# Towards an Automatic Monitoring of the Neurological State of Parkinson's Patients from Speech

J.R. Orozco-Arroyave, J.C. Vásquez-Correa F. Hönig, J.D. Arias-Londoño, J.F. Vargas-Bonilla, S. Skodda, J. Rusz, and **E. Nöth**

**Shanghai, 22.03.2016**

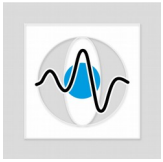
UNIVERSIDAD DE ANTIOQUIA
1803

FRIEDRICH-ALEXANDER UNIVERSITÄT ERLANGEN-NÜRNBERG
TECHNISCHE FAKULTÄT

# Outline

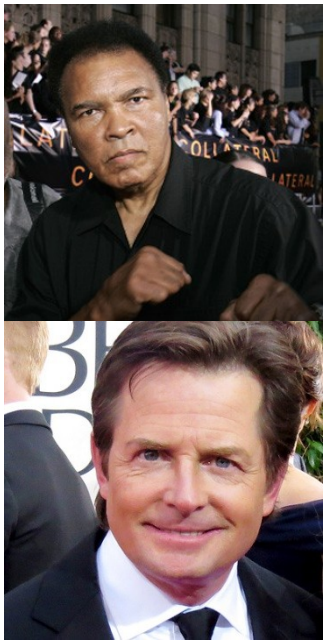
- Introduction
- Aims of the work
- Methodology
- The data
- Experiments and results
- Discussion and future work

# Introduction

# Parkinson's Disease: Motivation and Prevalence

- The second most prevalent neurological disease worldwide
- Four million patients worldwide
- Speech impairments are one of the earliest manifestations
- If the impact of the disease is delayed in 20%, the total cost during the life of the patient could decrease in up to $USD 60K



UNIVERSIDAD DE ANTIOQUIA
1803

FRIEDRICH-ALEXANDER UNIVERSITÄT ERLANGEN-NÜRNBERG
TECHNISCHE FAKULTÄT

# MDS-UPDRS: Standard Neurological Scale for PD

- MDS-UPDRS: Movement Disorders Society – Unified Parkinson's Disease Rating Scale
- Section III (MDS-UPDRS III) is only about motor activities
- Ranges from 0 to 132 (33 different items)
- Speech is evaluated only in one of those items!

# Characteristics of Parkinson's Speech

Czech

German

Colombian

- Reduced loudness
- Monotonic speech
- Breathy voice

- Imprecise articulation
- Accelerated or slowed
- Stutter-like

## Dysarthric Speech

UNIVERSIDAD DE ANTIOQUIA
1803

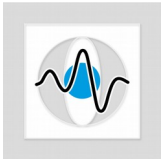
FRIEDRICH-ALEXANDER UNIVERSITÄT ERLANGEN-NÜRNBERG
TECHNISCHE FAKULTÄT

# Aim of this Study

To assess the neurological state of Parkinson's patients (according to the MDS-UPDRS-III score) from speech in three different languages:

- Spanish
- German
- Czech

# Patients and speech tasks

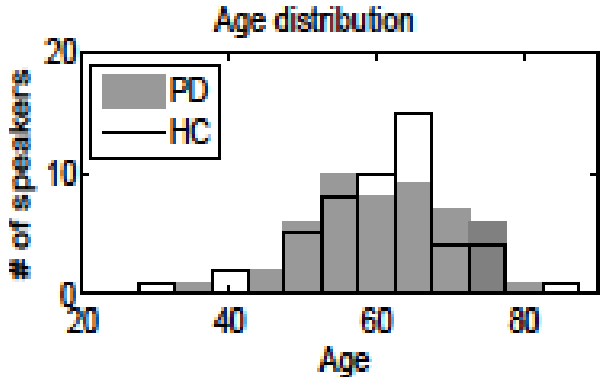



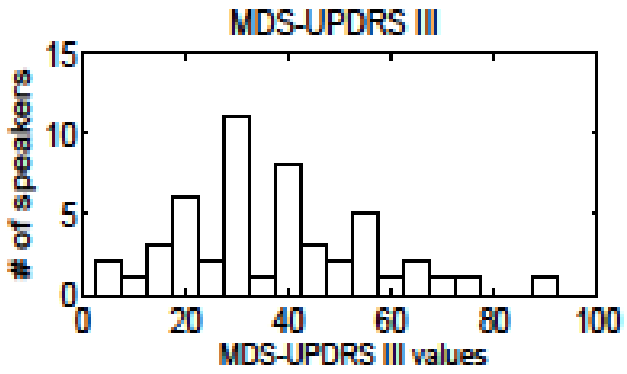
Three databases with recordings in three languages: Spanish, German, and Czech

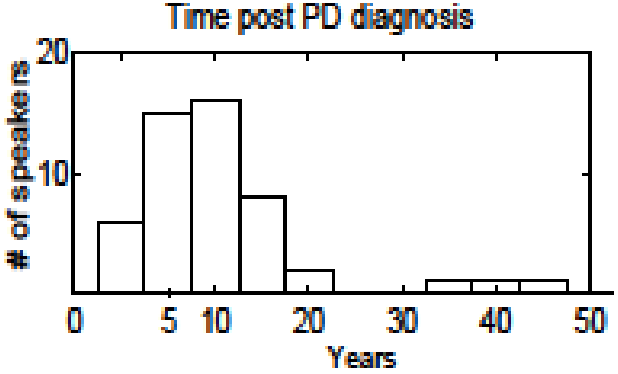
Isolated words, sentences, read texts, and monologues

# SPANISH (PC-GITA)

- 50 patients and 50 healthy controls
- Age around 60 years
- Sound-proof booth
- Speech tasks: 21 words, 6 sentences, read text, and monologue



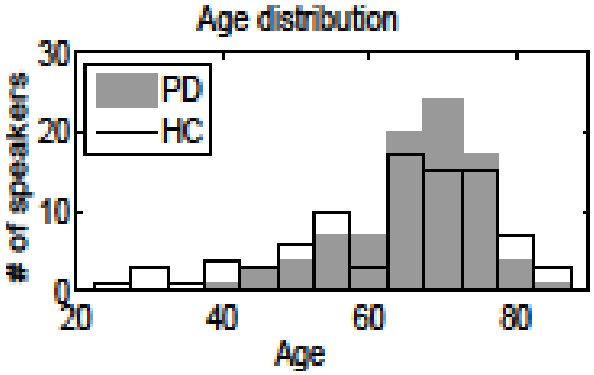



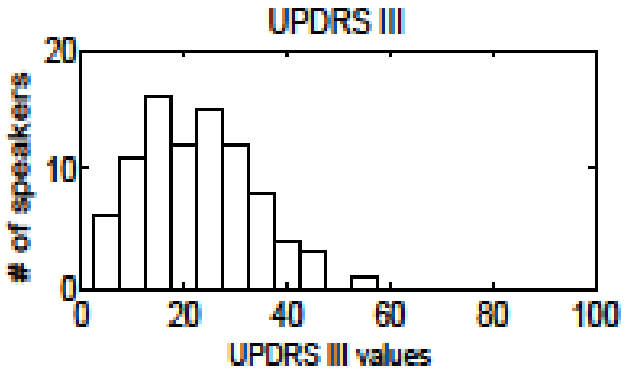

UNIVERSIDAD DE ANTIOQUIA
1803

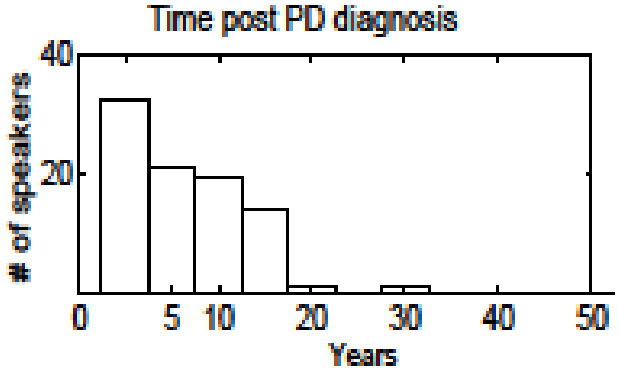
FRIEDRICH-ALEXANDER UNIVERSITÄT ERLANGEN-NÜRNBERG
TECHNISCHE FAKULTÄT

# GERMAN

- 85 patients and 85 healthy controls
- Age around 64 years
- No sound-proof booth
- Speech tasks: 6 words, 5 sentences, read text, and monologue



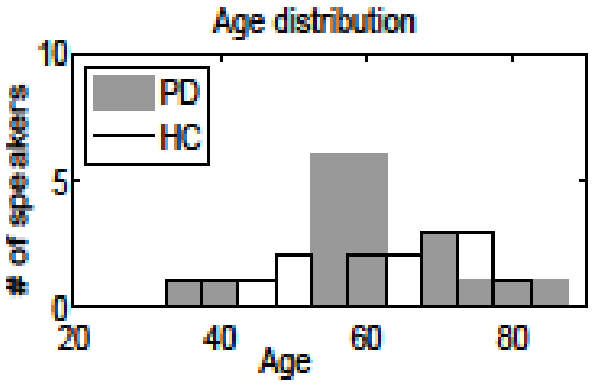

UNIVERSIDAD DE ANTIOQUIA 1803

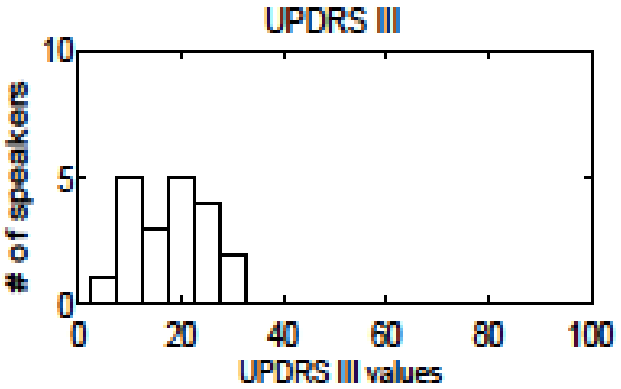
FRIEDRICH-ALEXANDER UNIVERSITÄT ERLANGEN-NÜRNBERG

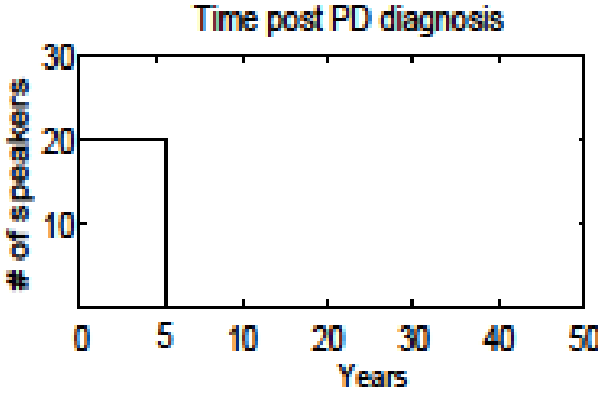
TECHNISCHE FAKULTÄT

# CZECH

- 20 patients (diagnosed at the recording session) and 15 healthy controls
- Age around 60 years
- No sound-proof booth
- Speech tasks: 7 words, 3 sentences, read text, and monologue




UNIVERSIDAD DE ANTIOQUIA
1803

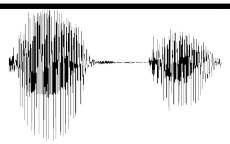
FRIEDRICH-ALEXANDER UNIVERSITÄT ERLANGEN-NÜRNBERG
TECHNISCHE FAKULTÄT

# Methodology

UNIVERSIDAD DE ANTIOQUIA
1803

FRIEDRICH-ALEXANDER
UNIVERSITÄT
ERLANGEN-NÜRNBERG
TECHNISCHE FAKULTÄT

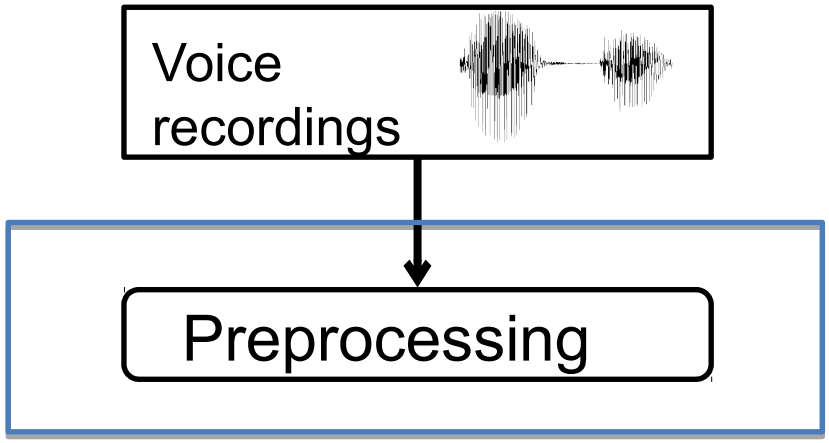
Voice recordings



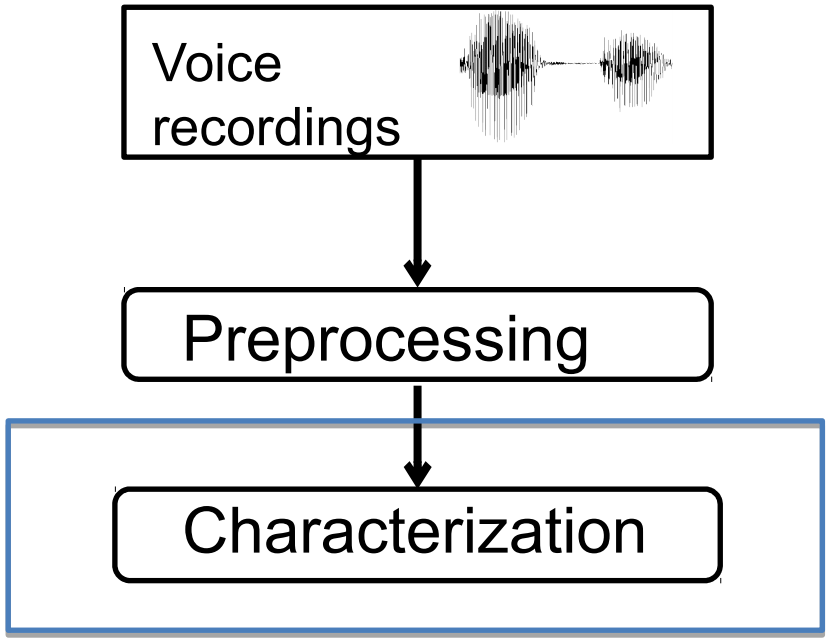
Spanish
German
Czech

UNIVERSIDAD DE ANTIOQUIA
1803

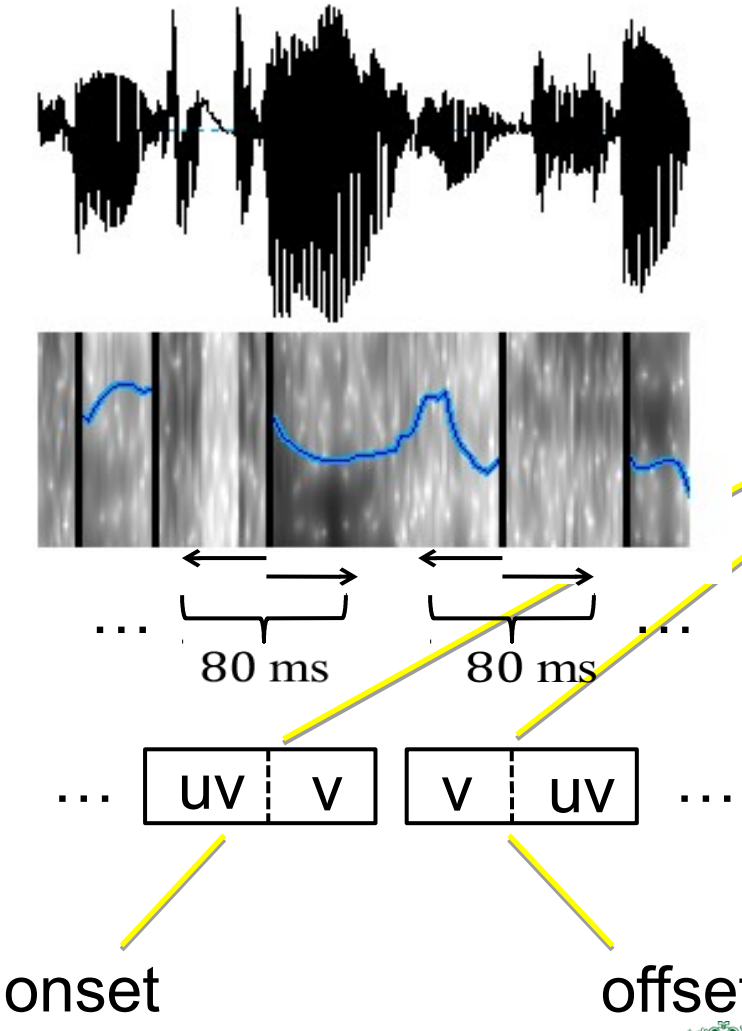
FRIEDRICH-ALEXANDER UNIVERSITÄT ERLANGEN-NÜRNBERG
TECHNISCHE FAKULTÄT




- Amplitude normalization
- Re-sampling down to 16kHz.
- Mean Cepstral Subtraction


Voice recordings
Preprocessing
Characterization

# Characterization: Onsets & Offsets




- 12 MFCC: Mel Frequency Cepstral Coefficients
- 25 BBE: Bark Band Energies

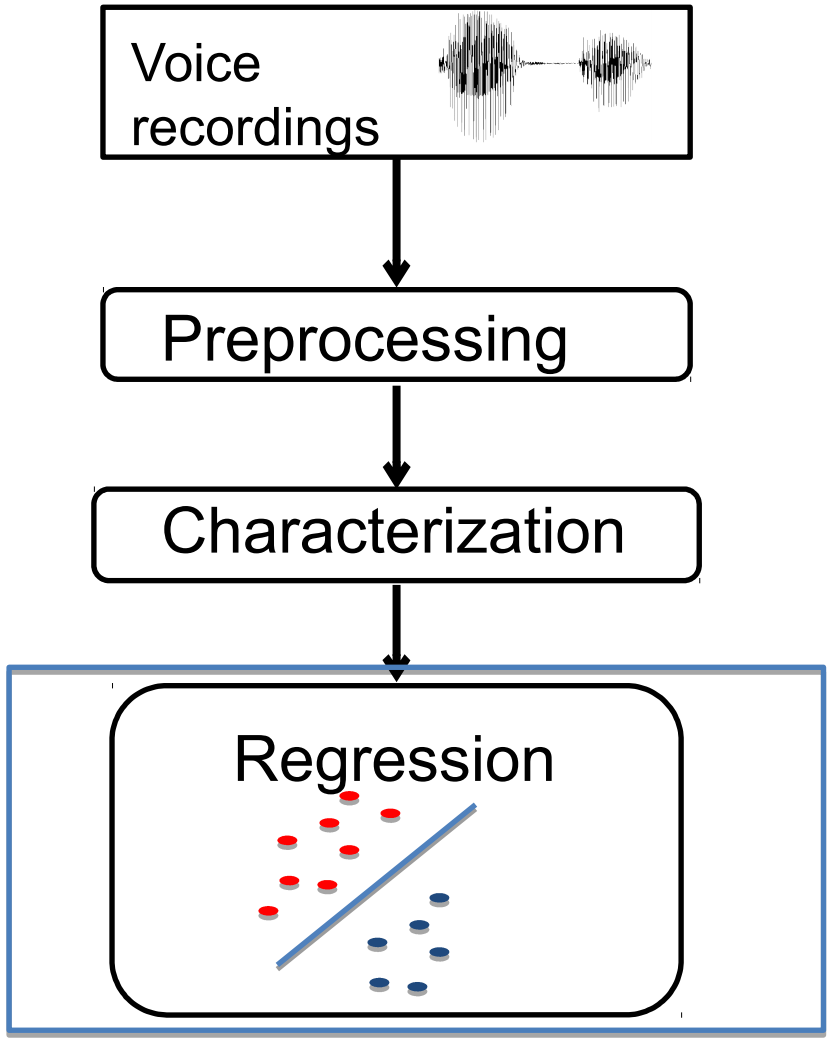
UNIVERSIDAD DE ANTIOQUIA 1803

FRIEDRICH-ALEXANDER UNIVERSITÄT ERLANGEN-NÜRNBERG

TECHNISCHE FAKULTÄT

# Characterization: Speech Intelligibility (Monologues are not Considered)

- Off-the-shelf speech recognizer: Google Inc.® (API v2)
- es-CO: Colombian; de-DE: German; cs: Czech
- Correctly pronounced words are counted and Word Accuracy (WA) is computed




ε-SVR with linear kernel

UNIVERSIDAD DE ANTIOQUIA 1803

FRIEDRICH-ALEXANDER UNIVERSITÄT ERLANGEN-NÜRNBERG

TECHNISCHE FAKULTÄT

# Experiments and Results

UNIVERSIDAD DE ANTIOQUIA
1803

FRIEDRICH-ALEXANDER UNIVERSITÄT ERLANGEN-NÜRNBERG
TECHNISCHE FAKULTÄT

# Experiment 1: Cross Validations and LOSO

## Validation strategies

- Spanish & German: 10-fold cross validation
- Czech: LOSO (leave-one-speaker-out)

## Optimization of SVR meta-parameters ($C$ and $\varepsilon$)

Grid search: $C \in \{10^{-4}, 10^{-3}, \ldots, 10\}$ and $\varepsilon \in \{1, 10, 20, 30\}$

## Evaluation of the system

Spearman's ($\rho$) correlation between the predicted values and the MDS-UPDRS-III labels

UNIVERSIDAD DE ANTIOQUIA
1803

FRIEDRICH-ALEXANDER UNIVERSITÄT ERLANGEN-NÜRNBERG
TECHNISCHE FAKULTÄT

## Spearman's Correlations:

| | Energy of onset transitions | | | |
|---|---|---|---|---|
| | Words | Sentences | Read text | Monologue |
| Spanish | 0.44 | 0.49 | 0.44 | 0.56 |
| German | 0.36 | 0.28 | 0.41 | 0.55 |
| Czech | 0.29 | 0.25 | 0.21 | 0.23 |
| | **Energy of offset transitions** | | | |
| Spanish | 0.46 | 0.46 | 0.53 | 0.74 |
| German | 0.37 | 0.24 | 0.36 | 0.31 |
| Czech | 0.24 | 0.18 | 0.35 | 0.35 |
| | **Intelligibility: WA** | | | |
| | Words | Sentences | Read text | All |
| Spanish | 0.39 | 0.20 | 0.07 | 0.49 |
| German | 0.12 | 0.18 | 0.19 | 0.31 |
| Czech | 0.22 | 0.16 | 0.15 | 0.25 |

# Experiment 2 (only Spanish): Train with PC-GITA and Test with the INTERSPEECH Challenge 2015

| Train and Development sets | Test set |
|---|---|
| PC-GITA | INTERSPEECH Challenge 2015 |
| Recorded in a sound proof booth | Eleven new patients |
| Mean MDS-UPDRS-III: 37 | Mean MDS-UPDRS-III: 39 |
| ε and *C* are optimized on development | Different acoustic conditions |
| Only words, sentences, read texts, and monologues are considered | Only words, sentences, read texts, and monologues are considered |

Evaluation of the systems: Spearman's correlation.
Baseline:  & Winners:

UNIVERSIDAD DE ANTIOQUIA
1803

FRIEDRICH-ALEXANDER UNIVERSITÄT ERLANGEN-NÜRNBERG
TECHNISCHE FAKULTÄT

## Spearman's Correlations: $\rho$

| | Energy of onset transitions | | | |
|---|---|---|---|---|
| | Words | Sentences | Read text | Monologue |
| *C*; ε | | | | |
| Train/Dev | 0.28 | 0.44 | 0.48 | 0.42 |
| Train/Test | 0.33 | 0.63 | 0.23 | 0.39 |
| | **Energy of offset transitions** | | | |
| *C*; ε | | | | |
| Train/Dev | 0.24 | 0.23 | 0.23 | 0.62 |
| Train/Test | 0.22 | 0.52 | -0.24 | 0.12 |
| | **Intelligibility: WA** | | | |
| | | Sentences | Read text | All |
| *C*; ε | | | | |
| Train/Dev | 0.40 | 0.22 | 0.28 | 0.44 |
| Train/Test | 0.56 | 0.31 | 0.51 | 0.69 |

UNIVERSIDAD DE ANTIOQUIA 1803

FRIEDRICH-ALEXANDER UNIVERSITÄT ERLANGEN-NÜRNBERG

TECHNISCHE FAKULTÄT

# Experiment 3: Considering the Ranks of the Predictions

- The aim of the community is to develop technology to monitor the neurological state of the patients
- It is not imperative to predict the real value of the UPDRS score, but to state whether the patient's state is improved or not
- The same SVR with no further optimization is used
- Measurements: WA and Energy of onsets

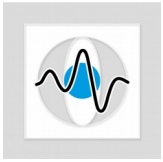
## Experiment 3 (cont.)

- The obtained results are ranked and correlated with the original UPDRS scores

$$\rho_{WA} = 0{,}59 \quad \text{and} \quad \rho_{Onsets} = 0{,}69$$

- When WA and Onsets are combined: $\rho_{all} = 0{,}72$

# Discussion and Future Work

- The energy content of the onset transitions is suitable to predict the neurological state of PD patients

- The results on Spanish recordings are better because those patients are in a more advanced stage of the disease

- An off-the-shelf Speech Recognizer can be used to develop an intelligibility-based test for telemonitoring PD patients

- The proposed approach is slightly better than the method proposed by the winners of the ComParE 2015

# Future work

- Multi-modal analyses: speech, handwriting, and gait
- Longitudinal analyses
- Recordings before, during, and after surgery

# Thank you for your attention!