J. Mekyska[1], Z. Smekal[1], M. Kostalova[3, 4], M. Mrackova[2, 3], S. Skutilova[4], I. Rektorova[2,3]

[1]Signal Processing Laboratory, Department of Telecommunications, Faculty of Electrical Engineering and Communication, Brno University of Technology, Czech Republic

[2]First Department of Neurology, Masaryk University and St Anne's Hospital, Czech Republic

[3]CEITEC, Masaryk University, Czech Republic

[4]Department of Neurology, Masaryk University and University Hospital Brno, Czech Republic

Hypokinetic Dysarthria in Parkinsonian Speech: From Theory to Feature Selection

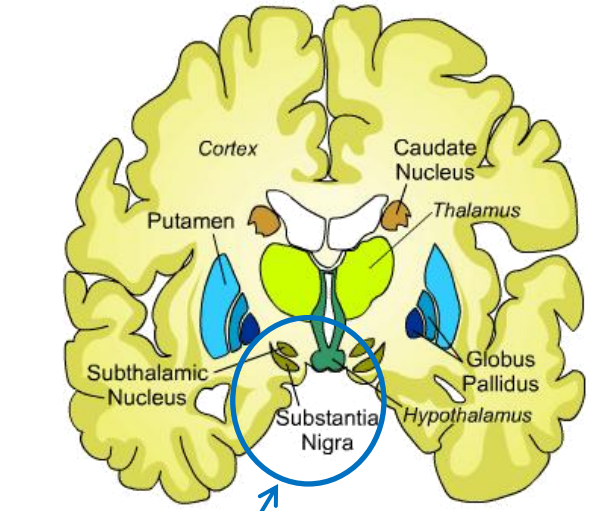
SP LAB

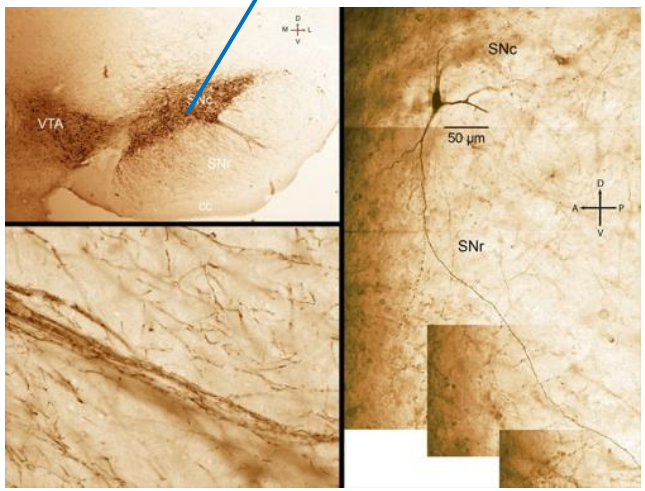
Outline

1. **Parkinson disease**
2. Analysis of Parkinsonian speech
3. Czech Parkinsonian speech corpus
4. Selection of optimal features for Parkinsonian speech analysis
5. Future work

Parkinson disease

- James Parkinson: *An Essay on the Shaking Palsy* (1817)
- Neurodegenerative brain disease
- Progressive death of dopaminergic neurons in the area of substancia nigra pars compacta and in the other areas of brain





Parkinson disease

- Symptoms
  - Rigidity – inability to initiate movement due to difficulty selecting and/or activating motor programs in the central nervous system
  - Akinesia – loss of the ability to create muscular movement
  - Bradykinesia – slowness of movement
  - Tremor
  - Hypokinetic dysarthria – reflected in speech
- Progression is not linear (quick start)

Outline

1. Parkinson disease
2. **Analysis of Parkinsonian speech**
3. Czech Parkinsonian speech corpus
4. Selection of optimal features for Parkinsonian speech analysis
5. Future work

Hypokinetic Dysarthria (HD)

- Speech disorder
- Nearly 70 % of Parkinsonians have HD
- Reflected in many areas
  - Phonation
  - Articulation
  - Prosody
  - Fluency
  - Faciokinesis

Hypokinetic dysarthria – example

- What should we concentrate on:
  - No facial expressions
  - No emotions
  - Late responses
  - Slow movement of articulatory organs
  - Silent voice
  - Long pauses in the speech

Hypokinetic dysarthria – phonation

- **Increased mean(F_0)**
- **Speech tremor – increased jitter**
  - Inability to maintain laryngeal muscles in a stable position for a longer period of time
  - Prolonged vowels
- **Hypophonia** – low speech intensity
- **Dysphonia** – hoarse or weak, excessively breathy, harsh or rough speech

Hypokinetic dysarthria – phonation

- **Small value of shimmer**
- Inability to effectively work with a breath
  - Insertion of undesirable pauses
  - Uttering a small number of syllables in one breath
- Sudden increase or decrease of speech rate
- Repetition of initial syllables
- **Hypernasality** – highlight of some formants
  - A sudden release of air flow through the nasal cavity

Hypokinetic dysarthria – articulation

- Bad quality of consonants [p], [t], [k], [b], [d], [g]
  - Improper working of articulatory organs
  - Lips, tongue tip, the centre of the tongue, tongue base, epiglottis and larynx
- **DDK (Diadochokinetic Tasks)**
  - Sequences with plosive-vocal combination
  - Pa ta ka pa ta ka pa ta ka …
  - DDK used to judge the precision of articulation

Hypokinetic dysarthria – prosody, fluency

- **Prosody**
  - Unable to express emotions (anger)
  - **Disprosody**
- **Speech fluency**
  - **Hesitation** – unintentionally breaks
  - **Palilalia** – hurriedly repeated words or single syllables
  - **Bradyphemia** – suddenly decelerated speech
  - **Tachyphemia** – suddenly accelerated speech

Features used for the description of Parkinsonian speech

- Not clear, which speech features are useful
- No statistically significant tests – small databases (10 speakers)

- **Fundamental frequency F_0**
  - Increased **mean(F_0)**
  - Increased **jitter** and **std(F_0)** in short signals
  - Decreased jitter and std(F_0) in long signals – monotonous speech from the global view

Features derived from F_0 and speech intensity

- Variation and normalization of F_0

$$F_o\mathrm{VR} = F_{o\_max} - F_{o\_min}$$

$$\mathrm{rel}F_o\mathrm{SD} = \frac{\mathrm{std}(F_o)}{\mathrm{mean}(F_o)}$$

$$\mathrm{rel}F_o\mathrm{VR} = \frac{F_o\mathrm{VR}}{\mathrm{mean}(F_o)}$$

- Speech Intensity
  - Decreased intensity – low **Short-time energy** and **Teager-Kaiser energy operator**
  - Small variation - low **shimmer**
  - Necessary to do the microphone calibration (1 Pa, 1 kHz)

Speech tempo

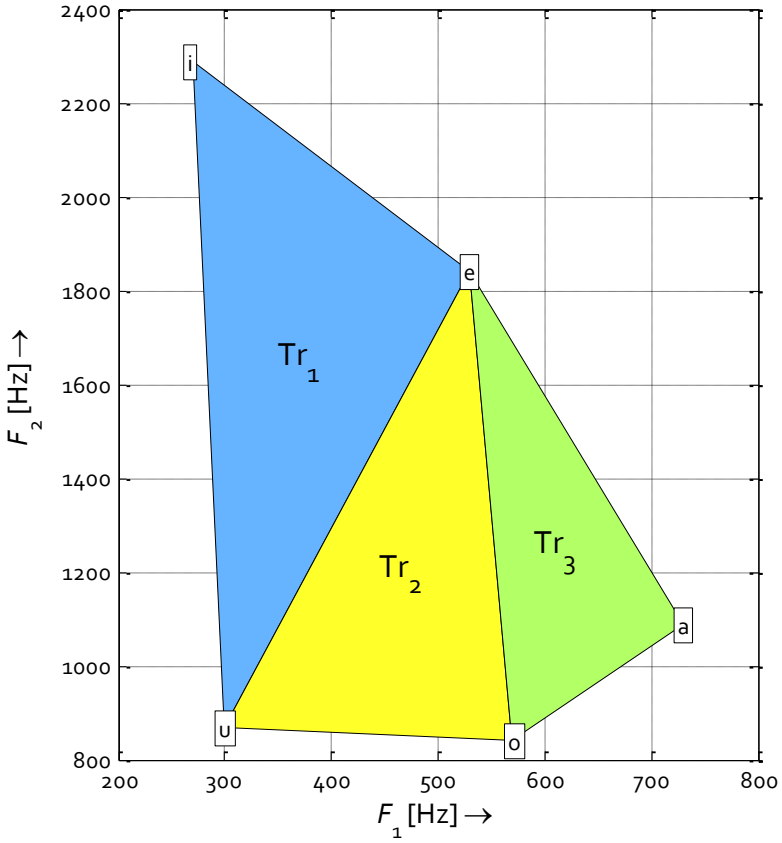
- Not clear if should be increased or decreased
- Some features derived from the speech tempo
  - **Total Speech Time** (TST)
  - **Total Pause Time** (TPT) – pauses longer than 10 ms
  - **Net Speech Time** (NST) – speech without pauses
  - **Total Intra-word Pause Time** – total length of pauses inside the multi-sillable words
  - **Total Speech Rate** (TSR) – number of syllables per second
  - **Net Speech Rate** (NSR) – total number of syllables/NST
  - **Articulation Rate** (AR) – similar to NSR, but the pauses are longer than 50 ms

Pausing

- Unintentionally breaks
- Prolonged pauses
- Some features derived from pausing
  - **Percentual Pause Ratio** (PPR) – (Total Pause Time)/(Total Speech Time)
  - **Ratio of Itra-word Pauses** (RIWP) – (Total Intra-word Pause Time)/(Total Pause Time)
  - **Interpause Speech Duration** (ISD) – time between two pauses
  - **Speech Index of Rhytmicity** (SPIR) – number of inter-word pauses per second

Formants and Vowel Space Area (VSA)

- Can be used to monitor the tongue movement
- Dependent on age and gender
- Normalization F_{2i}/F_{2u}
- **Vowel Space Area (VSA)**
  - Decreased value in case of HD
  - Monitoring the borders of tongue movement
  - Usually used lnVSA



Formant Centralization Ratio (FCR)

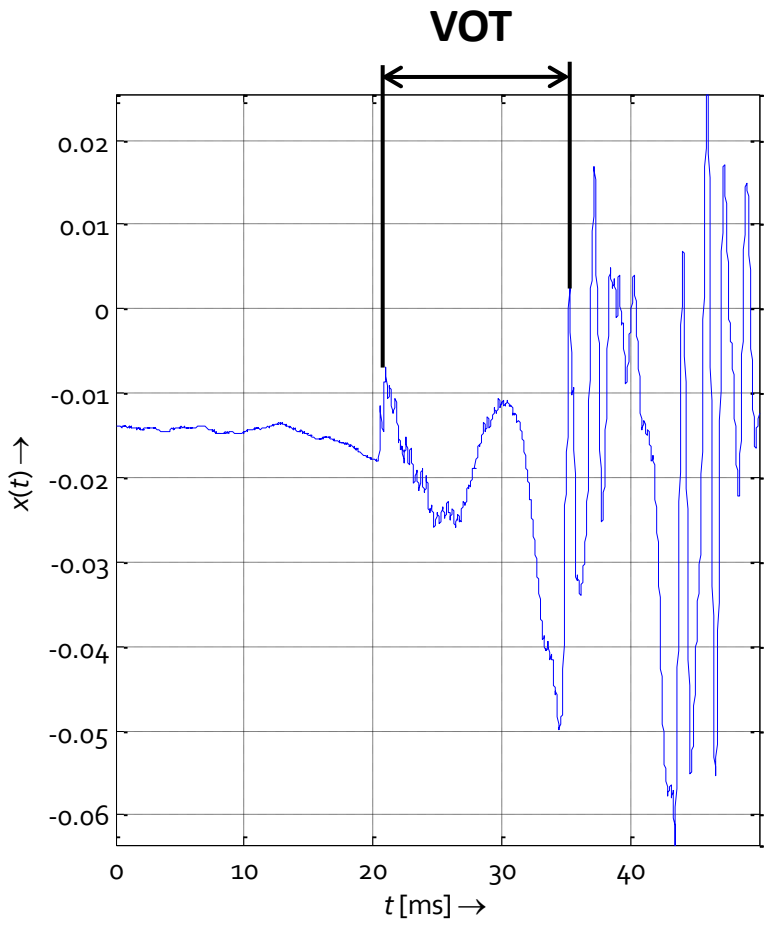
- VSA is dependent on age and gender
- Necessary to introduce a normalization

$$\mathrm{FCR} = \frac{F_{2u} + F_{2a} + F_{1i} + F_{1u}}{F_{2i} + F_{1a}}$$

- Vowel Articulation Index (VAI)

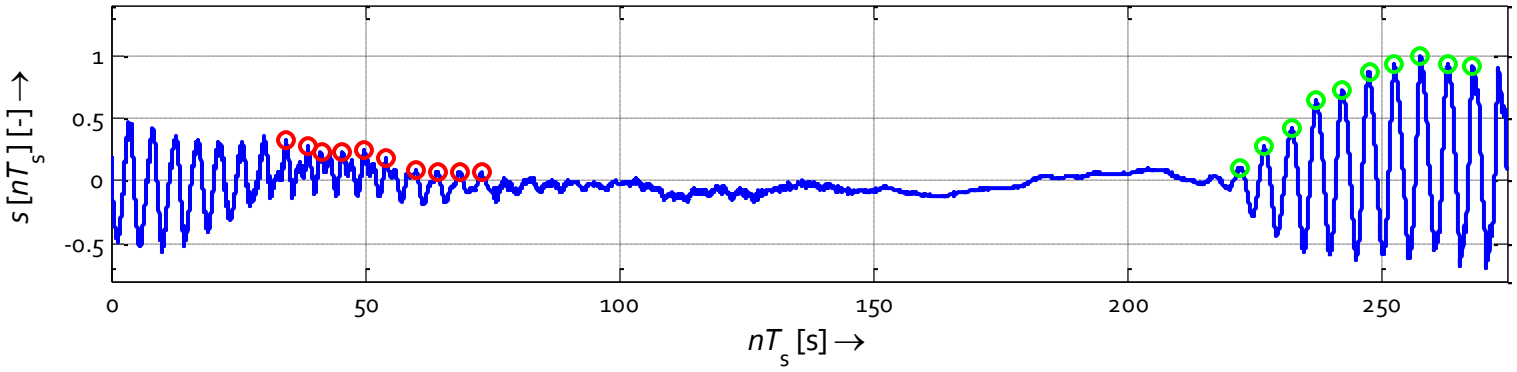
$$\mathrm{VAI} = \frac{1}{\mathrm{FCR}}$$

Voice Onset Time (VOT)



- Interval between the initial articulatory release of a stop consonant and the onset of voicing for the subsequent vowel
- Value dependent on the type of consonant
  - Short in case of bilabial consonants ([b], [p])
  - Medium in case of alveolar consonants ([d], [t])
  - Long in case of velar consonants ([g], [k])

Phonatory Offset and Phonatory Onset (POPO)



- Parkinsonians are not able to quickly end the phonation
- Phonatory offset – last 10 pulses
- Phonatory onset – first 10 pulses
- Using the semitones, there are monitored the changes of F_0 in phonatory offset/onset

Features used for the description of Parkinsonian speech

- Mentioned only speech features that can be useful according to the literature
- Many other features
  - Cepstral Peak Prominence
  - segmental features – MFCC, PLP, ACW…
  - features based on attractors
- It is necessary to find **optimal features** that can distinguish between Parkinsonians and control speakers
- Necessary to build robust **speech corpus**

Outline

1. Parkinson disease
2. Analysis of Parkinsonian speech
3. **Czech Parkinsonian speech corpus**
4. Selection of optimal features for Parkinsonian speech analysis
5. Future work

Czech Parkinsonian speech corpus

- In cooperation with St Anne's Hospital, Czech Republic
- Proposed data set
  - Parkinsonians: 100 (50 males, 50 females)
  - Control speakers: 100 (50 males, 50 females)
- To our best known, it will be the biggest and most complex database
- In some cases synchronization with fMRI
- Data acquisition
  - Speech, face, fMRI, hand writing (EUPMT)

Czech Parkinsonian speech corpus

- f_s = 48 kHz
- Spatial resolution: 768 x 576 px

- Procedure
  - Firstly the presence of Parkinson disease is tested
  - In case of Parkinsonians, the degree of disease is checked
  - The participants are doing many tasks according to the protocol

The protocol

- Protocol contain 14 main tasks, in some cases with subtasks
- **Spontaneous speech without any control**
  - What are your hobbies?
  - Useful for the detection of speech rate and prosodic characteristics
- **Reading the text**
  - Reading the same text at the beginning and at end of the protocol
  - Reading the poem - detection of rhythm
  - 1 sentence uttered in 3 different ways - monitoring the intonation
  - 1 sentence with 3 different stresses
- **Repetition according to the doctor**
  - Sentences complicated for an articulation
  - Words complicated for an articulation
  - Words with combination plosive-vocal at the beginning (VOT)
  - Short and prolonged vocals [a], [e], [i], [o], [u]
  - Silent and loud prolonged vocals [a], [e], [i], [o], [u]

The protocol

- **Diadochokinetic Tasks**
  - Ba ba ba ba ba ba ba
  - P t k p t k p t k p t k p t k
  - Pa ta ka pa ta ka pa ta ka pa ta ka
- **Maximum time of phonation**
  - Prolonged mmmmmmmmmmmmmmmm
  - Prolonged iiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiii
  - Prolonged aaaaaaaaaaaaaaaaaaaaaaaaaaaa
  - Prolonged combination of consonant and vocal Fiiiiiiiiiiiiiii
  - Prolonged combination of consonant and vocal Ryyyyyyyy

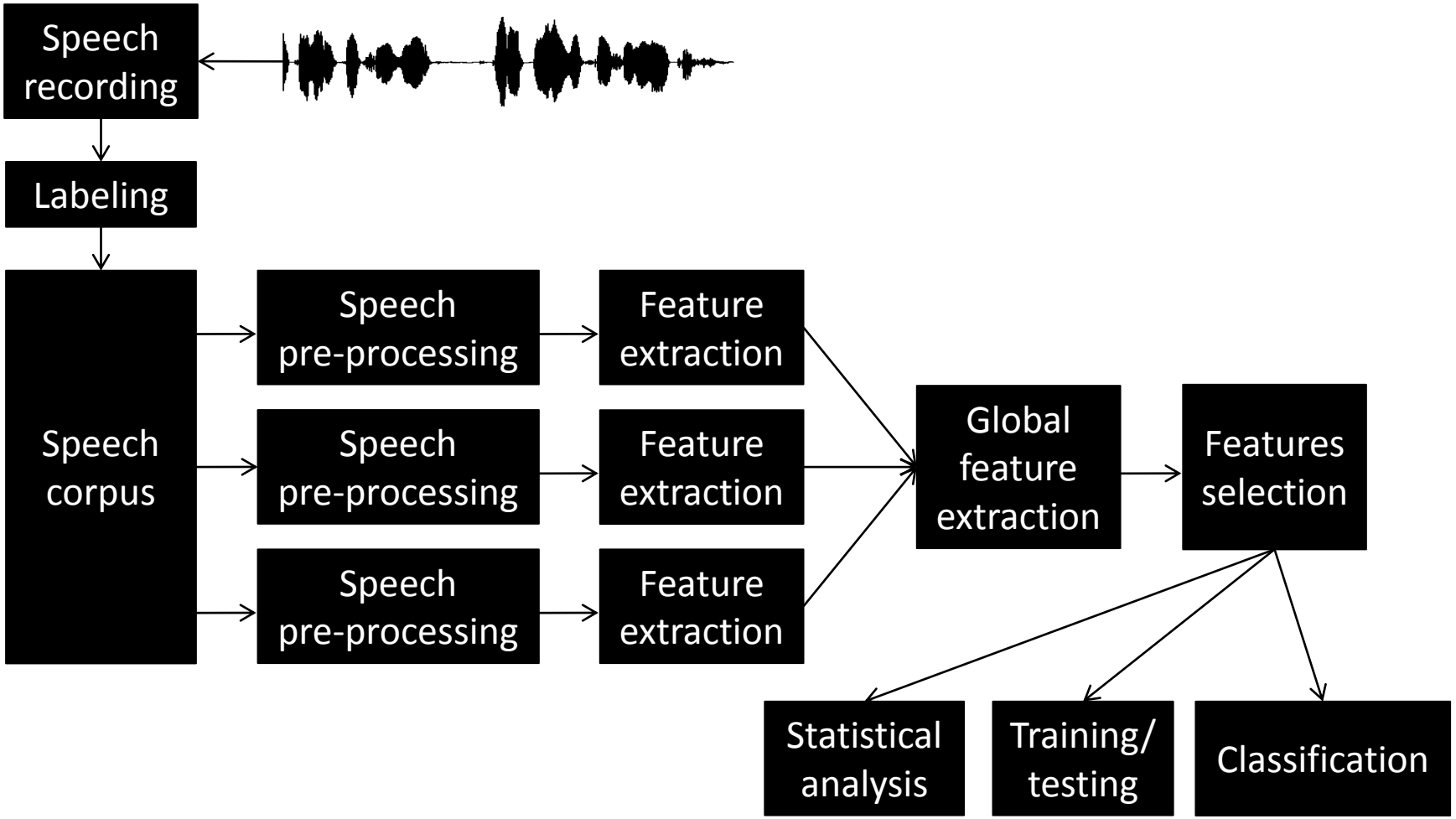
Data post-processing

- Each speaker identification is coded depending on:
  - Presence of Parkinson disease
  - Gender
  - Order
- Data are manually labeled
  - start sample – stop sample – label of task
  - Corpus XML is generated
- According to the tasks, the documentation is automatically generated (list of all present tasks)

Outline

1. Parkinson disease
2. Analysis of Parkinsonian speech
3. Czech Parkinsonian speech corpus
4. **Selection of optimal features for Parkinsonian speech analysis**
5. Future work

Selection of optimal features for Parkinsonian speech analysis



Selection of optimal features for Parkinsonian speech analysis

- Which features are suitable for diagnosis of Parkison disease?
- Input features:

  mean, median, variance, standard deviation, maximum and minimum of F_0, F_1, F_2, F_3; F_0 -VR, relF_0-SD, relF_0-VR, mean local jitter (jitter ML), absolute mean local jitter (jitter AML), mean rap jitter (jitter MR), mean ppq5 jitter (jitter MP), mean ddp jitter (jitter MD); mean, median, variance, standard deviation, maximum and minimum of bandwidth B of formants F_1, F_2, F_3; global features VSA (Vowel Space Area), lnVSA (logaritmic VSA), VAI (Vowel Articulation Index), FCR (Formant Centralization Ratio), F_{2i}/F_{2u}, mean autocorrelation, mean noise-to-harmonic ratio (mean NHR), and mean harmonic-to-noise ratio (mean HNR)

Selection of optimal features for Parkinsonian speech analysis

- Database
  - 12 male speakers with hypokinetic dysarthria
  - 42 male control speakers
- 2 sets of Czech vocals [a], [e], [i], [o], [u]
  - Natural voice intensity (sn)
  - Loud pronunciation (ln)
- 510 features for each speaker
- Selection of 20 best features
  - minimum Redundancy Maximum Relevance (mRMR)
  - Inter-Intra Class Distance Ratio (IICDR)
- In case of normal distribution (Jarque-Bera test) ANOVA was applied

Results – IICDR, mRMR, ANOVA

	IICDR METHOD		mRMR METHOD	
Order	Label	Parameter	Label	Parameter
1	u_ ln	max F_1	a_ sn	mean F_0
2	u_ ln	HNR	e_ sn	jitter MR
3	o_ ln	jitter ML	u_ ln	jitter MR
4	u_ sn	std F_1	a_ sn	jitter MP
5	u_ sn	max F_1	i_ sn	jitter MP
6	o_ ln	HNR	u_ sn	jitter MP
7	u_ ln	std F_1	e_ ln	jitter MR
8	a_ ln	std F_1	i_ ln	jitter MP
9	u_ sn	var F_1	o_ ln	jitter MR
10	a_ ln	HNR	i_ ln	local jitter
11	a_ sn	std F_1	e_ sn	jitter MD
12	o_ ln	jitter ML	u_ ln	jitter MD
13	o_ sn	mean B-F_1	e_ sn	jitter ML
14	u_ ln	jitter MR	a_ sn	jitter ML
15	u_ ln	jitter MD	o_ sn	jitter ML
16	o_ ln	jitter MD	o_ sn	jitter MR
17	o_ ln	jitter MR	o_ sn	jitter MP
18	o_ ln	jitter MP	u_ sn	jitter MR
19	i_ sn	jitter MR	a_ ln	jitter MP
20	i_ sn	jitter MD	e_ ln	min F_0

Analysis of variance (ANOVA) for parameters chosen by the ICCDR method

Label	Parameter	*p*-value
o_sn	mean B-F_1	0.0000
a_ln	mean HNR	0.0567
u_ln	jitter MD	0.4178
u_ln	jitter MR	0.4205
u_ln	mean HNR	0.5008
o_ln	mean HNR	0.9339

Analysis of variance (ANOVA) for parameters chosen by the mRMR method

Label	Parameter	*p*-value
e_ln	mean F_0	0.0880
u_ln	jitter MR	0.5189
u_ln	jitter MD	0.5226

Selection of optimal features for Parkinsonian speech analysis – results

- Significant features
  - Features based on the jitter
  - Features based on the harmonicity – reflecting the speech quality
  - Features based on F_1
- The database is still quite small
- The tests were made just for males

Outline

1. Parkinson disease
2. Analysis of Parkinsonian speech
3. Czech Parkinsonian speech corpus
4. Selection of optimal features for Parkinsonian speech analysis
5. **Future work**

Future work

- Acquire more speakers (proposed number: 200)
- Try the other features
  - Segmental features
  - Features based on attractors
  - Features describing the quality of speech
  - Features describing tempo and pausing
  - Features describing articulation
- Introduce new features
  - Feature universal for both genders
  - Involve the genetic programming

Future work

- Find better statistical analysis
- Monitor the progression of disease
- Synchronize the results obtained by the speech analysis with other results
  - fMRI
  - Face analysis
  - Hand writing analysis
- Prepare software suitable for the Parkinsonian speech analysis
  - Complex – recording, labeling, statistical analysis, classification, monitoring the progression
  - User friendly – GUI

Later questions: j.mekyska@phd.feec.vutbr.cz

Thank you for your attention

SP LAB