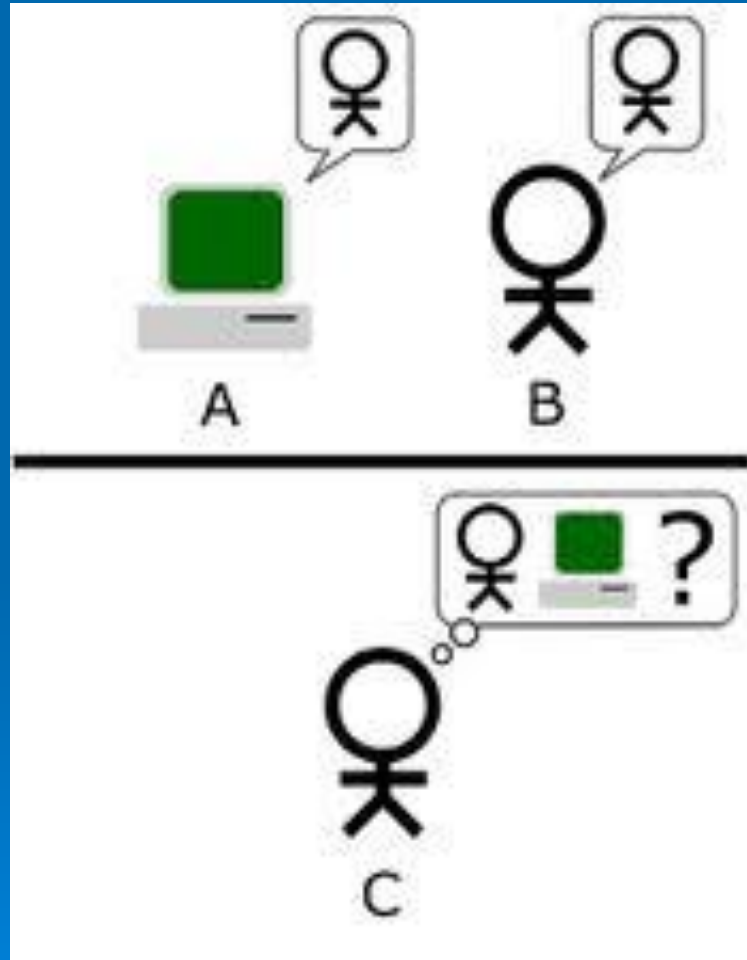
- 1) создание систем искусственного интеллекта

«Может ли машина мыслить?»  
1950-ый г.

Алан Тьюринг



# Тест Тьюринга





ДОПОЛНИТЕЛЬНЫЕ ОЧКИ  
В ТЕСТЕ ТЬЮРИНГА:

УБЕДИТЬ ЭКЗАМЕНАТОРА,  
ЧТО ЭТО ОН КОМПЬЮТЕР.

ЗНАЕШЬ, ТЫ ГОВОРИШЬ  
ОЧЕНЬ УБЕДИТЕЛЬНО.

/

Я... Я УЖЕ НЕ УВЕРЕН,  
КТО Я НА САМОМ ДЕЛЕ.



# «Элиза» 1966 г.



# Гипотеза Ньюелла-Саймона 1976 г.

- Если мы полагаем, что животное, или человек, или машина действуют осмысленно, то значит, они каким-то образом выполняют символьные вычисления. И наоборот, так как компьютер способен к подобным вычислениям, то на его основе *может* быть создан *искусственный интеллект*.

# Машина VS Человек



# DeepBlue



# Философия

«Каковы последствия создания  
ИИ для человечества?»



# Информационные технологии и ЛИНГВИСТИКА

- 2) создание систем автоматического перевода;
- 3) создание систем автоматического аннотирования и реферирования текстов;
- 4) создание систем порождения текстов;
- 5) создание систем обучения языку;
- 6) создание систем понимания устной и письменной речи;
- 7) создание систем генерации речи;

## □ Синтаксис:

«Мы отдали бананы обезьянам, потому что они были голодные» и «Мы отдали бананы обезьянам, потому что они были перезрелыми».

**Свободный порядок слов:** «Бытие определяет сознание» — кто кого определяет?



## Лексика, морфология:

В речи могут встретиться неологизмы, например, глагол «Пятидесятирублею» — то есть вышлай 50 рублей. Система должна уметь отличать такие случаи от опечаток и правильно их понимать.

## Омонимы. Фонетические омонимы.

«Серый волк в глухом *лесу* встретил рыжую *лису*»

(Кроме того, что лиса может быть рыжей, а лес — глухим, лес также может быть рыжим (характеристика, в данном случае обозначающая преобладающий цвет листвы в лесу), в то время как лиса может быть глухой, что порождает дополнительную проблему, вытекающую из предыдущей.)

# Информационные технологии в филологии

- 8) создание автоматизированных информационно-поисковых систем;
- 9) создание систем атрибуции и дешифровки анонимных и псевдоанонимных текстов;
- 10) разработка различных баз данных (словарей, карточек, каталогов, реестров и т.п.) для гуманитарных наук;
- 11) разработка различного типа автоматических словарей;
- 12) разработка систем передачи информации в сети Интернет т.д.

Эти комплексные задачи включают целый ряд более мелких проблем. К их числу относится автоматизация следующих процессов:

- 1) построение словарей текстов;
- 2) морфологический анализ слова;
- 3) определение значения многозначного слова;
- 4) синтаксический анализ предложения;
- 5) поиск слова в словаре;
- 6) порождение предложения и т. д.