

# Квантитативная лингвистика и новые информационные технологии

Ворожцова Ирина Борисовна

# Содержание дисциплины

- **Прикладная лингвистика**
- **Компьютерная лингвистика**
- **Квантитативная лингвистика**
- **Корпусная лингвистика**
- **Машинный перевод**

# Тема 1. Прикладная лингвистика

- Прикладная лингвистика (ПЛ) как раздел языкознания для решения практических задач, связанных с языками и речевой деятельностью: преподавание языков, перевод, информационный поиск и обработка данных.

# Из истории ПЛ

- Термин «прикладная лингвистика» появился в 50-х годах XX века.

*applied linguistics,*  
*angewandte Linguistik (anwenden)*  
*linguistique appliquée,*  
*lingüística aplicada,*  
*приложна лингвистика,*  
*прикладной кылтодон,*  
*Примењена лингвистика,*  
*anvendt lingvistik,*  
*linguistica applicata*

- Первая конференция по прикладной лингвистике прошла в начале 60-х.

В западной лингвистике в 20-х годах прошлого века и относится к областям:  
преподавание иностранных языков: методика преподавания, особенности описания грамматики для учебных целей, преподавание языка как родного и иностранного и т.д.

# Определение ПЛ (2012)

- **Applied linguistics** is an interdisciplinary field which identifies, investigates, and offers solutions to language-related real-life problems. Some of the academic fields related to **applied linguistics** are education, psychology, communication research, anthropology, and sociology. Major branches of **applied linguistics** include bilingualism and multilingualism, conversation analysis, contrastive **linguistics**, sign **linguistics**, language assessment, literacies, language assessment, literacies, discourse analysis, language assessment, literacies, discourse analysis, language pedagogy, language assessment, literacies, discourse analysis, language pedagogy, second language acquisition, language assessment, literacies, discourse analysis, language pedagogy, second language acquisition, language planning, language assessment, literacies, discourse analysis, language pedagogy, second language acquisition,

# ПЛ в СССР

В СССР прикладная лингвистика была заявлена в 50-е гг. Разработка компьютерных технологий и появление систем автоматической обработки информации (систем машинного перевода, систем информационного поиска, автоматизированных систем обработки текста и пр.).

«Прикладная лингвистика» = «компьютерная лингвистика», «вычислительная лингвистика», «автоматическая лингвистика», «инженерная лингвистика».

В. А. Звегинцев: «Под прикладной лингвистикой чаще всего понимают все виды автоматической обработки речевой информации (Language-data Processing) — машинное распознавание устной речи, машинный перевод, автоматическую классификацию технических и иных документов, автоматическое аннотирование текстов, автоматическое кодирование и пр. Но все же было бы неправильно замыкать прикладную лингвистику в пределах данной проблематики».

# ПЛ в разных традициях

Прикладная лингвистика в советское время — это больше исследование по взаимодействию с ЭВМ на естественном языке. Оно активно развивалось в 1960–1970-е годы. Позже исследования, связанные с автоматической обработкой текстов, то есть использование знаний о языке для создания программ, которые достают информацию из текста, обеспечивают взаимодействие человека с компьютером на естественном языке. Работы по машинному переводу, который сейчас очень активно развивается. Был период с 1990-х по 2000-е годы, когда работ было мало.

- сфера политической лингвистики, и это исследование политического языка.
- изучение языка рекламы,
- лингвистическое консультирование (лингвистическая экспертиза текста, обеспечение политиков необходимыми знаниями о языке, работа спичрайтеров)

На Западе (и в Германии, и в США, и во Франции) это обучение иностранному языку и переводу. в Дублинском университете *School of Applied Languages* – изучение языка с прикладными целями, в первую очередь это школа перевода. *applied linguistics* в англоязычной традиции – обучение языку; в начале 1990-х годов, — это упрощение языка законов и законодательства.

# Область применения ПЛ

- Решение практических задач, связанных с использованием языка и речи:
- Преподавание языка
- Создание алфавитов для бесписьменных языков

Чаще всего  
междисциплинарного  
характера

Оптимизация знаний о языке, и именно это объединяет абсолютно разные направления, которые существуют в сфере прикладной лингвистики.

Баранов А.Н.

«Множественность грамматик» в толковании Ю.В. Рождественского



# Кто? Что читать? Что слушать?

- **Анатолий Баранов**, доктор филологических наук, заведующий отделом экспериментальной лексикографии Института русского языка им. Виноградова РАН.
- Баранов А.Н. Введение в прикладную лингвистику. Учебное пособие. – Либроком, 2016.
- Баранов А.Н. Прикладная лингвистика на ютубе 2015 г. Лекция на «Постнауке»

<https://www.youtube.com/watch?v=3BzrupwFSu>

8

# Тема 2: Компьютерная лингвистика

**Компьютерная лингвистика** — прикладная область лингвистики, это набор задач, которые можно решить при помощи компьютера, используя встроенные ИКТ и параметры обработки текстов, созданных на определенном языке.

Компьютерная лингвистика: информационные технологии в лингвистике.

Междисциплинарная область, связанная с необходимостью обработки данных на естественном языке

Занимается задачами, решаемыми с помощью компьютера.

АОТ – автоматическая обработка текстов: информационный поиск, информационное реферирование и аннотирование текстов (сжатие текстов), рубрицирование, автоматическое извлечение фактов из текста,

Анализ определенных текстовых параметров (тональность текста)

Инженерная лингвистика: АСУ

Вычислительная лингвистика...

Математическая лингвистика

Что от лингвистики? Фонология, морфология, синтаксис, семантика, прагматика, лексикография, психолингвистика

Информатика

Искусственный интеллект

# Из истории использования ЭВМ в лингвистических исследованиях

- Создание точных описаний языкового и речевого материала – математические методы в лингвистике, математические модели в лингвистике, машинный перевод, информационный поиск, информационная обработка текстов (создание аннотаций и рефератов), начиная с 50-х годов. Коммерческого продукта не состоялось: требовались ЭВМ с огромной памятью, автоматическим вводом данных и быстродействием.
- Сегодняшняя ситуация: есть компьютеры, отвечающие на запрос лингвистических исследований: бездонная память, автоматический ввод данных, быстродействие.

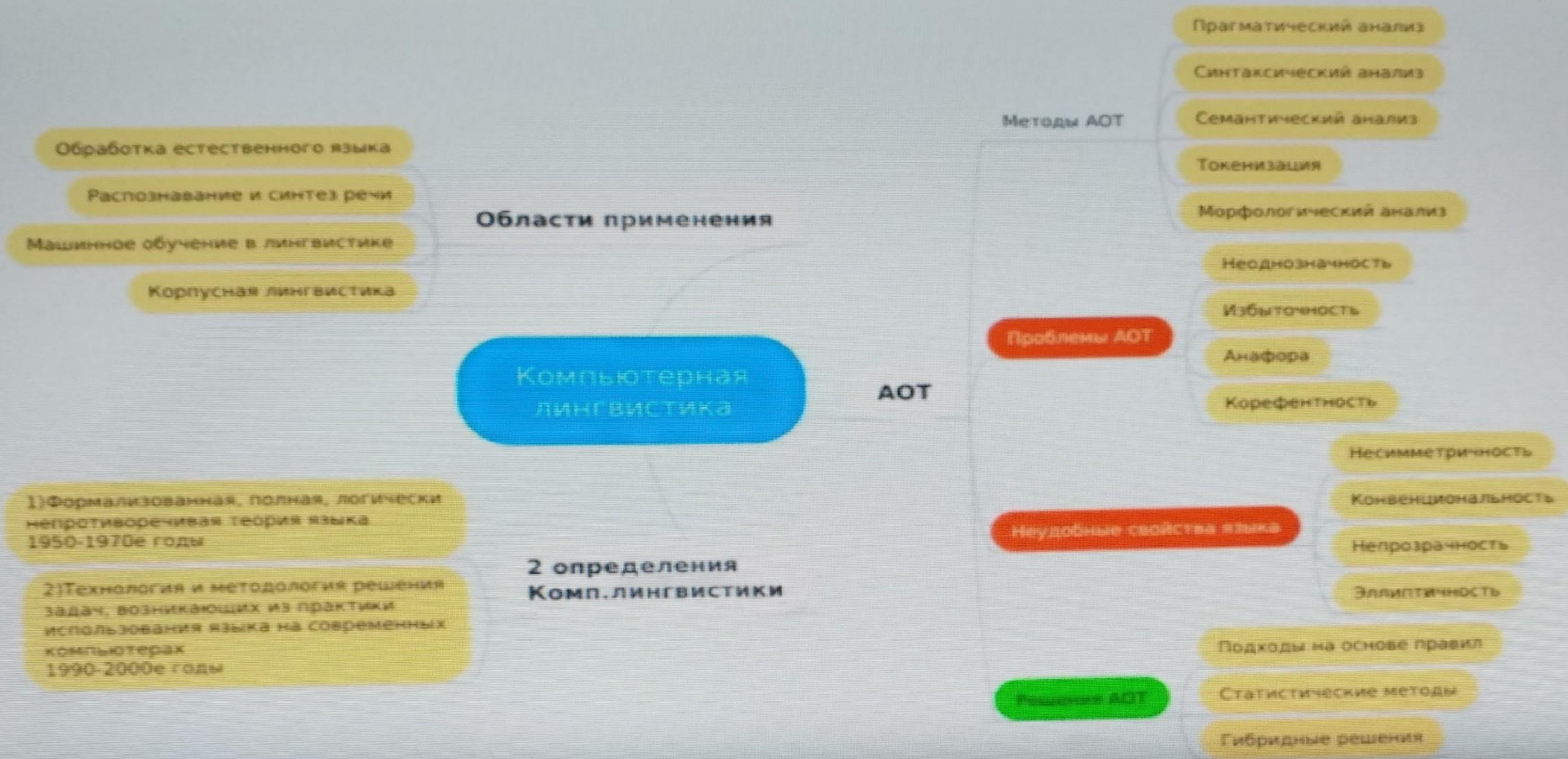
# Селегей Владимир Павлович

- Директор по лингвистическим исследованиям компании АВВУУ, заведующий кафедрами компьютерной лингвистики РГГУ и МФТИ, председатель оргкомитета конференции по компьютерной лингвистике «Диалог».
- Селегей: Теория и практика компьютерной лингвистики (Чердак). 09.08.2016.



<https://www.youtube.com/watch?v=zwgbbg0pD8A>

# Компьютерная лингвистика: общая схема



# Данные схемы

- Определения КЛ: 1) формализованная, полная, логически непротиворечивая теория языка (1950-1970 гг.); 2) технология и методология решения задач, возникающих из практики использования ЕЯ на современных компьютерах.
- Области применения: обработка ЕЯ; распознавание и синтез речи; машинное обучение в лингвистике; корпусная лингвистика.



# Данные схемы (АОТ)

- Методы АОТ: прагматический А. – синтаксический А. – семантический А. – токенизация – морфологический А.
- Проблемы АОТ: неоднозначность – избыточность – анафора – кореференция
- Неудобные свойства языка: несимметричность – конвенциональность – непрозрачность – эллиптичность (*Петя любит Машу, а Маша – Петю*)
- Решения АОТ: подходы на основе правил – статистические методы – гибридные решения

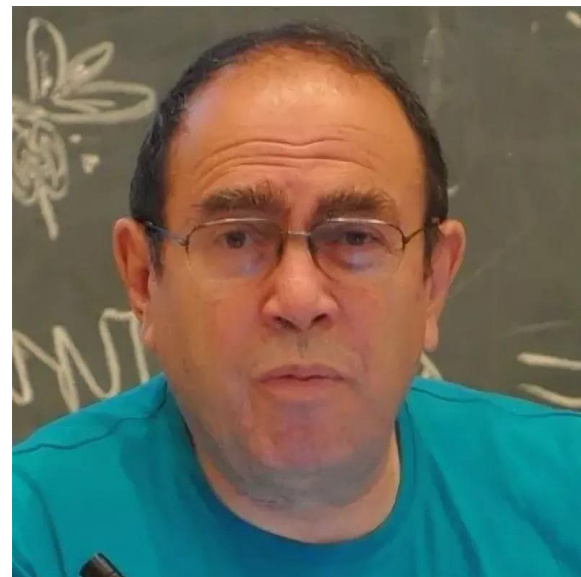
# Словечки

- Бот от робот – специальная программа Бот от робот – специальная программа, выполняющая автоматически Бот от робот – специальная программа, выполняющая автоматически и/или по заданному расписанию Бот от робот – специальная программа, выполняющая автоматически и/или по заданному расписанию какие-либо действия через интерфейсы Бот от робот – специальная программа, выполняющая автоматически и/или по заданному расписанию какие-либо действия через интерфейсы, предназначенные для людей.
- Спеллчекер
- Токен - компактное устройство, предназначенное для обеспечения информационной безопасности пользователя, также используется для идентификации его владельца (авторизация)
- Блокчейн – цепочка блоков данных, которые



# Иомдин Леонид Лейбович

- **ИОМДИН** Леонид Лейбович, зав.лаб. компьютерной лингвистики Института проблем передачи информации (ИППИ) РАН, МФТИ, РГГУ, ШАД (Школа Анализа Данных) Яндекса
- Компьютерная лингвистика, на радио «Говорит Москва», источник Sinus, от 19 окт. 2018 г.  
<https://www.youtube.com/watch?v=UMBsM2x5kr4>
- КЛ: К.морфология, К. синтаксис, К представление значений, Распознавание и синтез речи, Машинное обучение в лингвистике, Корпусная Л., МП, Информационный поиск, Извлечение информации, Диалоги и чат-боты, Анализ тональности, К.текстология, Квантит Л.(что можно сосчитать в языке и речи?), Речевое воздействие и манипулирование
- Компьютерная лингвистика №3. Синтаксический анализ (Чердак). 01.07.2016  
<https://www.youtube.com/watch?v=ZFBOT01bT-g>



# Предмет КЛ: взгляд лингвиста

- Распознавать тексты? Их надо описать на языке, который понимает машина. Есть правила для текстов, но машина так не понимает. Зато умеет обрабатывать большие объемы с огромной скоростью
- Например, столкнулись с тем, что люди говорят не словами, а конструкциями. Как выделить конструкции?
- Особым образом устроенная теория языка – формализованная. Полная, логически непротиворечивая, чтобы могла понять машина для обработки данных для определения правильности предложения, построения его языковой структуры, определения смыслового тождества высказываний и др.
- Слово как линейная последовательность знаков от пробела до пробела

# Как это работает?

- **Компьютерный лингвист** не пишет код и не исследует язык. Он пытается понять, как сложные языковые схемы можно сделать проще, понятней для компьютера. Лингвист привык работать с языковым материалом, поэтому ему легче улавливать тенденции и формализовать их, дело **программиста** – написать код, который эти схемы встраивал бы в компьютерный продукт.
- **Подход** — ввести только то, что активно используется носителями языка в определенных ситуациях. Эти слова, эта лексика активного словарного запаса помещается в те ячейки памяти, к которым легко обеспечить быстрый доступ. А все остальное скидывается уже в более долговременную память. И это позволяет программе очень быстро работать. Как в программах проверки орфографии, «спел-чекерах». Если вогнать весь словарь, например 200 тысяч слов, то все равно такой парсер даже при имеющихся очень эффективных программах обработки и очень эффективном «железе» будет работать довольно долго.

# Как лингвисты ставят задачи айтишникам

Рахилина Екатерина Влад., российский лингвист, д.ф.н., профессор НИУ ВШЭ.

Современная лингвистика и компьютерные технологии:

[https://www.youtube.com/watch?v=\\_bTVHL5jDtk](https://www.youtube.com/watch?v=_bTVHL5jDtk)

# Проблемы

- Неоднозначность языковых единиц
- Избыточность
- Метафоричность
- Понимание текста требует учета огромного количества знаний: о мире, системы представлений
- Как описать мир для машины?

# Что может компьютер?

- Огромная память
- Быстродействие
- Обработка огромных объемов данных: через тематическое моделирование, кластеризацию, изучение дистрибуции единиц, моделирование языка, статистическую обработку
- Выявление фактов,
- Выявление отношений,
- дистрибутивная семантика,
- тематическое моделирование,
- определение авторства.

# Что умеют компьютеры?

Умеют распознавать слова, их формы, их синтаксическую функцию, семантику. Одно из главных условий для машинного обучения компьютера распознаванию лингвистических данных – огромные объемы текстов в электронном виде. В середине века компьютеры занимали огромные помещения, были ламповыми, а жесткие диски на несколько сотен мегабайт стоили запредельные суммы, которые были не под силу индивидуальным пользователям. Сейчас хранение больших текстов в «цифре» не проблема – на телефоне, на флэшке можно легко уместить собрание сочинений. Поэтому, хотя сами идеи и были заложены достаточно давно, их начали серьезно применять только в последние 15–20 лет.

Как они это делают? По правилам. Эти правила введены в оболочку ИКТ. Или благодаря машинному обучению.

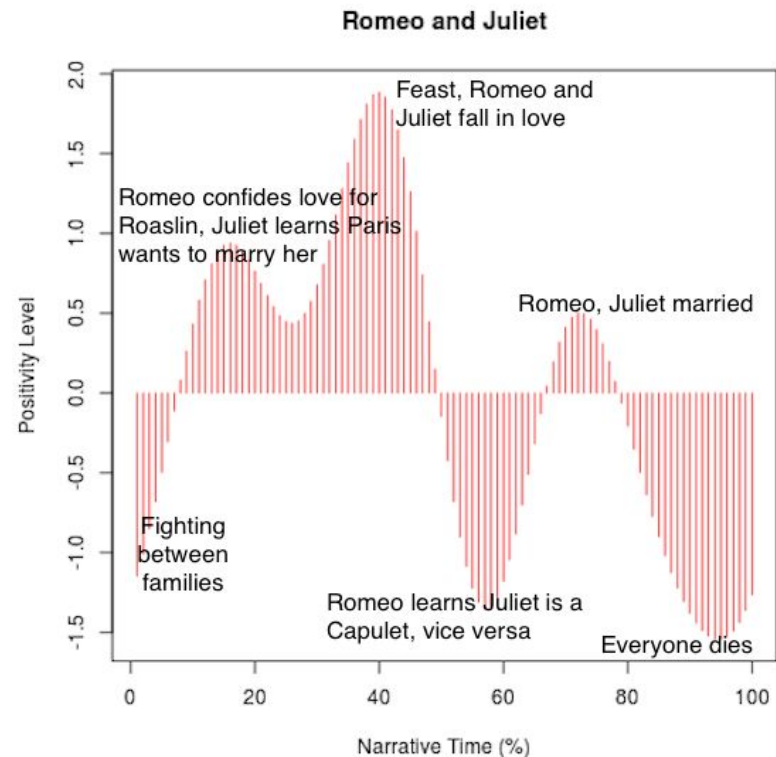
Компьютерные инструменты: МП; извлекать значимую информацию, разрешать неоднозначность (замок, стекло, кореференция «Отец Егора, очень важная фигура в жизни Егора, он его бесконечно любит», когда я первый раз читал эту фразу, не понимал, кто «он» именно, Егор или его отец?

В новостных текстах: упоминания разных лиц, мест или фактов, организаций. Коммерческая организация просит компьютерно-лингвистические компании мониторить Сеть и просто анализировать блоги, записи в социальных сетях, чтобы те, кто осуществляет для них эту деятельность, осуществлять поиск сообщений о настроениях людей, которые пишут в Сети, их отношении к продукции: за что его ругают, за что его хвалят. Это невозможно сделать вручную: записей очень много, пользователей в Сети очень много.

# «Анализ тональности»

определение эмоционального «фона» текста: насколько он положительный или отрицательный. Если упростить, то это способность компьютера определить, говорят сейчас о чем-то хорошем или о чем-то плохом.

применение этого инструмента к тексту «Ромео и Джульетты» дает диаграмму, описывающую сюжет трагедии в терминах эмоций, той самой «тональности»:





# Достижения

- Компьютерно-лингвистические продукты, особенно в интернете, находятся везде: пользовательские запросы, интерфейс, факты, которые встречаются нам в браузере, в интернете постоянно.
- обрабатывать большие объемы текстов для распознавания значений. И это не с помощью правил и словарей, а исходя из данных, которые он получит на вход, что два каких-то слова похожи между собой, что они являются синонимами (как говорят, квазисинонимами). Это то, что называется дистрибутивной семантикой или векторными моделями в лингвистике, в построении семантики.
- Можно ничего специально с помощью правил или словарей компьютеру не объяснять и получить списки синонимов, которые являются уже большим шагом к пониманию смысла текста, к пониманию того, с чем имеет дело компьютер.
- Мы не можем объяснить компьютеру все слова, описать все правила в языке, но используя компьютерные технологии, можно получать такие классификации слов и конструкций, которые продемонстрируют общность их значений.

# Ашманов Игорь Александрович

В Общественной палате РФ

- [https://www.youtube.com/watch?v=\\_9D739byOEs](https://www.youtube.com/watch?v=_9D739byOEs)

Аналитика Big Data

- <https://www.youtube.com/watch?v=ko1lkrwvEZg>

# Чего компьютерная лингвистика не умеет?

- Распознавать иронию; оценку того, сказано о чем-то хорошо или плохо, в каком контексте упоминаются названия организаций и т.п.
- «Ну хорош человек, хорош!», и для компьютера формально здесь будет все вполне себе пристойно, а на самом деле выражено недовольство.
- Описывать ситуацию внешнего мира...

# Что читать? Что слушать?

- Компьютерная лингвистика - Борис Орехов 2016 г.

[https://www.youtube.com/watch?v=74RUVKb\\_KfQ](https://www.youtube.com/watch?v=74RUVKb_KfQ)

- Вопрос науки. Почему компьютерная лингвистика развивается медленно

<https://www.youtube.com/watch?v=hSY95YH5Uts>

# Тема 3: **Квантитативная** **ЛИНГВИСТИКА**

Компьютерная лексикография:

- программы лексикографических работ;
- автоматические словари.

Информационные технологии в словарной работе: словари и энциклопедии; лингвострановедческие словари; терминологические словари; толковые словари (дескриптивные и нормативные, общие и частные); идеографические словари; дву- и многоязычные (переводные) словари; ассоциативные словари; частотные словари; исторические и этимологические словари; орфографические и орфоэпические словари; словообразовательные и грамматические словари; словари трудностей; сочетаемостные словари; ономастиконы; нетрадиционные типы словарей.

# Продукты

Разработка универсальных компьютерных лингвистических моделей (морфология Разработка универсальных компьютерных лингвистических моделей (морфология, синтаксис Разработка универсальных компьютерных лингвистических моделей (морфология, синтаксис, семантика Разработка универсальных компьютерных лингвистических моделей (морфология, синтаксис, семантика) для задач автоматической обработки текстов для разных языков. Эти модели использовались в составе разных программных продуктов и проектов: спелчекеров Разработка универсальных компьютерных лингвистических моделей (морфология, синтаксис, семантика) для задач автоматической обработки текстов для разных языков. Эти модели использовались в составе разных программных продуктов и проектов: спелчекеров (Litera, 1989, Институт Новых Технологий; LingvoCorrector, 1993, BIT Software), всех версий OCR Разработка универсальных компьютерных лингвистических моделей (морфология, синтаксис, семантика) для задач автоматической обработки текстов для разных языков. Эти модели использовались в составе разных программных продуктов и проектов: спелчекеров (Litera, 1989, Институт Новых Технологий; LingvoCorrector, 1993, BIT Software), всех версий OCR-системы ABBYY FineReader Разработка универсальных компьютерных лингвистических моделей (морфология, синтаксис, семантика) для задач автоматической обработки текстов для разных языков. Эти модели использовались в составе разных программных продуктов и проектов: спелчекеров (Litera, 1989, Институт Новых Технологий; LingvoCorrector, 1993, BIT Software), всех версий OCR-системы ABBYY FineReader, электронных словарей ABBYY Lingvo Разработка универсальных компьютерных лингвистических моделей (морфология, синтаксис, семантика) для задач автоматической обработки текстов для разных языков. Эти модели использовались в составе разных программных продуктов и проектов:

# Шайкевич Анатолий Янович

- советский и российский лингвист, доктор филологических наук советский и российский лингвист, доктор филологических наук, профессор. Член Научного совета по лексикологии и лексикографии РАН.
- В настоящее время формирование машинного корпуса современных русских текстов, статистическое исследование русской поэзии и прозы. Ряд исследований о статистике распределения личных имён В настоящее время формирование машинного корпуса современных русских текстов, статистическое исследование русской поэзии и прозы. Ряд исследований о статистике распределения личных имён, о статистике совместного цитирования В настоящее время



## Тема 4. Корпусная лингвистика

- Теория и практика создания корпусов национальных языков; задачи, решаемые с использованием корпусов текстов;
- Исходные понятия корпусной лингвистики: проблемная область; корпус данных; единица хранения корпуса данных; корпус текстов; исследовательские корпуса; иллюстративные корпуса; динамические и статические корпуса текстов; корпуса параллельных корпусов; способ представления и хранения корпуса данных; порог отображения; параметризация проблемной области.
- Требования к корпусу текстов с позиций пользователя: репрезентативность; полнота; экономичность; структуризация материала; компьютерная поддержка. Варианты корпусов текстов. Программное обеспечение корпуса; формирование корпуса текстов.



# Разноязычные корпуса

- [Национальный корпус русского языка](#)
- [Хельсинкский аннотированный корпус русских текстов ХАНКО](#)
- [Компьютерный корпус текстов русских газет конца XX века](#)
- [Национальный корпус русского литературного языка](#)
- [Регенсбургский диахронический корпус русского языка \(древнерусские тексты\)](#)
- [Параллельный корпус переводов «Слова о полку Игореве»](#)
- [Рукописные памятники Древней Руси: берестяные грамоты, летописи, рукописная книга](#)
- [Машинный фонд русского языка](#)
- [Тюбингенский и Упсальский корпуса](#)
- [Corpus of Contemporary American English \(COCA\)](#)
- [Corpus of Historical American English \(COHA\)](#)
- [TIME Magazine Corpus of American English](#)
- [BYU-BNC: British National Corpus](#)
- [Корпуса английского языка, доступные с сайта университета Лидс](#)
- [Британский национальный корпус](#)
- [Bank of English](#)
- [National American Corpus](#)
- [Международный корпус английского языка \(International Corpus of English, ICE\)](#)
- [Синтаксически размеченный корпус – банк деревьев \(Penn TreeBank\)](#)
- [Brown Corpus](#)
- [Саарбрюкенский корпус разговорного английского языка \(The Saarbrücken Corpus of Spoken English \(SCoSE\)\)](#)
- [Исторический корпус испанского языка CORPUS DEL ESPACOL](#)
- [Корпус испанского языка Corpus de Referencia del Español Actual \(CREA\)](#)
- [Диахронический корпус испанского языка Corpus Diacronic del Espacol \(CORDE\)](#)
- [ARTFL-FRANTEXT](#)
- [Lexicum](#)
- [Корпуса французского языка, доступные в проекте VISL](#)

# Что слушать?

**ПЛУНГЯН ВЛАДИМИР** Александрович, д.ф.н., Институт русского языка им. В.В. Виноградова РАН

- Корпусная лингвистика <https://www.youtube.com/watch?v=ico0zGpiBzc>
- Корпусная лингвистика и корпус РЯ

**СЕЛЕГЕЙ** Владимир Павлович

- Корпусные исследования языка. Малый ШАД  
Дмитрий **ДОБРОВОЛЬСКИЙ** (Постнаука)

- Параллельные корпуса текстов. 24.11.2015.  
<https://www.youtube.com/watch?v=ETDj85VpnEc>

**РАХИЛИНА ЕКАТЕРИНА ВЛАДИМИРОВНА**

- Национальный корпус РЯ, на полит.ру 2012 г.  
<https://www.youtube.com/watch?v=M0YboCvXTyA>

**ТИМОШЕНКО** Светлана

- Компьютерная лингвистика № 5. Корпус текстов (Чердак).  
15.07.2016. <https://www.youtube.com/watch?v=Hvrcl4WtMUQ>

# Знакомство с корпусами

- НКРЯ
- ГИКРЯ (Генеральный Интернет-корпус русского языка)  
<http://www.webcorpora.ru/>
- <https://www.english-corpora.org/bnc/>

# Тема 5: Машинный перевод

С конца 50-х годов стали готовить специалистов по машинному переводу и по структурной, прикладной и математической лингвистике.

Зачем это было нужно?

Что из этого получилось?

Проблемы решались лингвистически, но широких экспериментов не было.

Поскольку требовали: представления данных на машинном языке и быстроедействие. Таких машин не было.

Что требует компьютер?  
работать только с теми текстами, которые существуют в электронном виде.

Большие объемы электронных текстов

# МП: начало

Джоржтаунский эксперимент

7 января 1954 г.

Компания IBM

С РЯ на АЯ

IBM-701 поле для гольфа,  
ввод через перфокарты,  
рассыпались, мялись,  
выход на экране компа  
подготовка несколько  
месяцев, словарь 250  
русских слов, пословный  
перевод

Тв.п. = by

Событие освещалось во всех  
американских СМИ

- 1955 г. Ляпунов и Кулагина в журнале «Природа», статья
- Кулагина: МП с ФЯ на РЯ, Мельчук И.А., исп. отд. МГУ
- Институт прикладной математики АН СССР

# МП: состояние работ в 1958-1985 гг.

МП как в основном научная задача.

Убежденность в том, что машина не заменит переводчика.

МП без машины: алгоритмы разбора текстов;

- Система ФРАП: Н.Н. Леонтьева, ВЦП, Москва
- Система АМПАР: Марчук Ю.Н.
- Система НЕРПА

## электронные переводчики

Работают на методе анализа больших данных

Понимают ли текст? Вообще говоря, нет.

# Информационные ресурсы (лекции и интервью ученых)

- Баранов А.Н. Прикладная лингвистика. Постнаука. 2015 г. <https://www.youtube.com/watch?v=3BzrupwFSu8>
- Плунгян В.А. Корпусная лингвистика <https://www.youtube.com/watch?v=ico0zGpiBzc>
- Плунгян В.А. Корпусная лингвистика и корпус русского языка
- Селегей В.П. Компьютерная лингвистика: трудности перевода. <https://www.youtube.com/watch?v=hSY95YH5Uts> (передача «Вопрос науки» на Россия 24).
- Селегей В.П. Цикл лекций РГГУ «Компьютерная лингвистика» "Компьютерная лингвистика сегодня: от автоматической обработки текста до машинного понимания" [youtube.com](https://www.youtube.com/watch?v=b2pq0HJ-ABq) Селегей В.П. Цикл лекций РГГУ «Компьютерная лингвистика» "Компьютерная лингвистика сегодня: от автоматической обработки текста до машинного понимания" [youtube.com](https://www.youtube.com/watch?v=b2pq0HJ-ABq)
- Селегей В.П. Корпусные исследования языка. Малый ШАД
- Селегей В.П. Теория и практика компьютерной лингвистики (Чердак). 09.08.2016. <https://www.youtube.com/watch?v=zwqbbq0pD8A>
- Иомдин Л. Компьютерная лингвистика №3. Синтаксический анализ (Чердак). 01.07.2016 <https://www.youtube.com/watch?v=ZFB0T01bT-g>



# продолжение

- Тимошенко С. Компьютерная лингвистика №1. Зачем лингвисту компьютер (Чердак). 24.06.2016. <https://www.youtube.com/watch?v=kxSH7UTheLE>
- Тимошенко С. Компьютерная лингвистика № 2. Морфологический анализ (Чердак) . 27.06.2016. <https://www.youtube.com/watch?v=oj0Kjt9gMBc>
- Тимошенко С. Компьютерная лингвистика №4. Семантический анализ (Чердак) 10.07.2016. <https://www.youtube.com/watch?v=tUSEn6fVwSo>
- Компьютерная лингвистика № 5. Корпус текстов (Чердак). 15.07.2016. <https://www.youtube.com/watch?v=Hvrcl4WtMUQ>
- Добровольский Д. (Постнаука). Параллельные корпуса текстов. 24.11.2015. <https://www.youtube.com/watch?v=ETDj85VpnEc>
- Рахилина Е.В. Национальный корпус русского языка (полит.ру), 2012 г. <https://www.youtube.com/watch?v=M0YboCvXTyA>
- Орехов Б.В. (на сайте Постнауки <https://postnauka.ru/author/orehovb>) Орехов Б. В. (на сайте Постнауки <https://postnauka.ru/author/orehovb>). Компьютерная лингвистика [https://www.youtube.com/watch?v=74RUVKb\\_KfQ](https://www.youtube.com/watch?v=74RUVKb_KfQ)
- Орехов Б. Подсчеты легитимизируют гуманитарную наук. О цифровых гуманитарных исследованиях. <https://nplus1.ru/material/2017/12/28/dig-hum> Орехов Б. Подсчеты легитимизируют гуманитарную наук. О цифровых гуманитарных исследованиях. <https://nplus1.ru/material/2017/12/28/dig-hum>, материал от 28.12.2017, дата обращения 17.03.2019. <http://www.translate.ru>

# программное обеспечение

<http://www.lingvo.ru/>

- <http://www.google.ru/dictionary>
- <http://www.translate.ru>
- <http://www.translate.google.ru/>
- <http://babelfish.altavista.com/>

Программные продукты АBBYU Lingvo,  
Promt PM

Спасибо за внимание!

« *Quidquid recipitur ad modum recipientis recipitur* » (Tout ce qui est reçu est reçu en fonction de celui qui reçoit). Фома Аквинский