

Affective Computing

Olga Perepelkina
HSE, 2020

**Тема 4. Проявление эмоций в письменной речи и
определение тональности текстов.**

**Тема 6. Эмоции в устной речи, автоматическое
определение эмоций по голосу.**

План лекции

- Эмоции в речи
- Анализ текстов
- Анализ аудио
- Практика

<http://bit.ly/2GJeKz2>

Эмоции в речи

Эмоции в речи

Текст

Аудио

Эмоции в речи

Текст

Аудио

Наши руууляяят форееевааааа!!!!!!

*Приве-е-ед, Пушистый! Откуда такой ко-
о-шерный аватар надыбал?*

Жжжешшь смертельно!!!

Ржжжунемагу!

Какой он клааасссныыий!

Чууумовоой фильмец вчера видала.

Мазздааай косые нуббы!!!!

Сдохни уууррооооддд!!!!!!!

Такой но-о-осище не кирпича, а
железобетонного блока просит.

Изыыдиин! Тролль поганый!

Тра-а-аблы голимые накрыли нас с головой.

Ааааа... что значат все эти страшные
бууквывы?!!!

- Традиционно изучалось невербальное проявление эмоций (лицевые экспрессии)
- Вклад в исследования языка и эмоций пришел из социальной психологии (семантические дифференциал Осгуда, Osgood et al., 1975)
- Компьютерная лингвистика: определение эмоций – непростая задача, эмоции – не лингвистический конструкт

Семантический дифференциал

- Огуд с помощью метода семантического дифференциала пытался определять эмотивное пространство различными наборами парных слов
- Семантический дифференциал – метод построения семантических пространств
- Координаты объекта в семантическом пространстве – его оценки по ряду биполярных шкал, противоположные полюса которых заданы с помощью вербальных антонимов.
- Шкалы отобраны из множества пробных шкал методами факторного анализа.

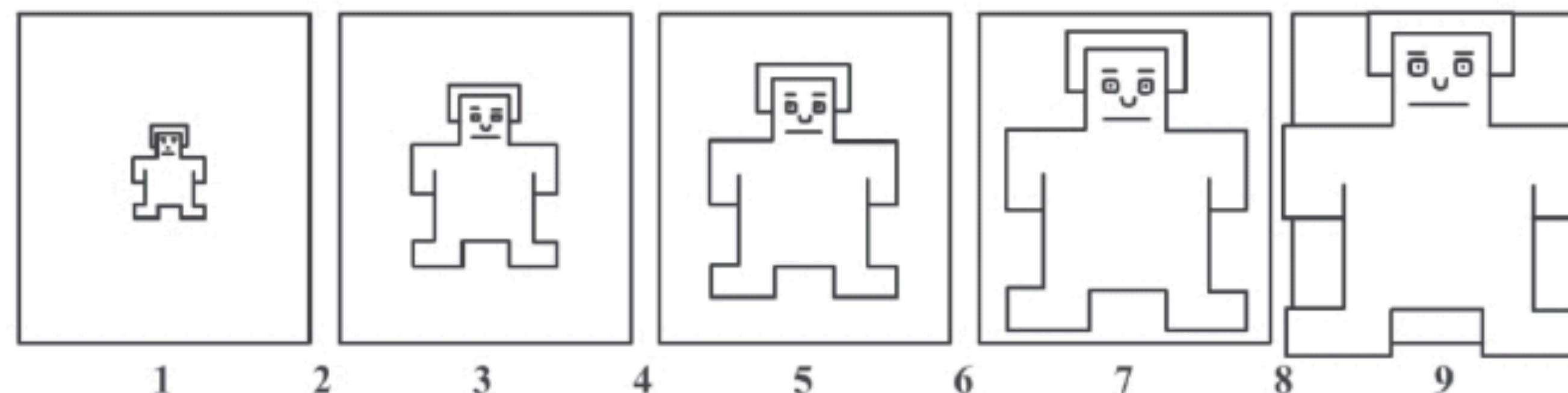
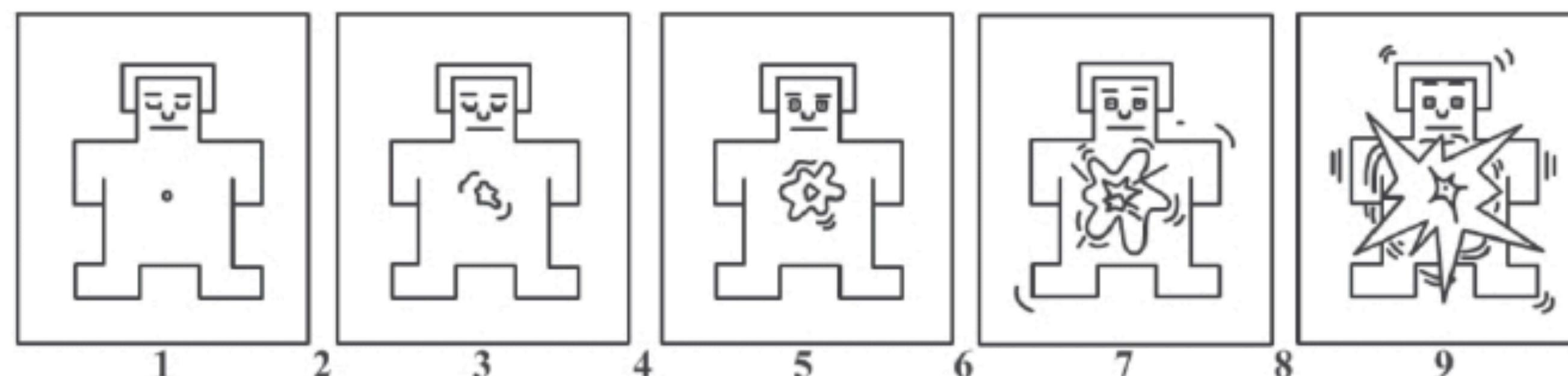
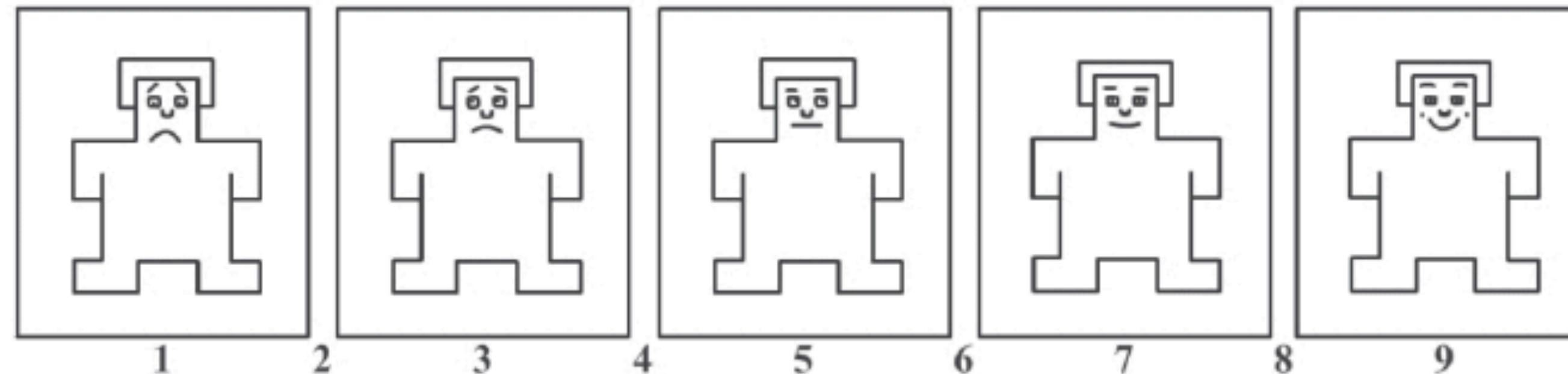
Семантический дифференциал

«оценка»:	<i>хороший</i>	3	2	1	0	-1	-2	-3	<i>плохой</i>
«сила»:	<i>сильный</i>	3	2	1	0	-1	-2	-3	<i>слабый</i>
«активность»:	<i>активный</i>	3	2	1	0	-1	-2	-3	<i>пассивный</i>

ANEW: Affective Norms for English Words

- ANEW – набор нормативных эмоциональных оценок для большого количества слов в английском языке (Bradley & Lang, 1999).
- Набор слов, которые были оценены людьми по шкалам удовольствия (pleasure), возбуждения (arousal) и доминирования (dominance).
- Основан на идее семантического дифференциала.
- Слова оценивались при помощи методики Self-Assessment Manikin

Self-Assessment Manikin

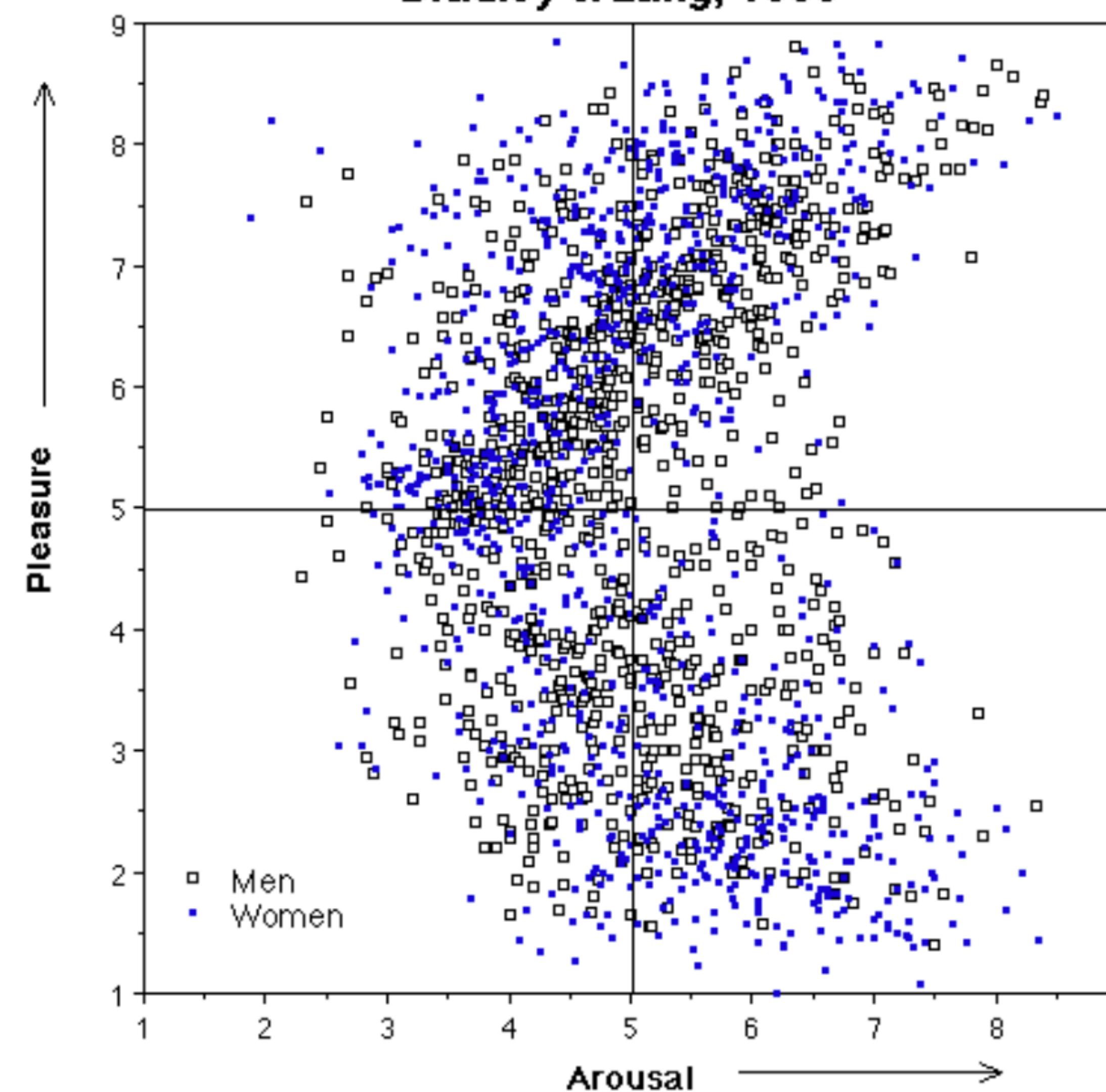


ANEW: Affective Norms for English Words

Word	Valence (Mean)	Valence (SD)	Arousal (Mean)	Arousal (SD)	Dominance (Mean)	Dominance (SD)
abduction	2.76	2.06	5.53	2.43	3.49	2.38
abortion	3.50	2.30	5.39	2.80	4.59	2.54
absurd	4.26	1.82	4.36	2.20	4.73	1.72
abundance	6.59	2.01	5.51	2.63	5.80	2.16
abuse	1.80	1.23	6.83	2.70	3.69	2.94
acceptance	7.98	1.42	5.40	2.70	6.64	1.91
accident	2.05	1.19	6.26	2.87	3.76	2.22
ace	6.88	1.93	5.50	2.66	6.39	2.31
ache	2.46	1.52	5.00	2.45	3.54	1.73
achievement	7.89	1.38	5.53	2.81	6.56	2.35

Affective Norms for English Words (ANEW)

Bradley & Lang, 1999



ENRuN: нормы на русском

№	Слово	Лингвистические характеристики			Валентность					Возбуждение				
		Частотность	Буквы	Слоги	M	SD	Min	Max	N	M	SD	Min	Max	N
1	Авария	29,6	6	4	1,74	1,01	1	7	91	5,59	1,51	1	7	90
2	Адрес	92,7	5	2	3,99	0,39	2	5	98	2,19	1,40	1	5	98
3	Альбом	23,7	6	2	4,51	0,92	1	7	91	2,56	1,69	1	7	91
4	Ангел	32,4	5	2	5,51	1,33	1	7	99	3,78	1,99	1	7	99
5	Анекдот	35	7	3	5,03	1,18	2	7	91	3,80	1,96	1	7	91
6	Арест	46,5	5	2	2,19	1,22	1	7	99	4,68	1,91	1	7	99
7	Атака	42,7	5	3	3,13	1,42	1	7	91	5,11	1,59	1	7	91
8	Бабочка	21,3	7	3	5,09	1,26	1	7	99	3,91	1,72	1	7	99
9	Балкон	28,2	6	2	4,22	0,78	2	7	99	2,31	1,60	1	7	99
10	Бандит	42,8	6	2	2,01	0,89	1	5	99	4,59	1,68	1	7	99
11	Банкротство	20,9	11	3	1,91	0,97	1	4	91	4,63	1,93	1	7	91
12	Башня	39,4	5	2	4,12	0,90	1	7	190	2,48	1,68	1	7	190
13	Беда	7,3	4	2	1,65	0,82	1	6	99	5,38	1,74	1	7	99
14	Безумие	15,1	7	4	3,08	1,59	1	7	91	5,68	1,40	2	7	91
15	Беседа	87,5	6	3	5,13	0,99	4	7	91	3,91	1,68	1	7	91

https://www.researchgate.net/publication/308722711_Baza_dann_yh_ENRuN

ENRuN: нормы на русском

№	Слово	Лингвистические характеристики			Радость					Грусть					Злость					Страх					Отвращение				
		Частотность	Буквы	Слоги	M	SD	Min	Max	N	M	SD	Min	Max	N	M	SD	Min	Max	N	M	SD	Min	Max	N	M	SD	Min	Max	N
1	Авария	29,6	6	4	0,09	0,46	0	3	47	3,36	1,69	0	5	47	1,72	1,65	0	5	47	3,47	1,54	0	5	47	0,89	1,60	0	5	47
2	Адрес	92,7	5	2	0,55	1,28	0	5	53	0,34	0,92	0	5	53	0,08	0,33	0	2	53	0,26	0,81	0	4	53	0,08	0,38	0	2	53
3	Альбом	23,7	6	2	2,57	1,77	0	5	47	1,19	1,50	0	5	47	0,04	0,20	0	1	47	0,15	0,51	0	3	47	0,04	0,20	0	1	47
4	Ангел	32,4	5	2	3,55	1,65	0	5	53	1,04	1,57	0	5	53	0,08	0,43	0	3	53	0,60	1,23	0	5	53	0,04	0,19	0	1	53
5	Анекдот	35	7	3	3,62	1,44	0	5	47	0,19	0,50	0	2	47	0,04	0,20	0	1	47	0,09	0,28	0	1	47	0,30	0,75	0	4	47
6	Арест	46,5	5	2	0,47	0,97	0	4	53	2,21	1,80	0	5	53	2,47	1,87	0	5	53	2,75	1,65	0	5	53	1,45	1,65	0	5	53
7	Атака	42,7	5	3	1,15	1,79	0	5	47	0,89	1,45	0	5	47	2,28	1,90	0	5	47	2,26	1,87	0	5	47	0,57	1,17	0	5	47
8	Бабочка	21,3	7	3	3,26	1,60	0	5	53	0,47	1,10	0	5	53	0,19	0,65	0	3	53	0,42	1,03	0	5	53	0,38	0,81	0	4	53
9	Балкон	28,2	6	2	1,92	2,01	0	5	53	0,66	1,25	0	5	53	0,04	0,19	0	1	53	0,89	1,25	0	5	53	0,08	0,55	0	4	53
10	Бандит	42,8	6	2	0,19	0,76	0	5	53	1,25	1,62	0	5	53	2,77	1,90	0	5	53	2,83	1,49	0	5	53	2,04	1,88	0	5	53
11	Банкротство	20,9	11	3	0,09	0,35	0	2	47	3,23	1,75	0	5	47	2,53	1,95	0	5	47	3,04	1,91	0	5	47	0,74	1,37	0	5	47
12	Башня	39,4	5	2	1,53	1,75	0	5	100	0,82	1,35	0	5	100	0,12	0,48	0	3	100	1,06	1,48	0	5	100	0,15	0,63	0	4	100
13	Беда	7,3	4	2	0,04	0,19	0	1	53	3,36	1,61	0	5	53	1,91	1,71	0	5	53	2,75	1,65	0	5	53	1,57	1,60	0	5	53
14	Безумие	15,1	7	4	1,43	1,80	0	5	47	1,26	1,52	0	5	47	1,32	1,63	0	5	47	2,38	1,94	0	5	47	1,15	1,67	0	5	47
15	Беседа	87,5	6	3	3,26	1,51	0	5	47	0,40	0,83	0	4	47	0,23	0,89	0	5	47	0,51	1,12	0	5	47	0,17	0,84	0	5	47
16	Беспокойство	18,5	12	4	0,23	0,67	0	3	53	2,11	1,54	0	5	53	1,19	1,48	0	5	53	2,25	1,48	0	5	53	1,06	1,42	0	5	53
17	Беспомощность	6	13	4	0,13	0,74	0	5	47	3,21	1,59	0	5	47	2,64	1,85	0	5	47	2,91	1,89	0	5	47	1,74	1,84	0	5	47
18	Беспощадность	1,4	13	4	0,50	1,24	0	5	52	1,83	1,88	0	5	52	2,50	2,13	0	5	52	2,52	1,73	0	5	52	2,08	1,92	0	5	52
19	Бессилие	8,1	8	4	0,09	0,45	0	3	53	2,77	1,56	0	5	53	2,30	1,98	0	5	53	2,53	1,74	0	5	53	1,74	1,84	0	5	53

https://www.researchgate.net/publication/308722711_Baza_dann_yh_ENRuN

Эмоции в тексте

- Эмоциональная оценка, выраженная в тексте, называется **тональностью**, или **сентиментом** текста (от англ. sentiment – чувство; мнение, настроение).
- Обычно одна шкала (positive/negative), но есть попытки и с большим количеством шкал.
- Автоматическое определение тональности текста: выделение тех фрагментов текста, которые выражают позитивную или негативную эмоциональность по отношению к объекту эмоциональной оценки (объекту тональности).
- Объект тональности – человек, продукт, организация, услуга и тд

Анализ текстов

Natural Language Processing, NLP

- Обработка естественного языка – общее направление искусственного интеллекта и математической лингвистики.
- Оно изучает проблемы компьютерного анализа и синтеза естественных языков.
- Применительно к искусственному интеллекту анализ означает **понимание языка**, а синтез – **генерацию грамотного текста**.

Natural Language Processing



Задачи NLP

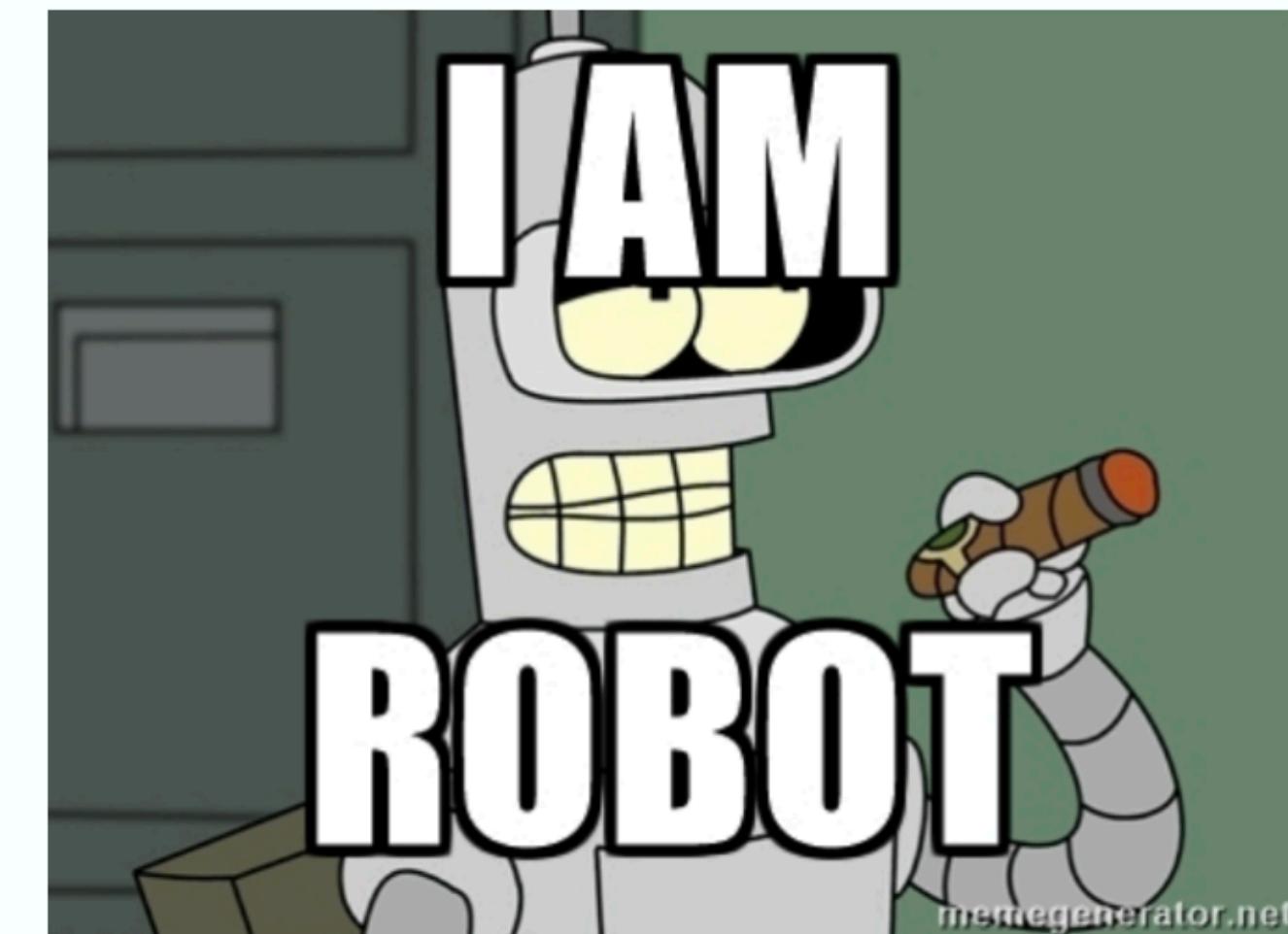
- Классификация текстов (темы книг, статей, позитивный/негативный)
- Регрессия на текстах: рейтинг статьи, количество лайков и просмотров
- Кластеризация текстов, выделение тематик в текстах
- Извлечение информации (фактов и событий)

Задачи NLP

- Поиск слов, похожих по смыслу на данное
- Диалоговые системы, вопрос-ответные системы
- Аннотация текста
- Автопереводы
- Генерация текстов

Основная идея

- Компьютер не понимает голый текст, он понимает только цифры.
- Нужно превратить текст в цифры!
- Работаем с документами. Документ состоит из токенов.

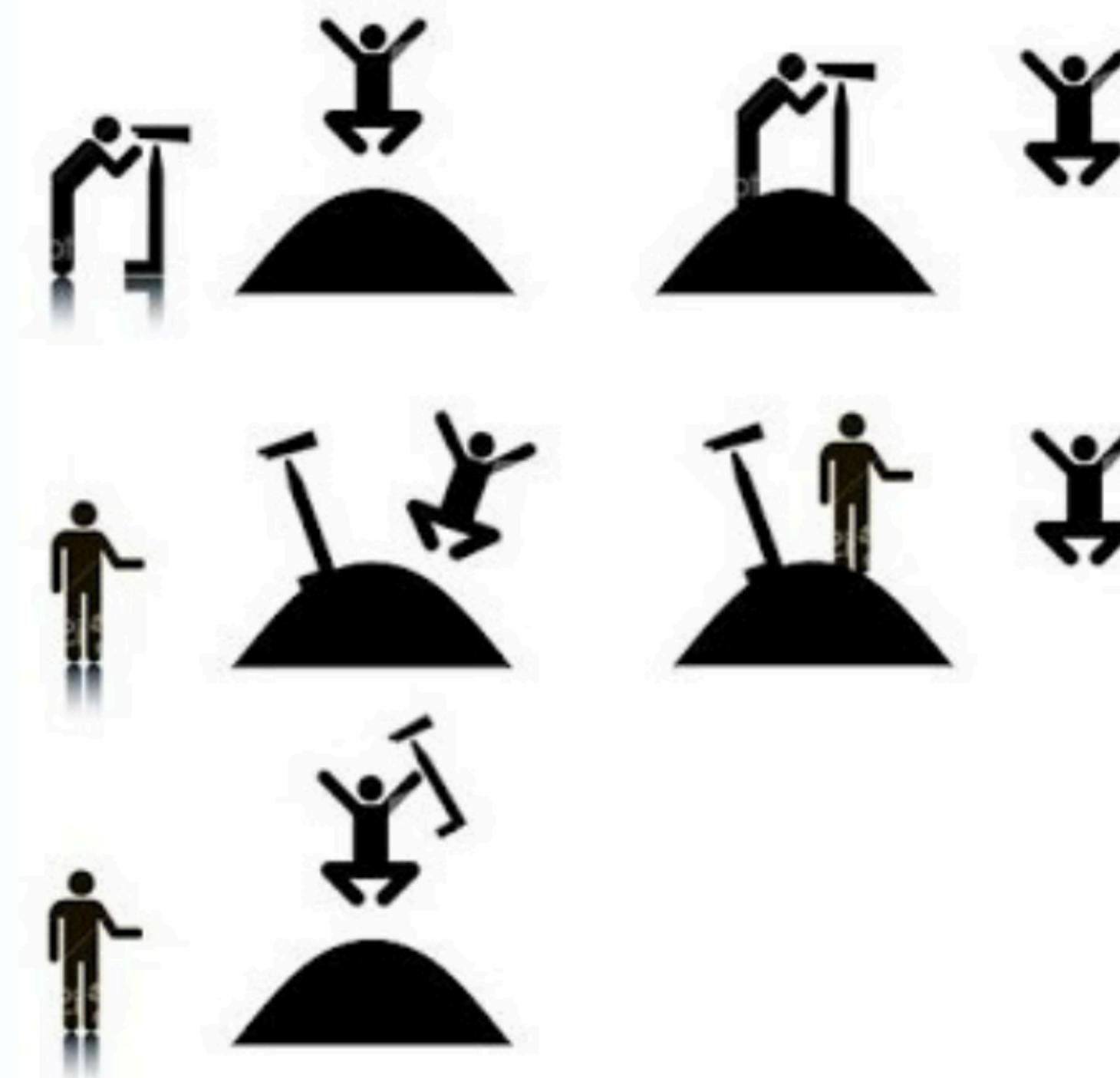


Основные проблемы

- Лексическая неоднозначность: орган, парить, рожки.
- Морфологическая неоднозначность: «хранить деньги в банке», «что делают белки в клетке».
- Синтаксическая неоднозначность: «Эти типы стали есть на складе».
- Неологизмы: печеньки, репостнуть, расшарить.
- Разное написание: Россия, РФ, Российская Федерация.
- Нестандартное написание, орографические ошибки и очепятки.

Синтаксическая неоднозначность

I saw a man on the hill with a telescope



Анализ текста

- Что является единицей анализа (юнитом)?
 - Весь документ
 - Предложение

Простой подход

- Документ – это множество слов
- Порядок слов неважен
- Порядок неважен слов
- Неважен слов порядок
- Слов порядок неважен
- Неважен порядок слов
- Слов неважен порядок
- Мешок слов (*bag of words*)



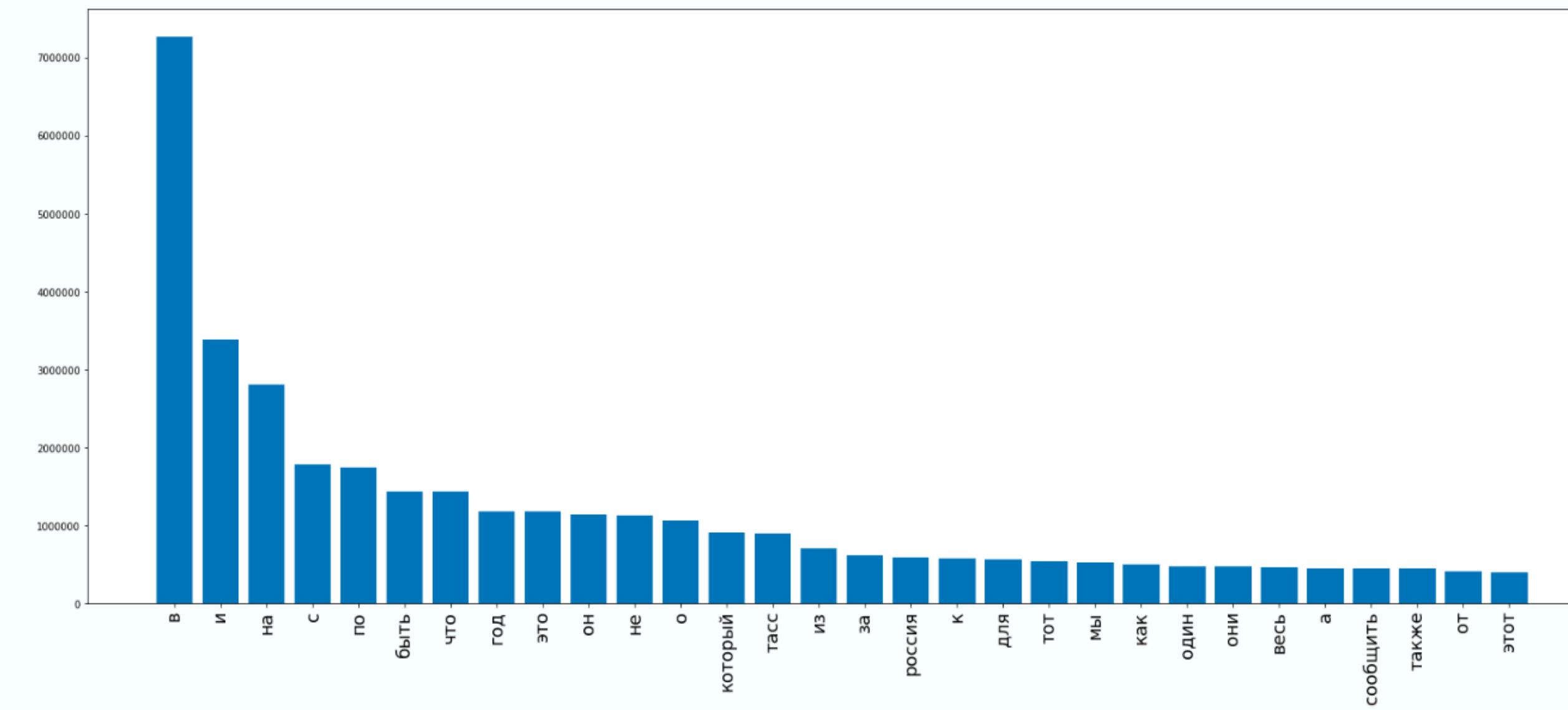
Bag-of-words

- Посчитаем частоту каждого слова в каждой статье!
- Готово, можно строить модели!

	я	просмотреть	выборы	вместе
Статья 1	4	1	1	1
Статья 2	6	2	0	0
Статья 3	7	0	1	0
....
Статья 150000	2	0	0	0

Стоп-слова

- На 40 000 наблюдений есть 2 058 294 уникальных слова в текстах



- Слова а, но, в, за можно найти в каждом тексте
- Выбрасываем стоп-слова

Нормализация

красивый, красива, красивые ...

банк, банков, банками ...

человек, люди ...

Стеминг

Обрезаем приставки
и окончания

Лемматизация

Меняем словоформу
по словарю



Классификация эмоций

- Частоты слов – в классификатор
- Проблема – даже с ухищрениями работает не очень...
- Что делать? Давайте использовать контекст!

Word embeddings

- Техника, возникшая при обработке естественного языка
- Представляет значение слова как вектор
- Использует интуитивную идею контекста вокруг слов

Word2Vec

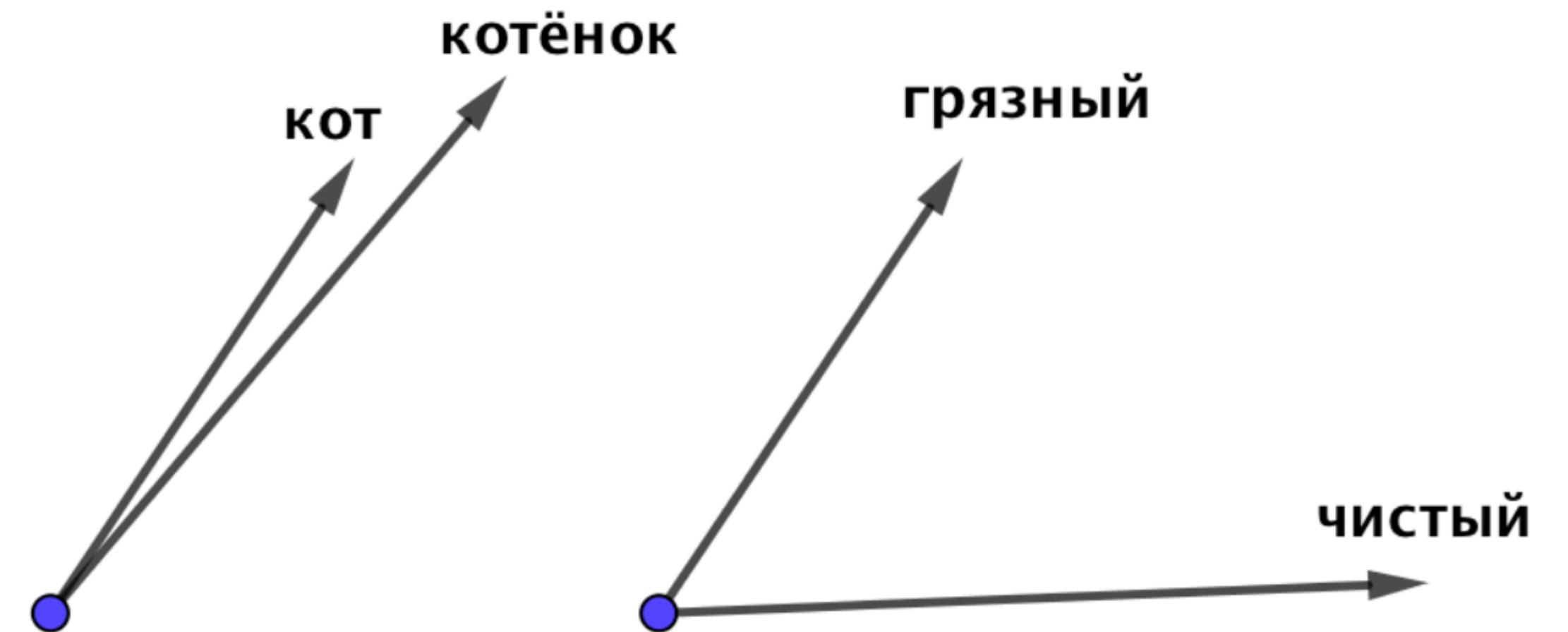
- Итак, хотим понимать смысл слов
- Опишем каждое слово вектором так, чтобы похожие слова были близкими векторами
- Как? Схожие слова имеют схожие контексты
- Давайте оценивать вероятность встретить слово i в контексте слова j

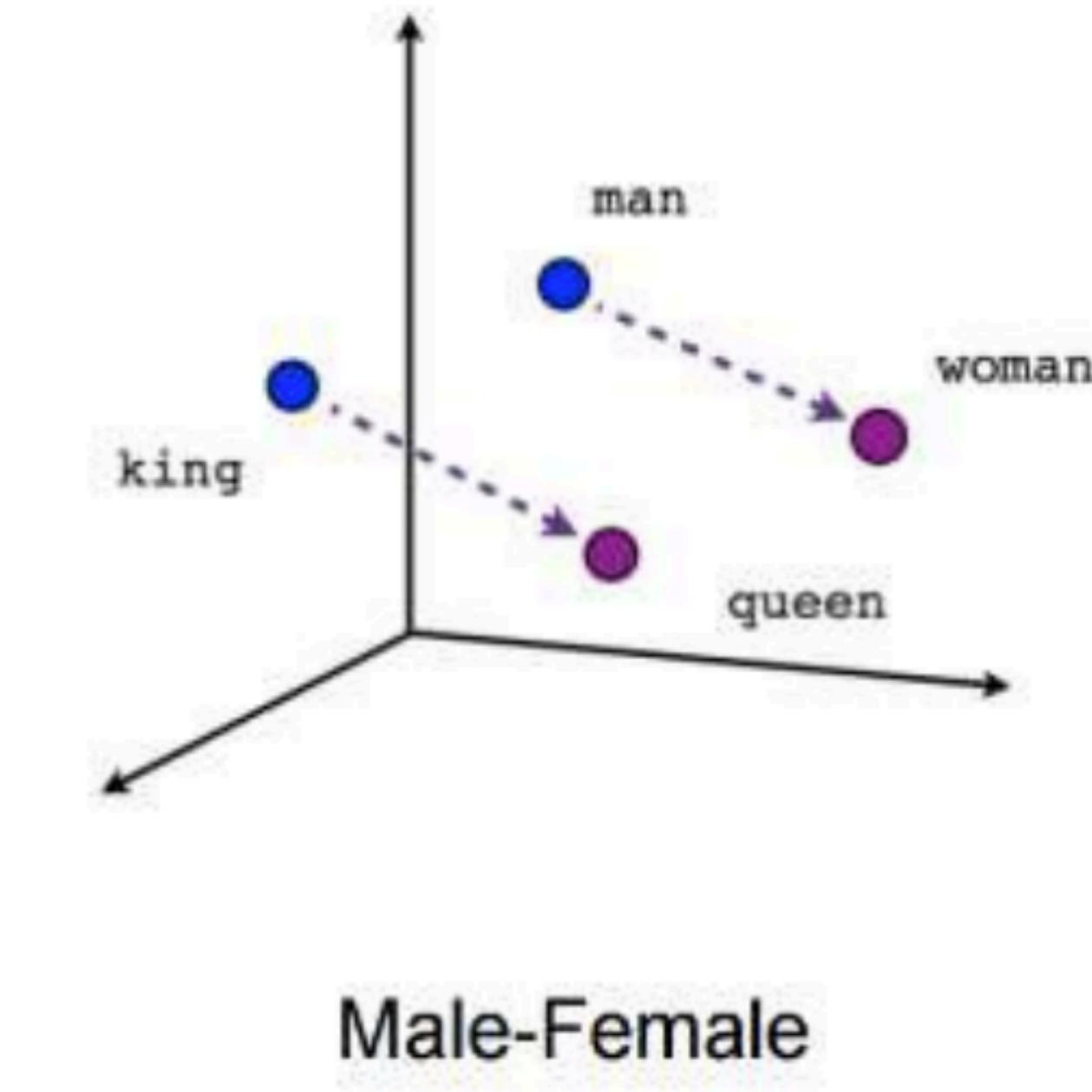
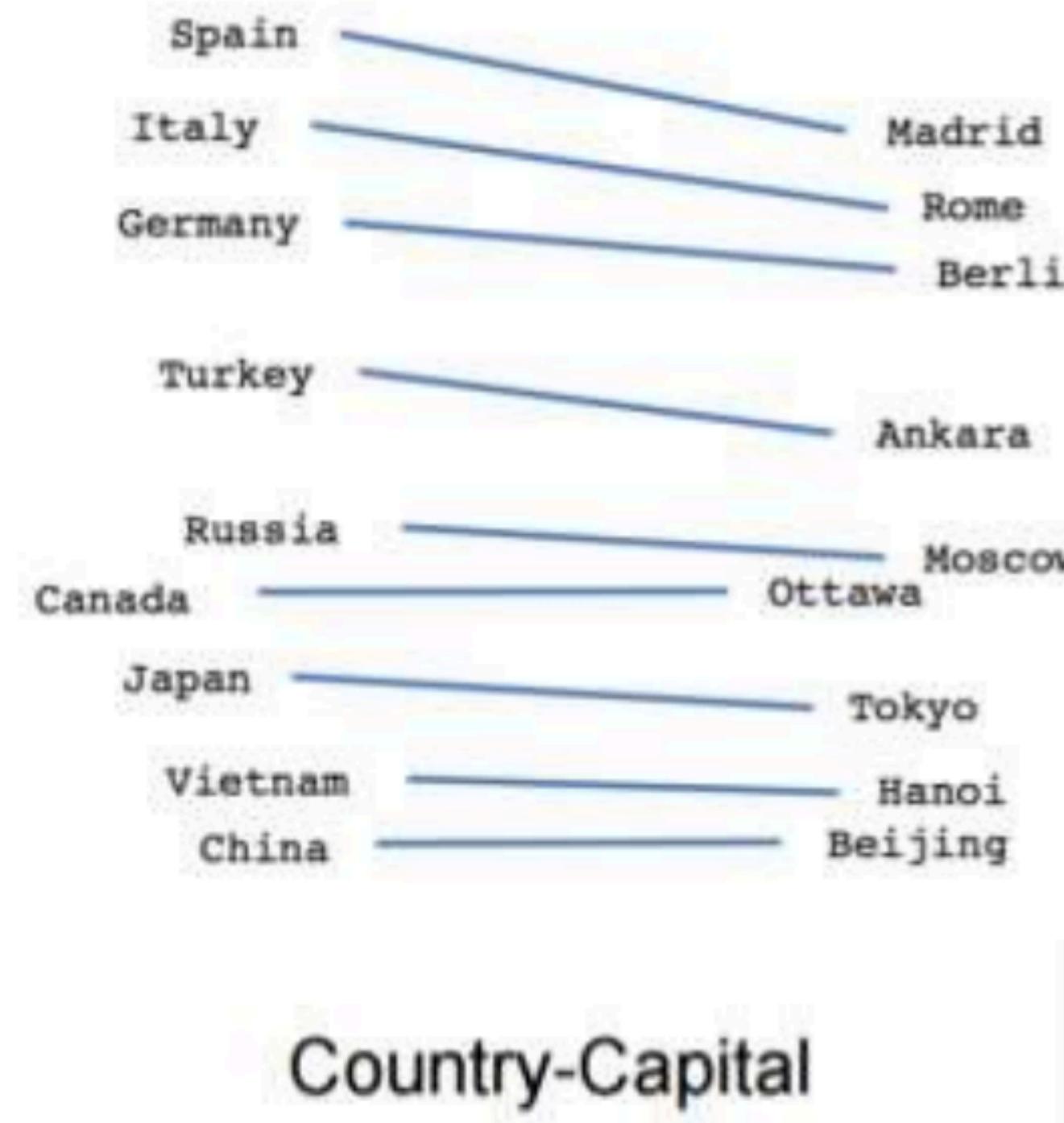
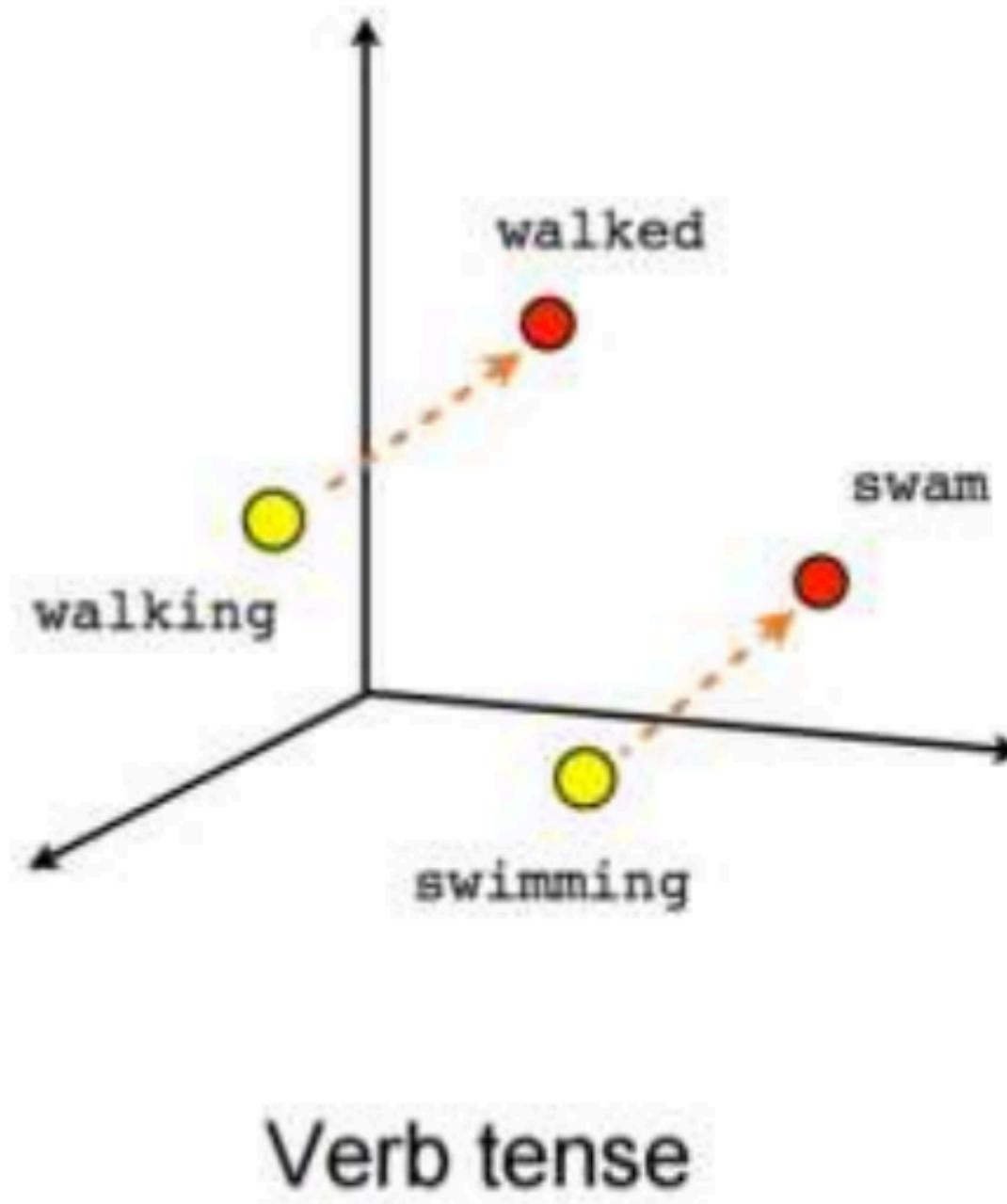
Word2Vec

- “The **cat** sat on the floor” }
- “The **dog** sat on the floor” }
- “The **person** sat on the floor” }
- “The **refrigerator** sat on the floor” }

Word2Vec

- Обучаем модель на корпусе текстов естественной речи (Google news, Tweets...)
- Обучаем ее так, чтобы похожие по смыслу слова лежали рядом (косинусное расстояние между векторами), а непохожие – далеко





$$\text{vector[King]} - \text{vector[Man]} + \text{vector[Woman]} = \text{vector[Queen]}$$

- Can give us insight into how ideas and language encode bias

[King] - [Man] + [Woman] = [Queen]

[Programmer] - [Man] + [Woman] = [Homemaker]

Extreme *she* occupations

- | | | |
|-----------------|-----------------------|------------------------|
| 1. homemaker | 2. nurse | 3. receptionist |
| 4. librarian | 5. socialite | 6. hairdresser |
| 7. nanny | 8. bookkeeper | 9. stylist |
| 10. housekeeper | 11. interior designer | 12. guidance counselor |

Extreme *he* occupations

- | | | |
|----------------|-------------------|----------------|
| 1. maestro | 2. skipper | 3. protege |
| 4. philosopher | 5. captain | 6. architect |
| 7. financier | 8. warrior | 9. broadcaster |
| 10. magician | 11. fighter pilot | 12. boss |

Игра в слова

<http://rusvectores.org/ru>

[https://www.ibm.com/cloud/watson-tone-
analyzer](https://www.ibm.com/cloud/watson-tone-analyzer)

Классификация эмоций

- Далее, когда текст преобразован в числа (частоты, векторы), можем подавать его на вход в классификатор
- Разметка текста – positive/negative, реже - базовые эмоции
- Challenges: юмор, сарказм, культурные различия

Эмоции в речи

Текст

Аудио

1



2



3



4



5



6



Surprised



Happy



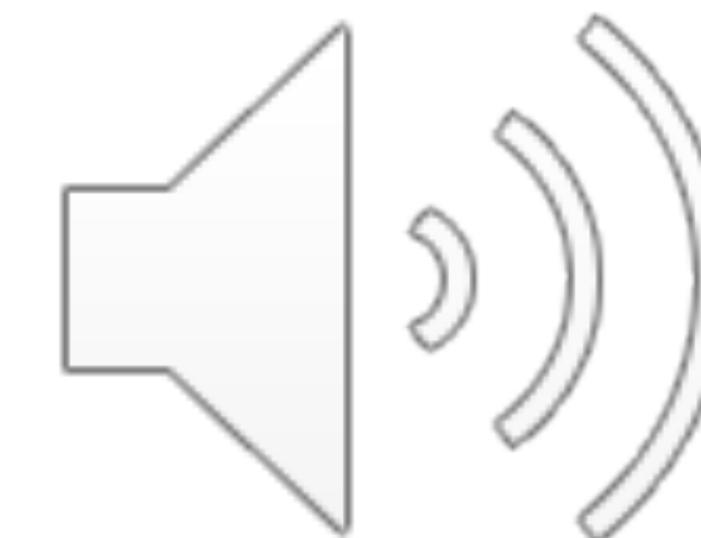
Sad



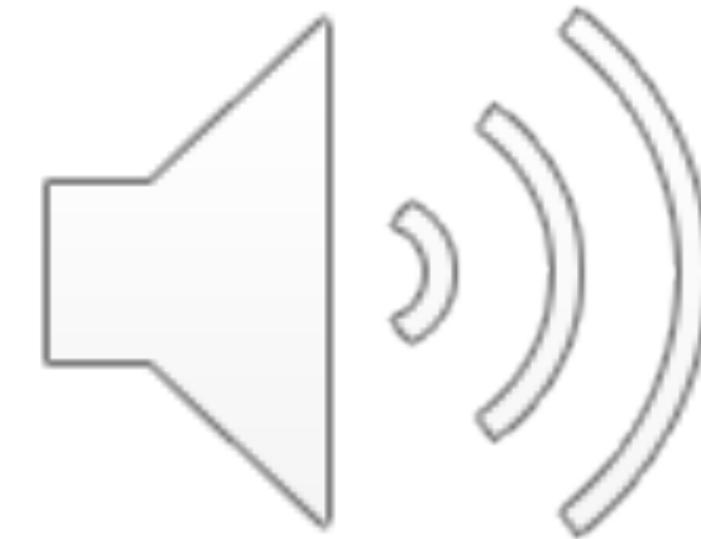
Angry



Fearful



Disgust



Речь

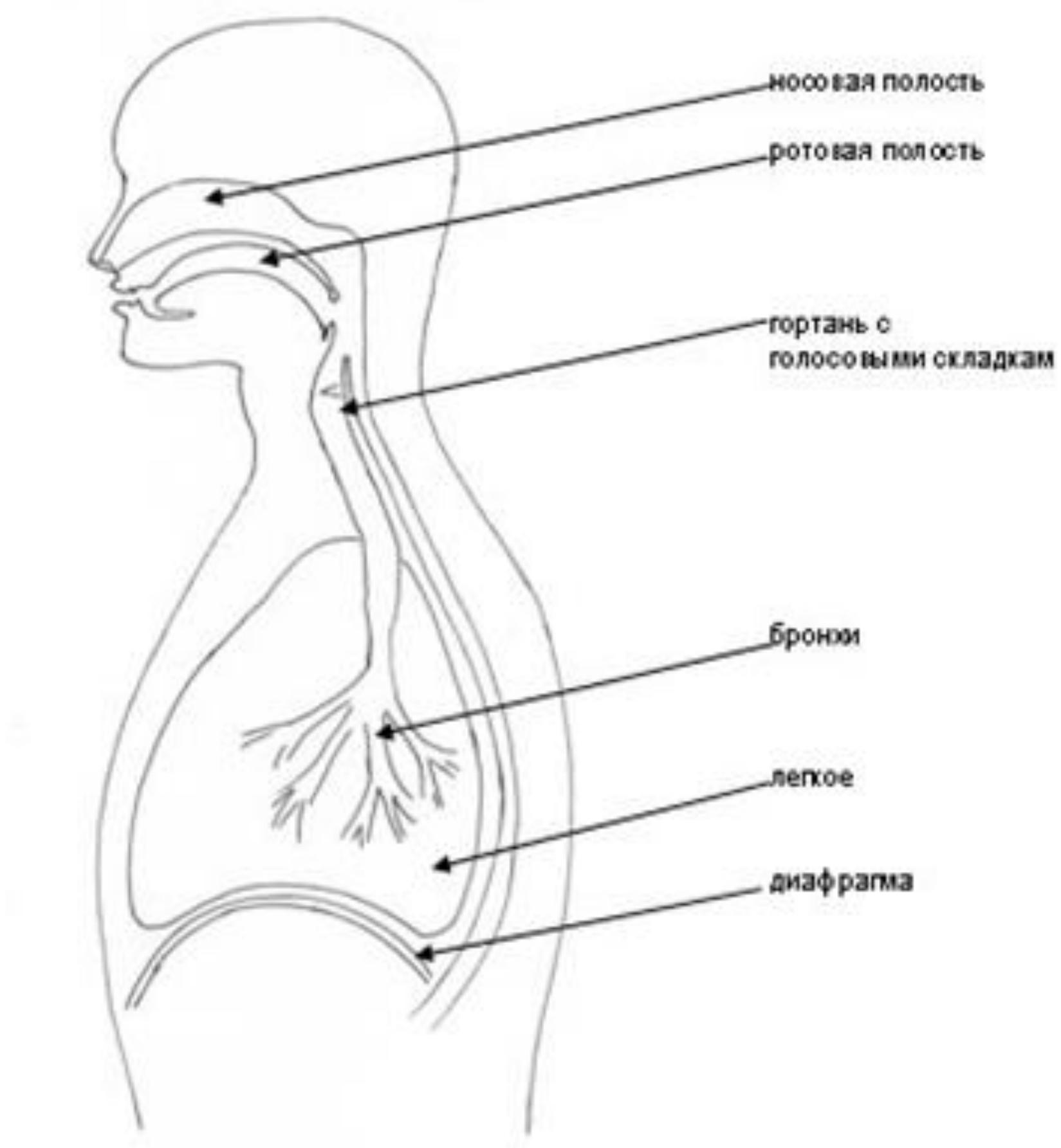
- Лингвистический компонент (текст)
- Паралингвистический компонент (характеристики возраста, пола, эмоциональной окраски и т.д.)

Речевой сигнал

- Речевой сигнал подвержен влиянию многочисленных факторов, включая:
 - Что было сказано (лингвистический компонент)
 - Кем было сказано (спикер, возраст, пол)
 - Как сигнал был записан (телефон, мобильный, тип микрофона)
 - Контекст (акустика комнаты, окружающие шумы)

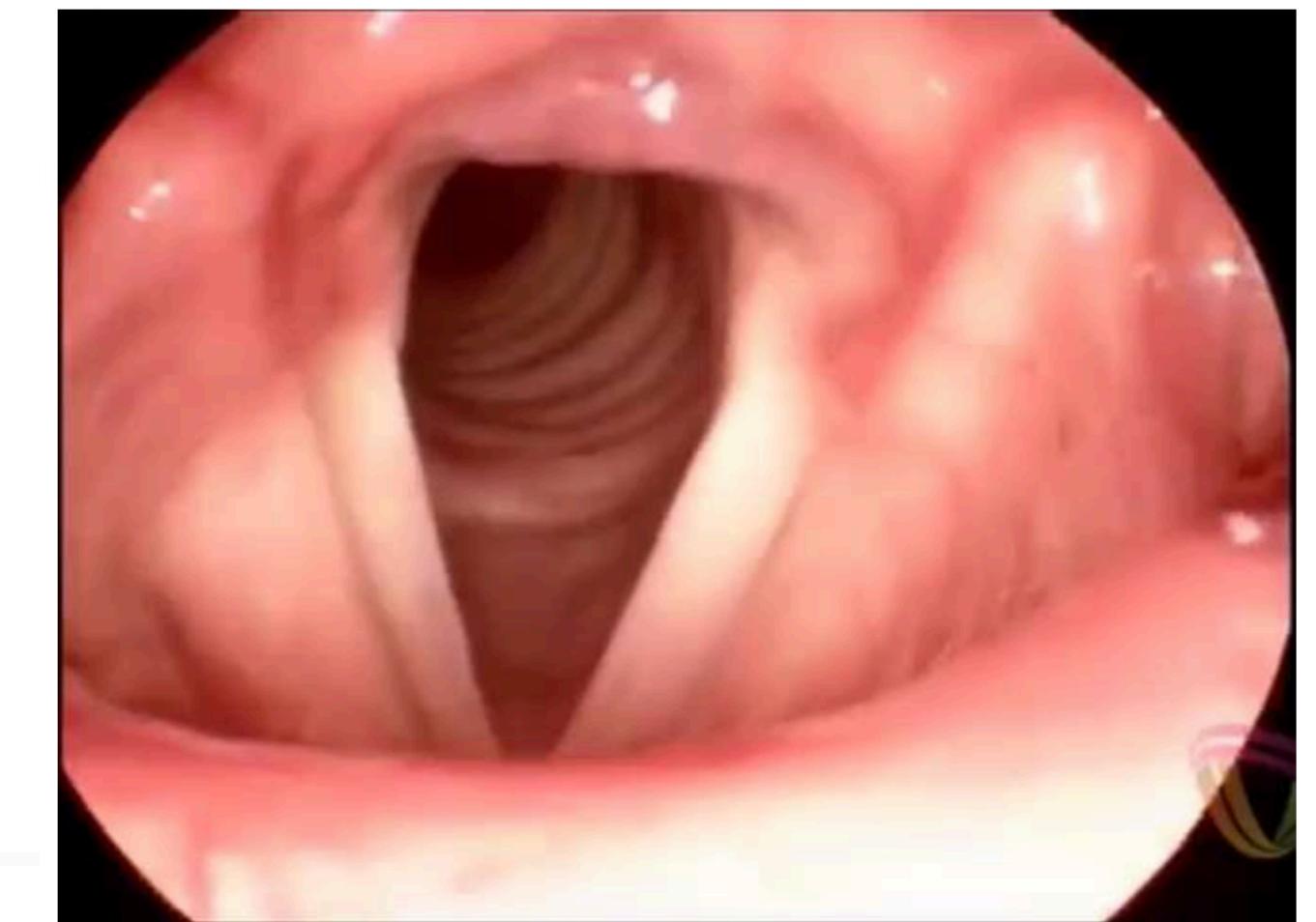
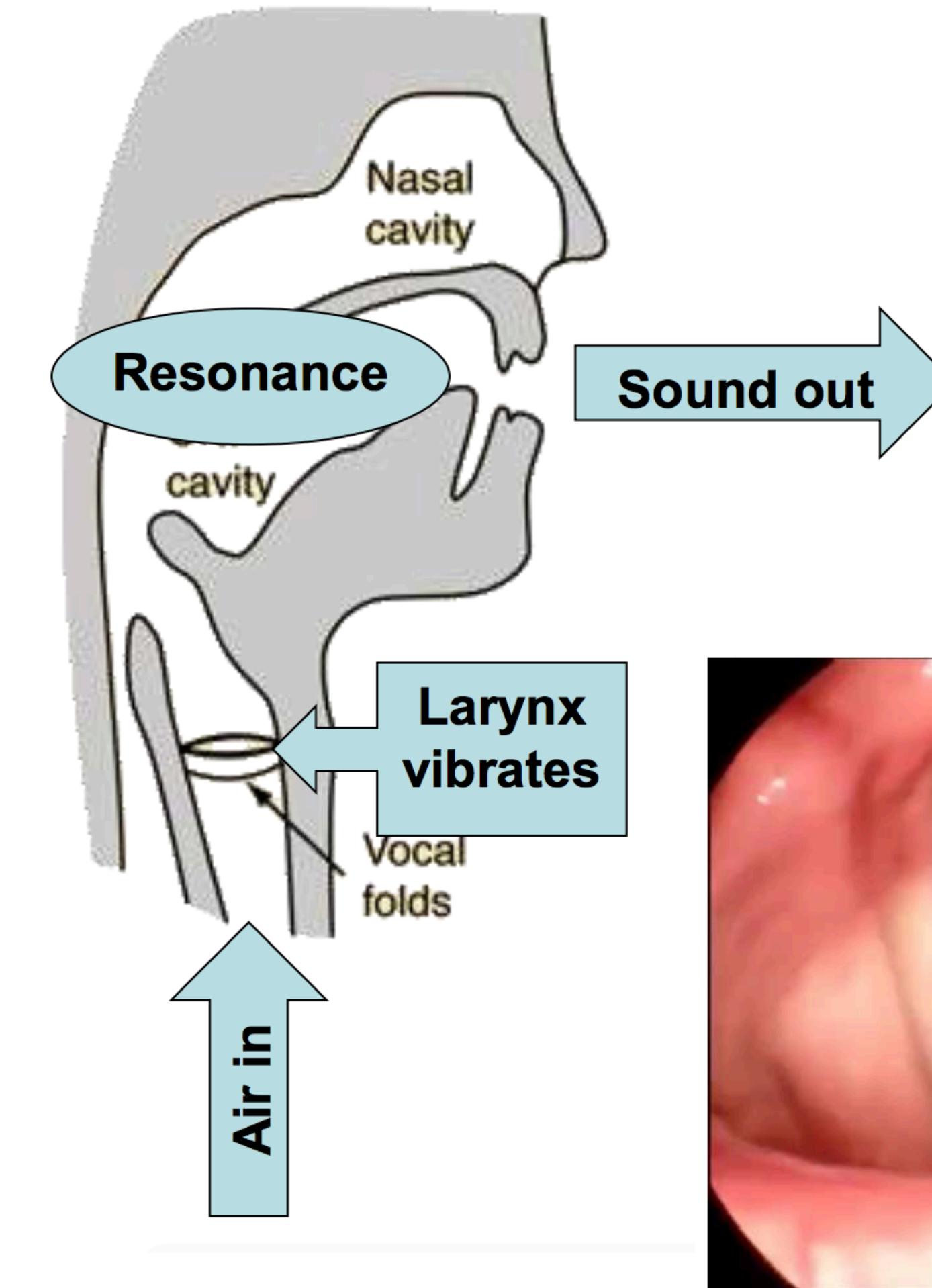
Речевой аппарат

- Для образования речи нужны те же условия, что и для образования любого звука: движущаяся сила, тело, колебания которого дадут тоны и шумы, и резонатор для оформления тембра звуков
- Речевой аппарат похож на духовой инструмент, состоящий из «мехов» (легкие), язычка (колебания, дающие тон) и резонатора (полость глотки, рта и носа).

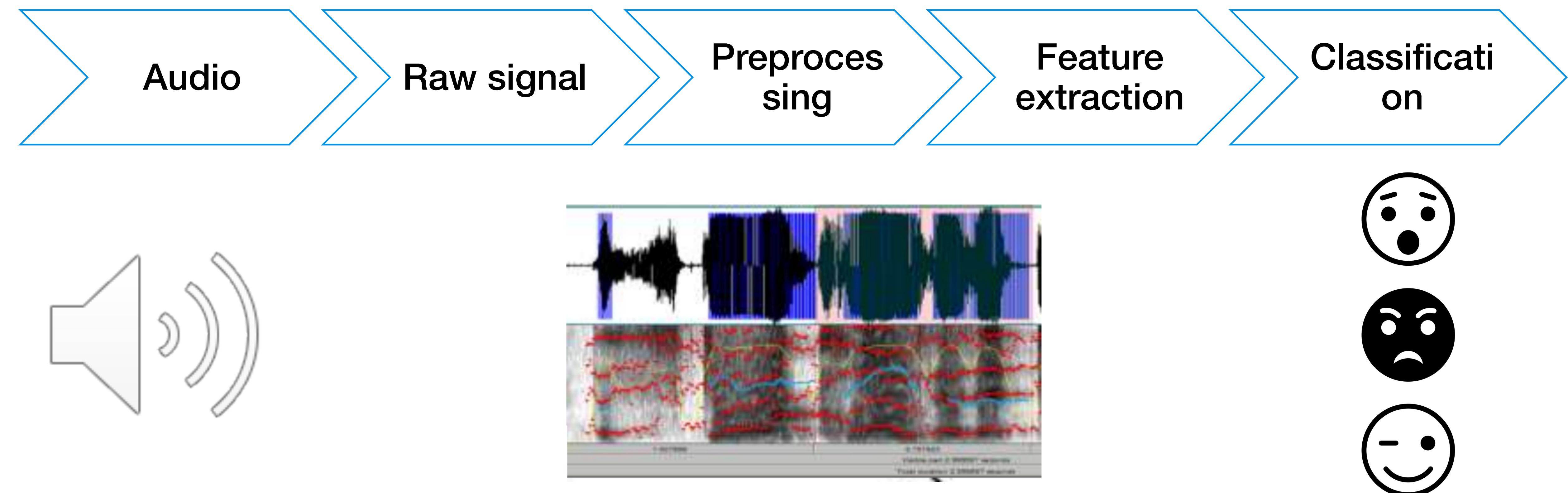


Речевой аппарат

- Легкие отправляют воздух к голосовым связкам и речевому тракту
- Голосовые связки открываются и закрываются, чтобы регулировать количество воздуха, проходящего сквозь них
- Речевой тракт меняет форму благодаря языку и небу



Анализ аудио: шаги



Анализ аудио: raw signal

- **Получение сырого сигнала: challenges**
 - Окружающие шумы (noise filtering)
 - Несколько говорящих (speaker diarization)

Анализ аудио: preprocessing

- Нормализация сигнала нужна, чтобы уменьшить вариативность
- Различные условия:
 - Разные типы записывающих устройств (телефоны, мобильные, микрофоны)
 - Разные настройки звукозаписи, расстояние до микрофона
- Различия между спикерами:
 - Половые различия
 - Различия строения голосового аппарата (гортань, например)
- Цель:
 - Снизить вариативность спикеров и окружающей среды (Interspeaker variability)
 - Оставить вариативность, связанную с эмоциями (Intraspeaker variability)

Анализ аудио: feature extraction

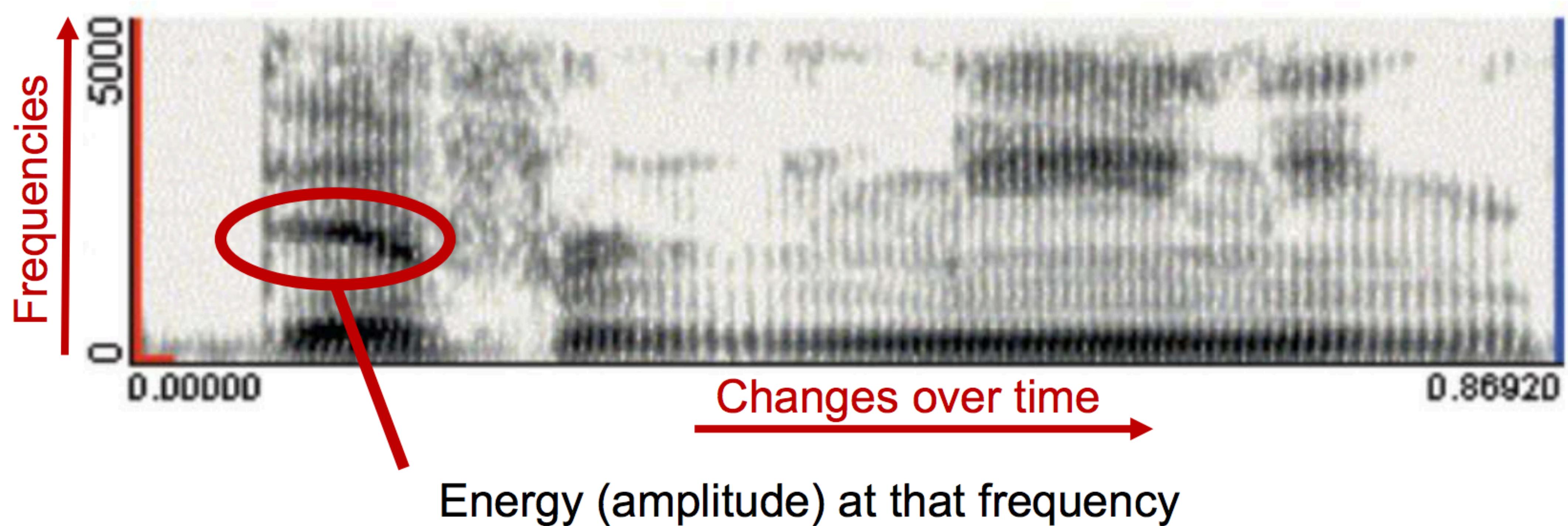
- Что делать с речевым сигналом?



Анализ аудио: feature extraction



Анализ аудио: feature extraction



Спектrogramма

- <https://musiclab.chromeexperiments.com/Spectrogram/>
- <https://tyomitch.github.io/#meklon.png>

Анализ аудио: feature extraction

- Гласные и согласные можно рассматривать как сегменты, из которых состоит речь
- Характеристики гласных и согласных называются **сегментными признаками (segmental features)**
- **Надсегментарные особенности (supra-segmental features)** (также известные как **просодические признаки**) - это параметры речи, которые включают в себя больше, чем согласные / гласные сигнала

Segmental features

- Сегментарные акустические признаки
- Гласные и согласные можно рассматривать как сегменты, из которых состоит речь
- Характеристики гласных / согласных называются сегментарными признаками
- Относятся к кратковременному спектру речи
- Различия между эмоциями в спектральных особенностях, особенно гласных

Supra-segmental features

- Большинство исследований эмоций связаны с этими признаками:
- **Pitch** (высота): среднее, медиана, стандартное отклонение, минимум, максимум и т.д.
- **Energy**: среднее, медиана, стандартное отклонение, минимум, максимум и т.д.
- **Длительность**: скорость речи, соотношение длительности вокализованной и невокализованной области, продолжительность самой длинной вокализованной области
- **Форманты**

Форманты

- Форманты представляют акустическую характеристику звуков речи (прежде всего гласных), связанную с уровнем **частоты** голосового тона и образующую **темпер** звука.
- Форманта обозначает определенную частотную область, в которой вследствие **резонанса** усиливается некоторое число гармоник тона, производимого голосовыми связками, то есть в спектре звука форманта является достаточно отчетливо выделяющейся областью **усиленных частот**.
- Принятое обозначение форманты – F. Считается, что для характеристики звуков речи достаточно выделения четырёх формант – **F_I, F_{II}, F_{III}, F_{IV}**

Практика

Практика «Речь»

<http://bit.ly/37aNhZg>

Практика «Тело»

<http://bit.ly/2GJeKz2>

HW_4

- Доделать практические задания (OpenPose, Praat)
- Прислать отчет на почту o.perepelkina@neurodatalab.com
- Дедлайн: 14 февраля