Национальный политехнический университет Украины
«Киевский политехнический институт»
Факультет Электроники

Статистический анализ вокализованных пауз и самоисправлений мужских и женских голосов спонтанной украинской и русской речи

Ольга Николаевна Ладошко

ladoshko@gmail.com

Содержание

Введение

Цель работы:

Статистически проанализировать вокализованные паузы и самоисправления спонтанной украинской и русской речи каждого диктора (мужчин и женщин) в отдельности.

Актуальность:

- **необходимость количественной оценки неинформативных (воклализ. паузы и самоисправления) с точки зрения автоматического распознавания речи особенностей спонтанной речи;**
- **необходимость правильного понимания результатов распознавания спонтанной речи (вариативность произношений дикторов) - возможно лишь в условиях распознавания каждого диктора отдельно**

Применение:

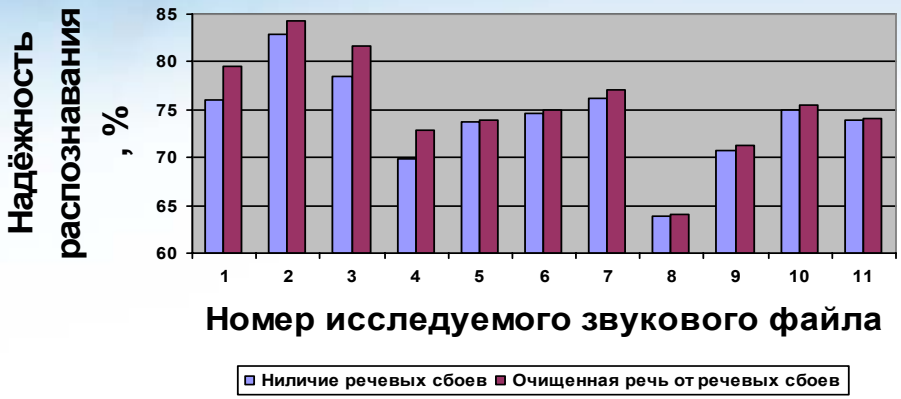
предварительная оценка корпуса по признакам спонтанности речевых данных

Предварительные исследования

Влияние речевых сбоев на процесс распознавания речи депутатов Верховной Рады Украины

Анализируемая выборка:

Влияние особенностей спонтанной речи на надёжность распознавания

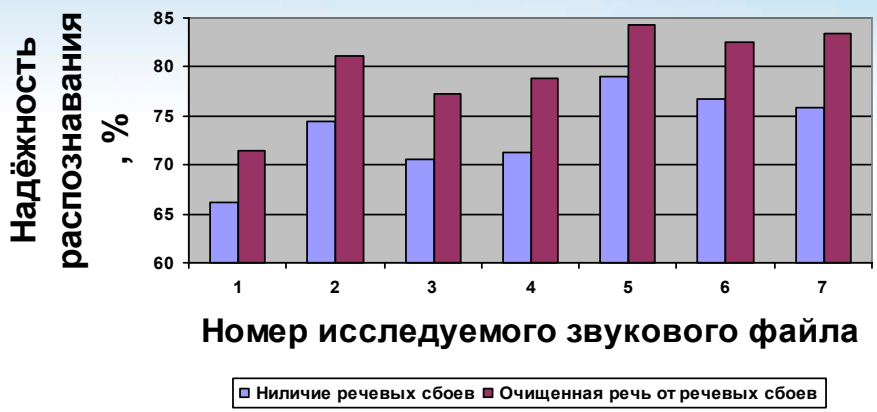


- 11 файлов заседаний ВРУ
- 10 часов 203 докладчика
- Надёжность распознавания - на 1,25%
- Наилучшее значение 3,5%
- Существенная зависимость от дикторов
- Общ. число речевых сбоев

1803 единиц – 2,44%

Контрольная выборка:

Влияние особенностей спонтанной речи на надёжность распознавания



- 7 файлов заседаний ВРУ
- 10 часов 149 докладчика
- Надёжность распознавания - на 6,45%
- Наилучшее значение 7,62%
- Существенная зависимость от дикторов
- Общ. число речевых сбоев

2156 единиц – 8,00%

Предварительные исследования спонтанной речи

Речевые сбои

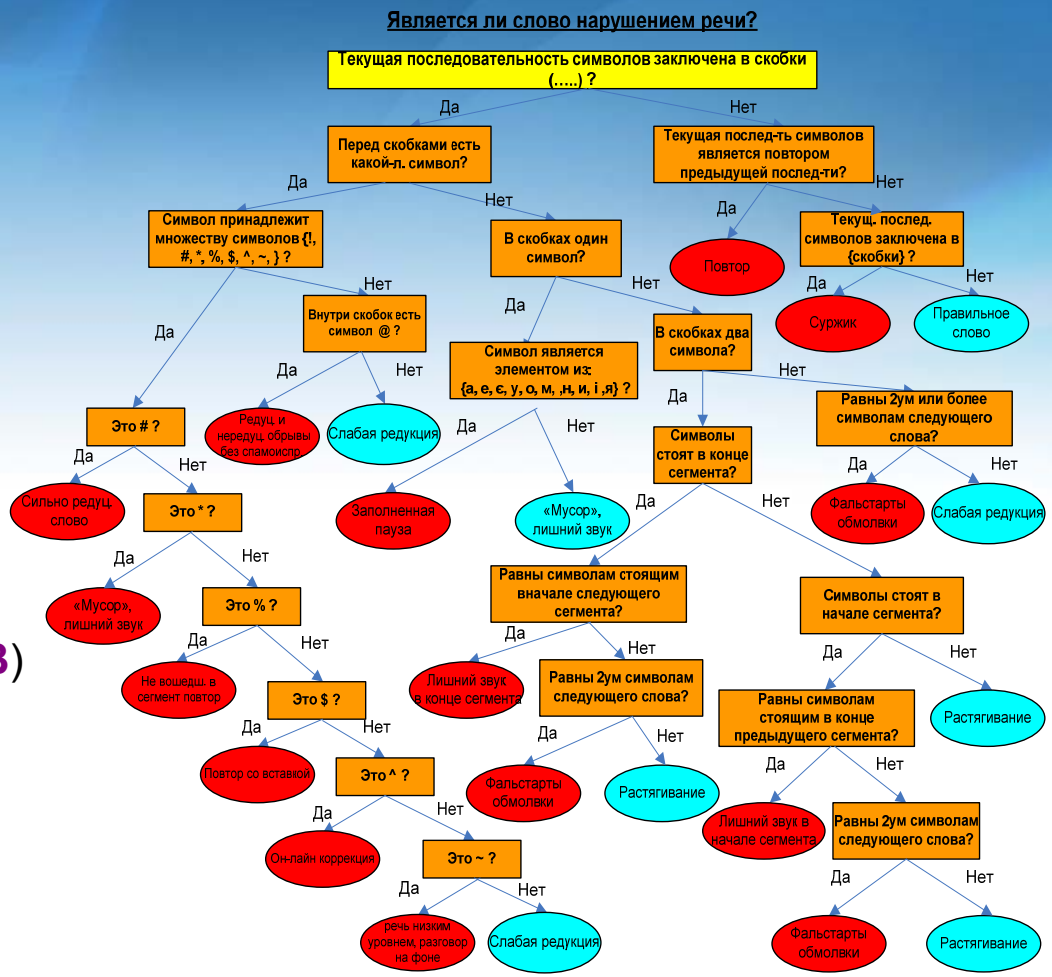
заполненные паузы («а», «ее»)
артефакты речи (кашель, вдохи)
редукция словоформ
обмолвки (фальстарты)
коррекции (самоисправления)

Дополнительные особенности

Суржик (шо, єсь)
Слова-аббревиатуры (СДПУ(о), ПДВ)
Фоновая речь
Неоднозначность сегментирования
Речь с малым уровнем

Обнаружено:

не менее 4,9% особенностей



Классификационное дерево особенностей спонтанной речи депутатов ВРУ

Вокализованные паузы

Вокализованные паузы – это паузы, в произношении которых участвуют голосовые связки

типа «а», «е», «о»

|(е) привернуть **|(е)** цю **|(е)** сферу

Виды:

- заполненные паузы, напоминающие фонемы **«а»**, **«е»**, **«о»** **и.т.п.**
- растягивание звуков **«ее»**, **«ме»**, **«аа»** (**«экание»**, **«мэкание»**, **«акание»**)
- растягивание звуков в словах

Функции:

- заполнителей промежутков в речи
- обдумывания и перепланирования речи
- позволяют избежать разрыва фраз
- могут входить в виде вставок в самоисправления диктора

Вокализованные паузы – наиболее распространённый вид нарушения плавности развёртывания речевого потока

Самоисправления

Фальстарты слов

процвітаючою, великою, **|(поту) потужною** державою

Коррекции со вставкой

Шановні колеги!Я хочу **|(попрос) |(е)** **|попросити**,

Он-лайн коррекция

Коли у видатках **|(бурже)** **|бюджету**

- **Самоисправления** - коррекции направленные на уточнение сказанного
- “Пробные шаги” диктора

- Использование вставок позволяет выделить большее время для осмысления варианта исправления

- Забракованный фрагмент речи полностью отличается от откорректированного варианта

Акустическое и текстовое наполнение

Текстовый и Аудио материал:

- Часть «Акустичного Корпусу Українського Мовлення» (148443 слова)

Особенности материала:

- Некоторым записям характерны нелинейный искажения ЧХ записывающего устройства (например, микрофона)
- Студийные условия записи (отсутствие естественных акустических условий записи)
- Присутствие характеристик подготовленной речи (актерское мастерство, чтение)
- Ограниченный тематикой лексикон (судебные заседания, политика)

Анализируемая выборка:

- 24 файла телепередач:
  - судебные заседания
  - политические дебаты
- длительность 62822 сек (~18 часов)
- Украинский язык:
  - Всего 83 дикторов

    Для анализа:
    - Женщины - 26
    - Мужчины - 38

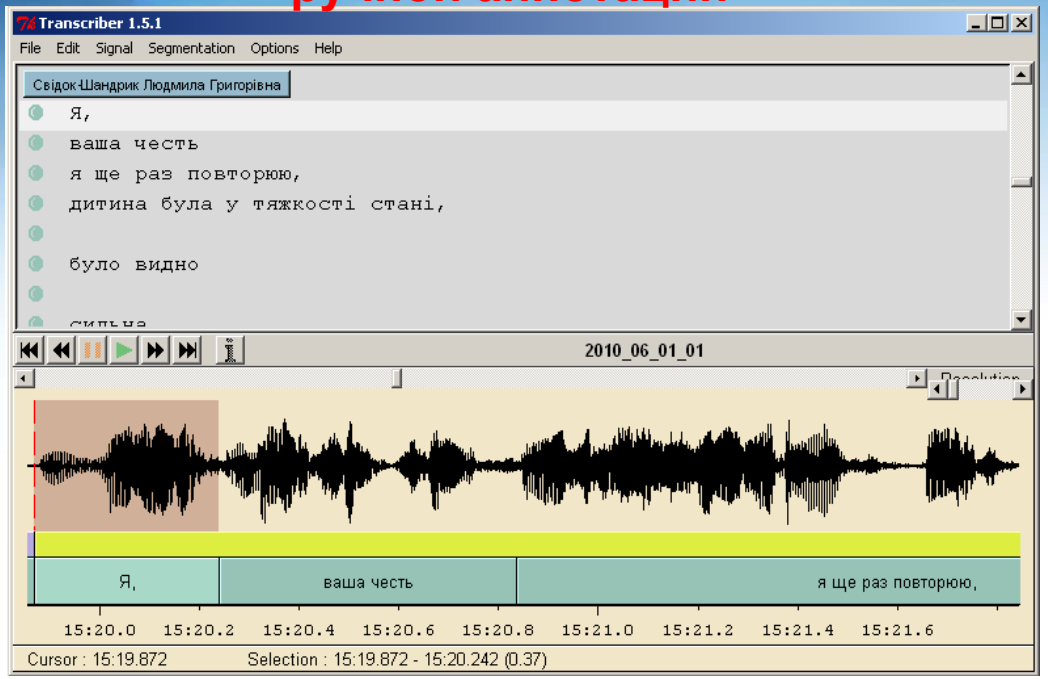
Русский язык:

- Всего 179 дикторов

  Для анализа:
  - Женщины – 68
  - Мужчины - 89

Обработка данных

Программа расшифровывания аудиозаписей и внесения необходимой ручной аннотации



Transcriber

(свободно распространяемая под GPL лицензией)

Этапы обработки материала

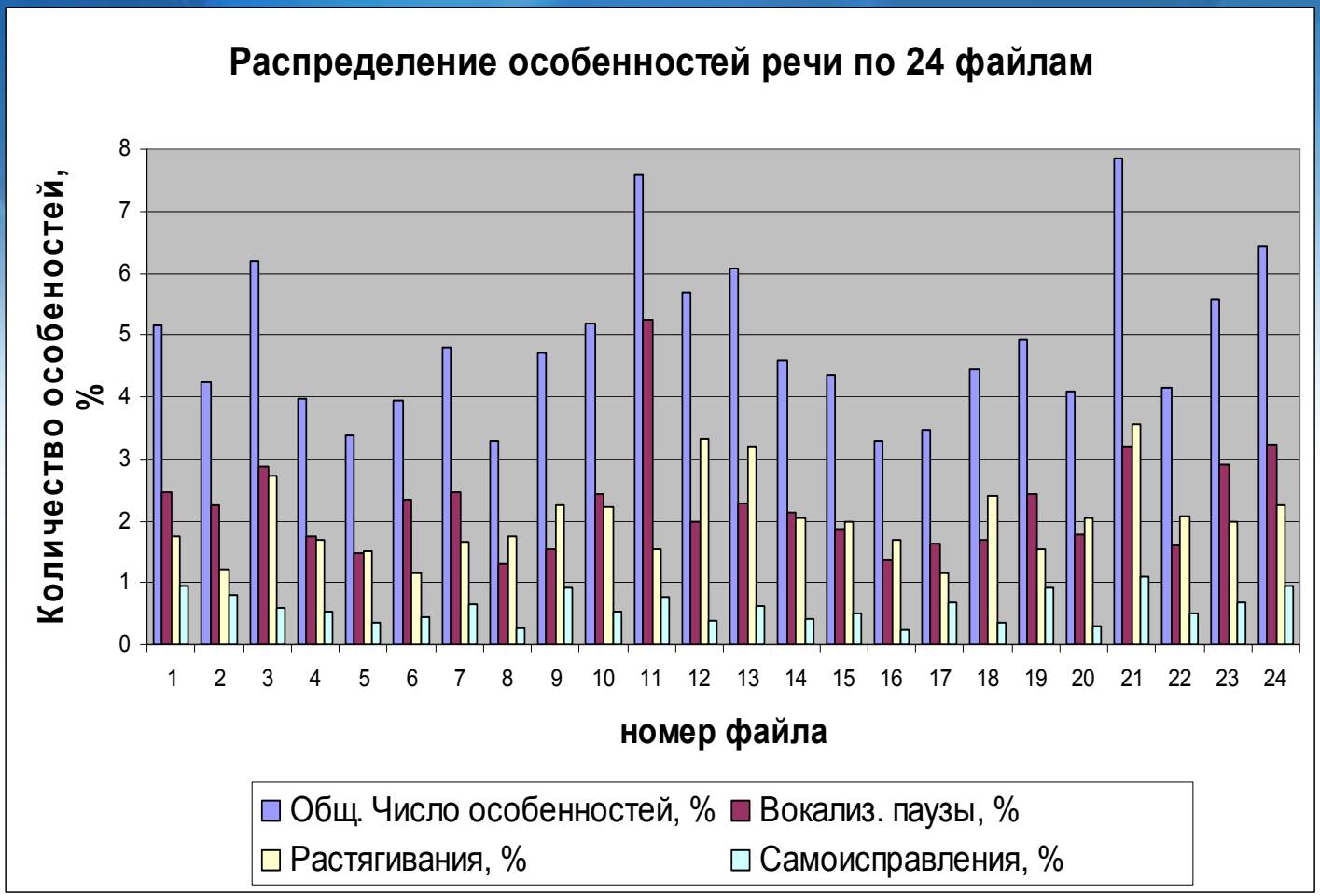
1. **Автоматическое преобразование текстовых данных Transcriber (*.trs) в требуемый для последующего анализа формат.**

2. **Автоматическое извлечение особенностей речи (программа):**

   - **Общей выборки**
   - **Каждого диктора в отдельности**

http://sourceforge.net/projects/trans/files/transcriber/1.5.1/Transcriber-1.5.1-Windows.exe/download

Статистический анализ



Всего особенностей 24 файла:
- **7485 единиц**
- **5,04% от 148443 слов**

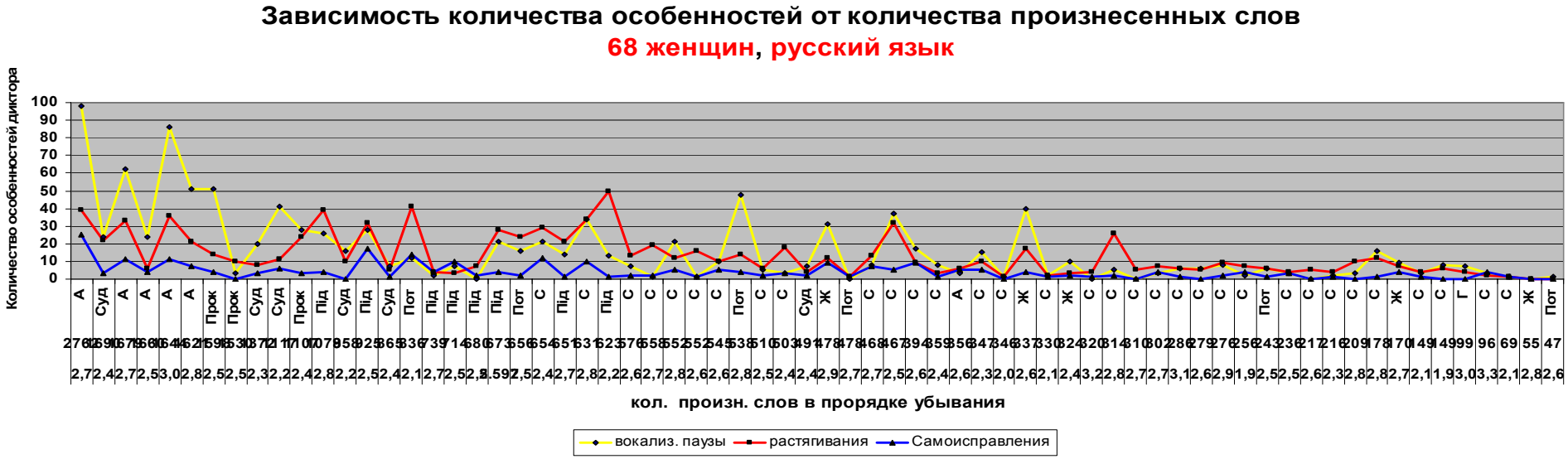
Вокализ. пауз:
- **3685 единиц - 2,48%**

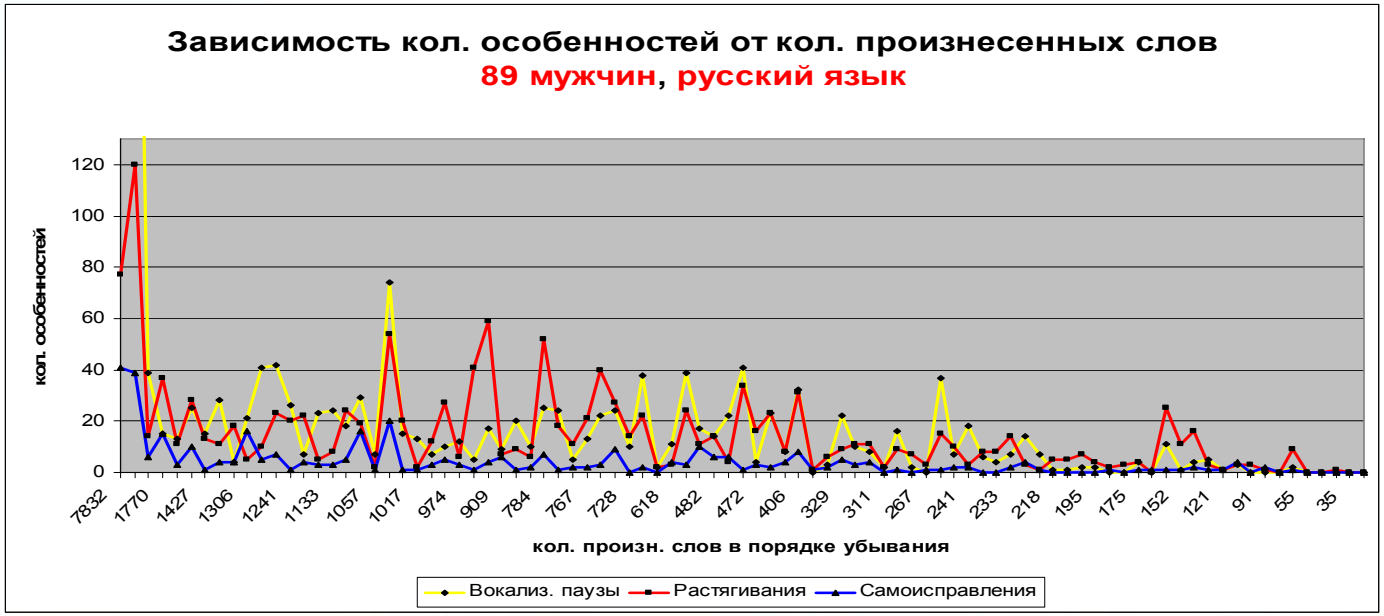
Растягиваний слов:
- **2915 единиц - 1,96%**

Самоисправлений:
- **885 единиц - 0,60%**

- **Материал характеризуется: преимущественно вокализованными паузами и растягиваниями звуков, а не самоисправлениями**

Статистический анализ

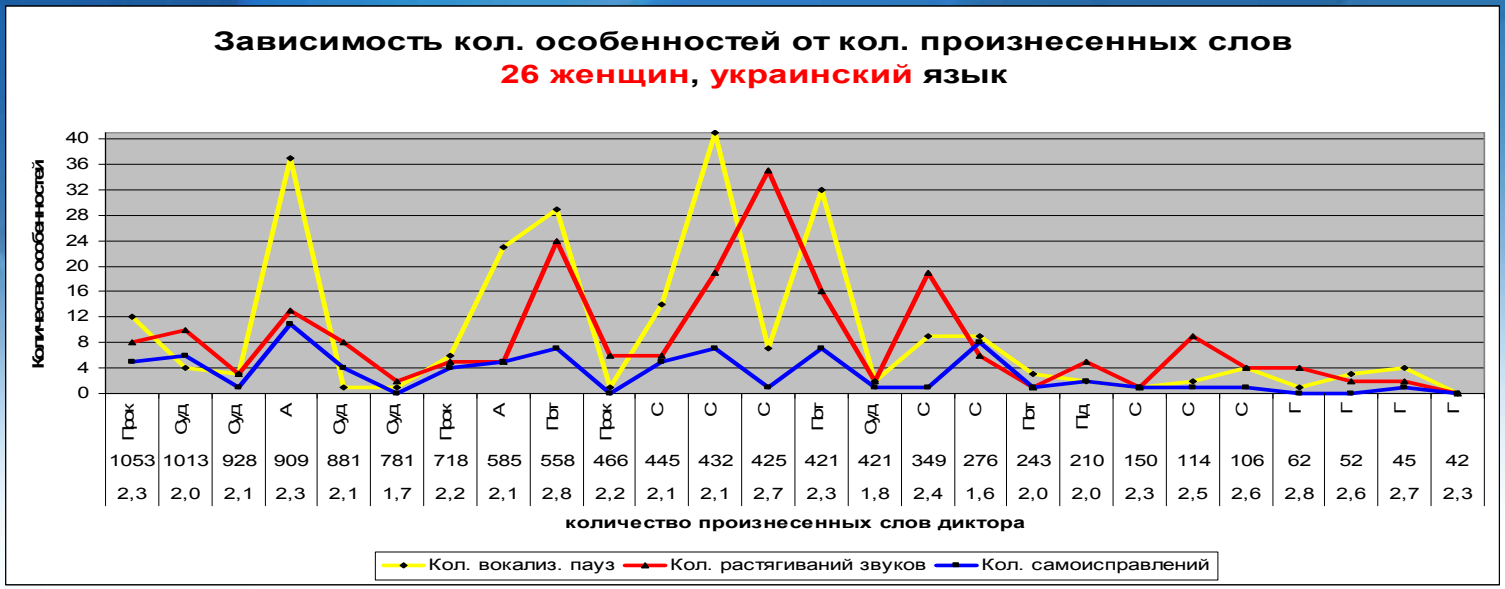




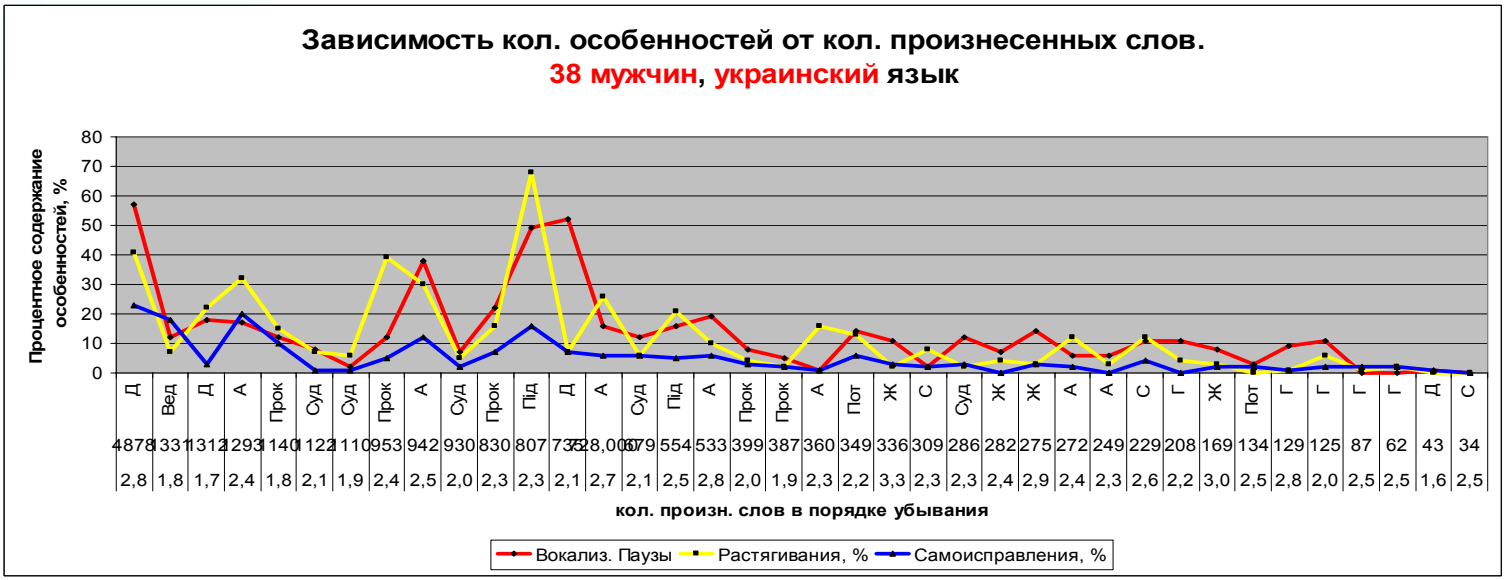
Уменьшение числа **особенностей** русской спонтанной речи с **уменьшением длительности речи** диктора

Наибольшее кол. вокализ. пауз у 1мужчины – 356 единиц

Статистический анализ

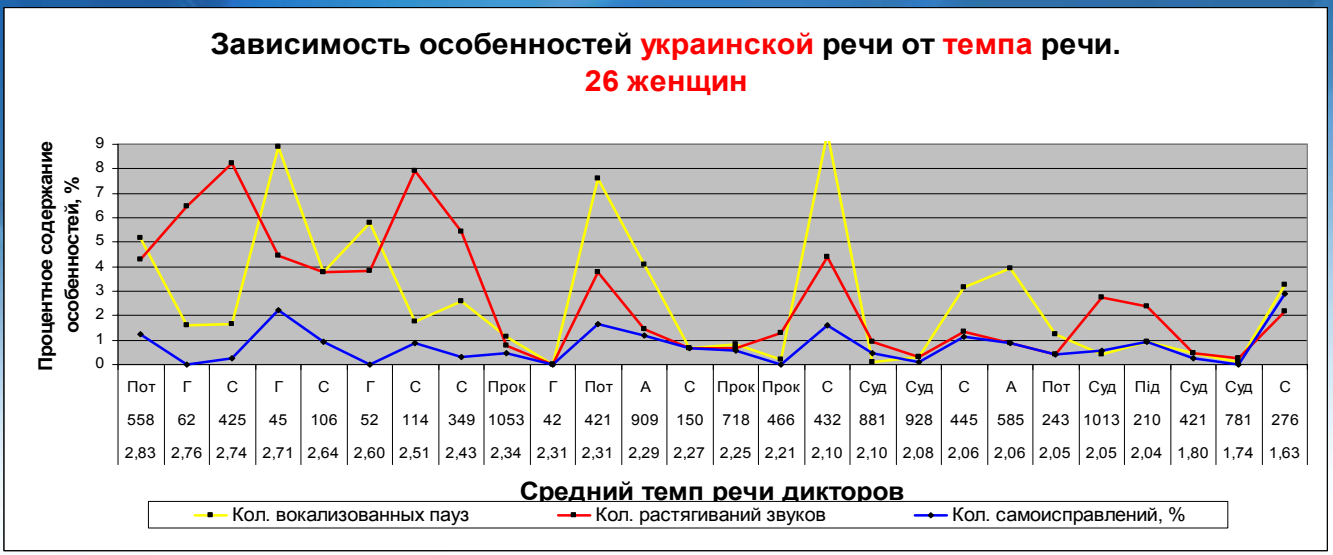


Уменьшение числа всех особенностей украинской спонтанной речи с уменьшением длительности речи диктора



Влияния типа речи (монолог, диалог, «вопрос-ответ» гостей телепередачи) на количество особенностей спонтанной речи

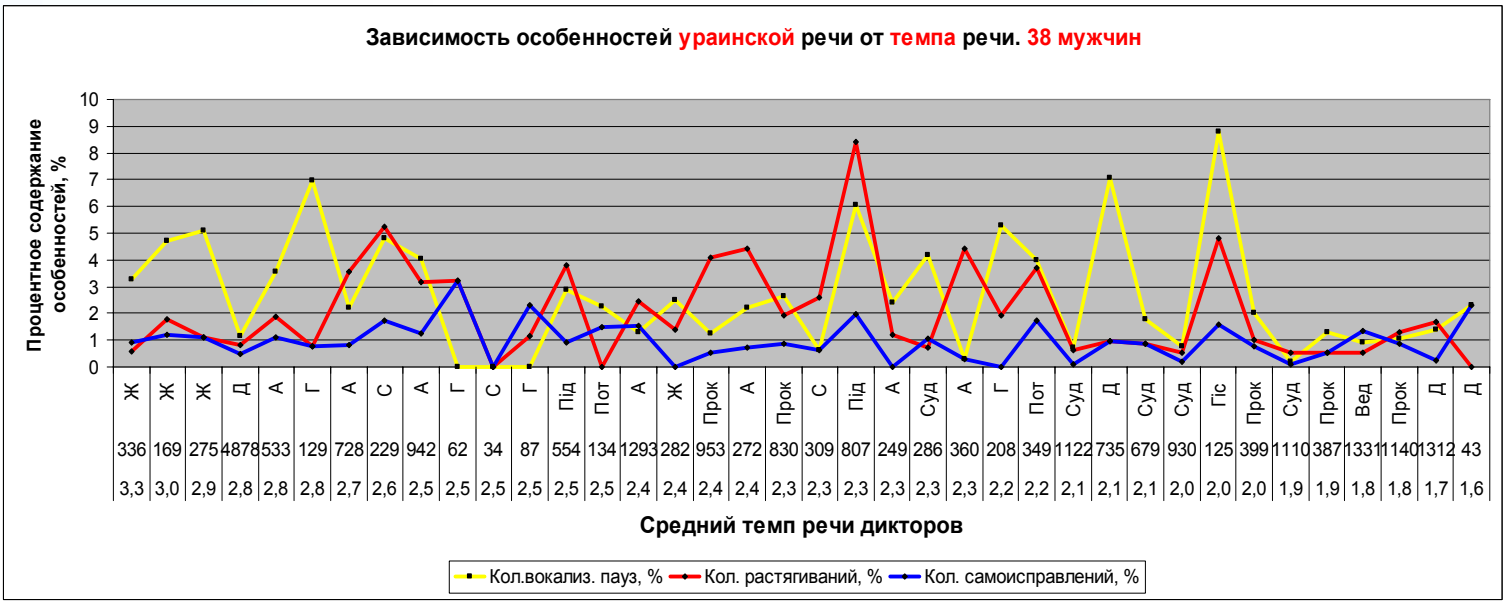
Статистический анализ



Зависимости особенностей укр. речи от трёх факторов:

- темп
- размер словаря
- соц. принадлежность (опытность диктора)

Изменение темпа речи с 2,0 -> 1,0 слова в секунду приводит к уменьшению количества особенностей речи

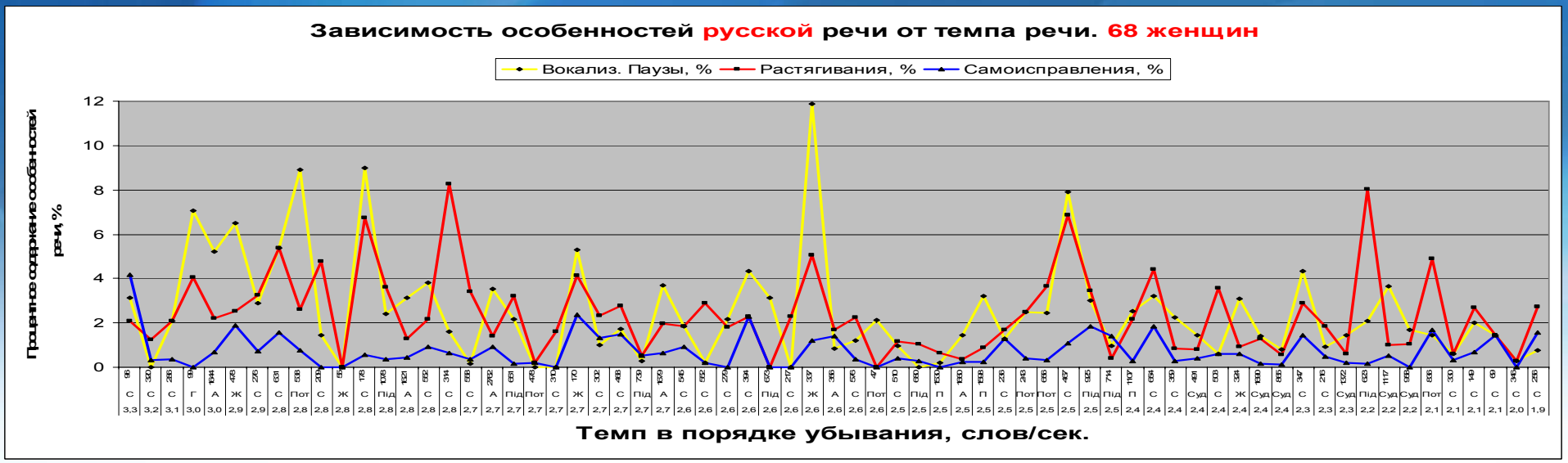


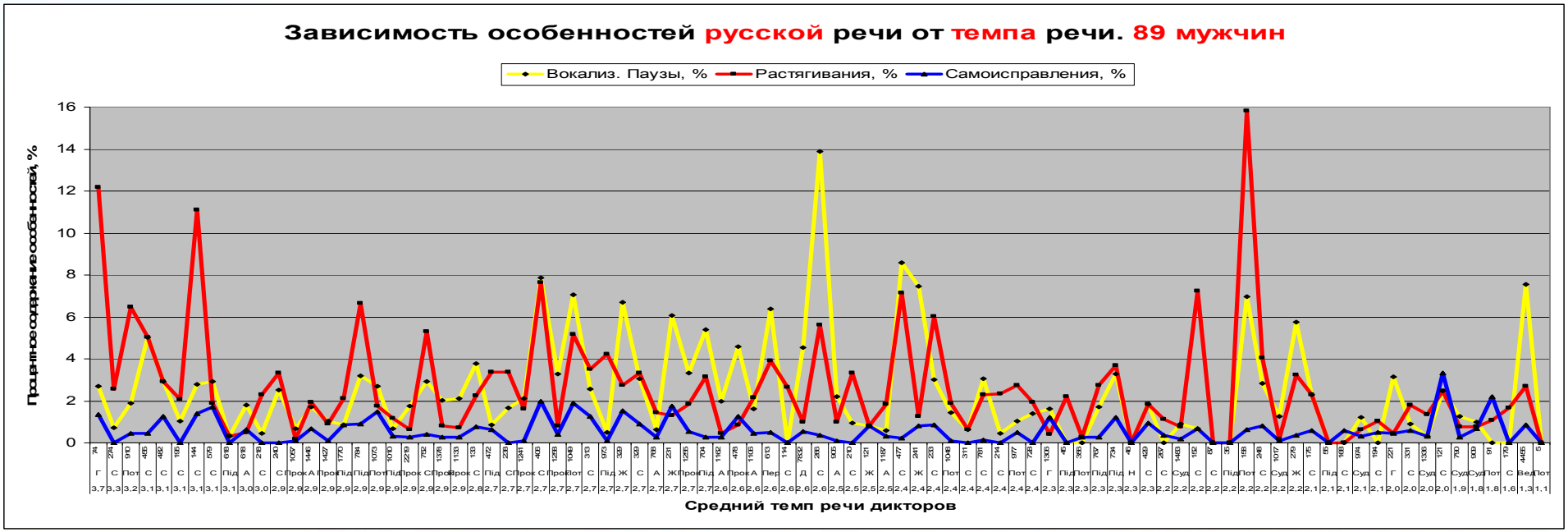
Обозначения:

Пот – потерпіла
Г – гість
С – свідок
Прок – прокурор
А – адвокат
Суд – суддя
Під – підсудний
Д – депутат
Вед – ведучий
Ж – журналіст
Н – невідомий
Прок - прокурор

Дикторы представлены незначительным разнообразием темпа речи (~2 слова/ сек)

Статистический анализ





Статистический анализ

Особенности спонтанной украинской и русской речи дикторов

		Все дикторы					Дикторы > 1000 слов			
№	Характеристики \ Дикторы	М-Рус 89чел	М-Укр 38чел	Ж-Рус 68 чел	Ж-Укр 26чел		М-Рус 12чел	М-Укр 9 чел.	Ж-Рус 6 чел.	Ж-Укр 3 чел.
1	Ср.кол. Слов	730	647	624	449		3240	2234	4048	2585
2	Ср.кол. Особенностей	39	43	33	19		160	138	207	63
3	Ср.кол. Особенностей, %	5,4%	6,7%	5,3%	4,3%		4,9%	6,2%	5,1%	2,4%
4	Ср.кол. Вокализ. пауз	21	13	16	10		97	29	124	30
5	Ср.кол. Вокализ. пауз, %	2,8%	2,1%	2,5%	2,1%		3,0%	1,3%	3,1%	1,2%
6	Ср.кол. Самоисправлений	4	5	4	3		17	14	22	12
7	Ср.кол. Самоисправлен., %	0,5%	0,8%	0,6%	0,7%		0,5%	0,6%	0,5%	0,5%
8	Ср.кол. Растягиваний	15	12	13	8		46	33	60	21
9	Ср.кол. Растягиваний, %	2,0%	1,9%	2,1%	1,8%		1,4%	1,5%	1,5%	0,8%
10	Ср.темп, слов/сек.	2,5	2,3	2,6	2,3		2,5	2,2	2,7	2,1

Наибольшее количество особенностей у мужчин, украинский язык 6,7% и 6,2%

Заключение и выводы

1. Автоматически извлечены особенности спонтанной украинской и русской речи отдельно для каждого диктора.

2. Выявлено, что на количество особенностей спонтанной речи влияют минимум 3-ри фактора:
темп, словарь, социальный статус диктора.

3. Статистический анализ части корпуса АКУМ показал:

- Неравномерное распределение особенностей спонтанной речи по файлам и по дикторам
- Уменьшение кол-ва особенностей с уменьшением словаря речи диктора и темпа речи до ~2 слов/сек.
- Материал характеризуется преимущественно вокализ.паузами
- Всего выявлено 7485 единиц особенностей - 5,04% от 148443 слов

4. Экспериментально полученные данные могут быть в дальнейшем использованы для исследований спонтанной речи.

Перспективы дальнейших исследований

1. **Исследования акустических свойств звуковых записей спонтанной речи**
   - **сформированных из имеющихся корпусов АУКМ и ВРУ разговорной речи большого объёма (не менее 10 часов речи)**
   - **для отдельных дикторов с существенным объёмом словаря речи**
2. **Выбор параметризации особенностей спонтанной речи для их численного моделирования**

Спасибо за внимание!