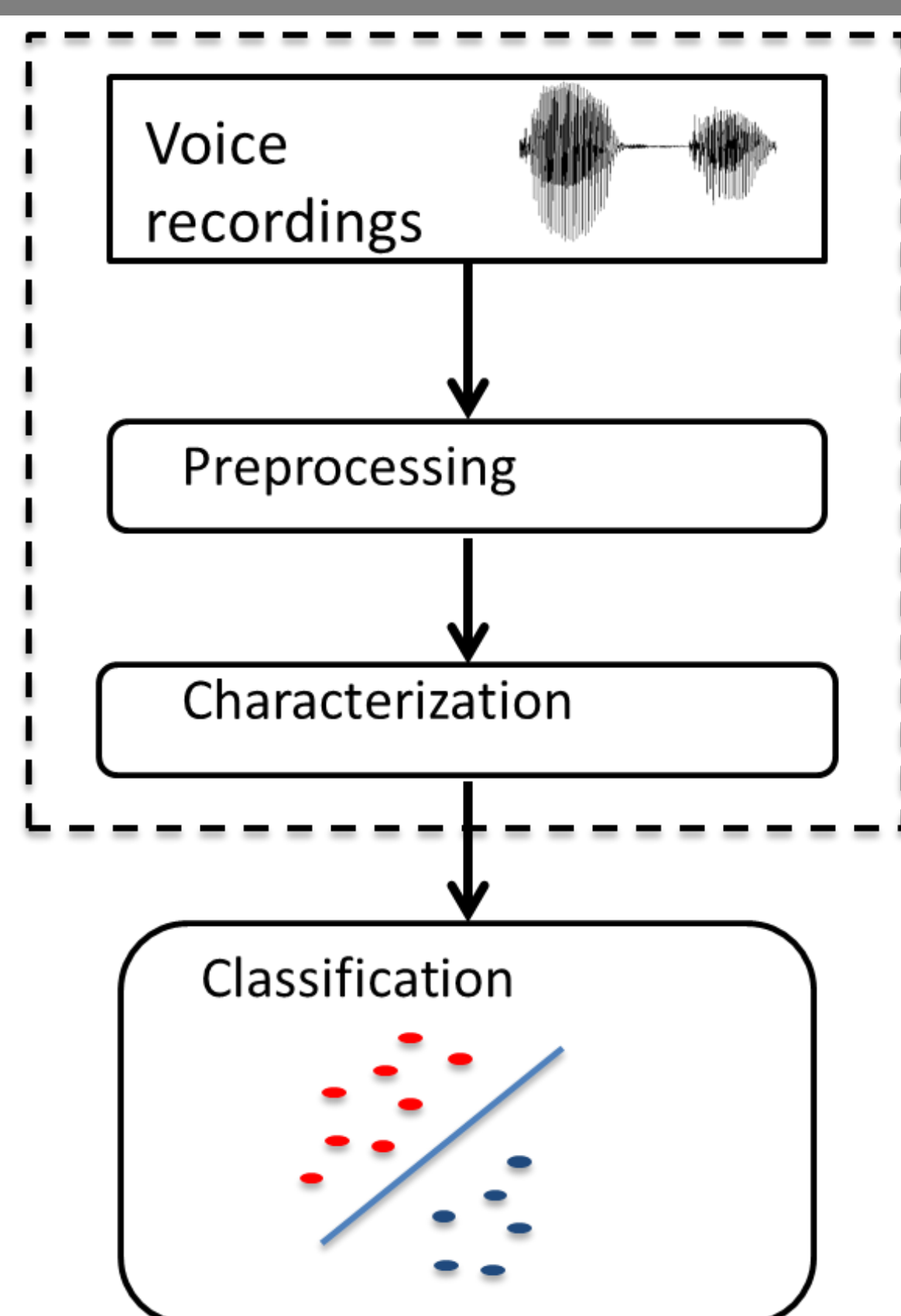


Abstract

About 90% of the people with Parkinson's disease (PD) develop speech impairments such as reduced loudness, monopitch, monoloudness, reduced stress, breathy, hoarse voice quality, and imprecise articulation. Such impairments affect different characteristics of speech including phonation, articulation, prosody, and intelligibility. Phonation problems of people with PD are related with vocal fold bowing and incomplete closing of vocal folds; articulation deficits are manifested as reduced amplitude and velocity of the articulators (lip, tongue and jaw) movements; prosody impairments reflect changes in speech rate and pause characteristics of people with PD. There are several studies addressing the problem of the automatic detection of PD from speech signals in order to develop computer aided tools for the assessment and monitoring of the patients. Recent works have shown that it is possible to detect PD from speech with accuracies above 90%; however, it is still unclear whether it is possible to make the detection independent of the spoken language. A methodology for automatic detection of PD considering speech recordings of words and sentences spoken in three languages (German, Spanish and Czech) is presented here. According to the results it is possible to classify between speech of people with PD and healthy controls (HC) with accuracies ranging from 84% to 99%, depending on the utterance.

Methodology



Validation strategy

10-folds Cross Validation: Spanish and German
One-Speaker-Out: Czech

Classifier

soft margin support vector machine with margin parameter C and Gaussian kernel with parameter γ .

The parameters are optimized up to powers of ten: grid search with $1 < C < 10^4$ and $1 < \gamma < 10^3$.

Databases (DB)

DB 1 – German

- 178 German native speakers, 88 PD patients and 88 HC.
- PD: 44 men and 44 women; HC: 45 men and 53 women.
- Age of PD: mean 66.5 ± 8.9 ; Age of HC: mean 63.2 ± 13.9 .
- Hoehn & Yahr: mean 2.4 ± 0.6 ; UPDRS-III: 22.7 ± 10.9
- Recorded in ON-State, $F_s=16\text{KHz}$, 16 bits.

DB 2 – Spanish

- 100 Spanish native speakers, 50 PD patients and 50 HC.
- PD: 25 men and 25 women; HC: 25 men and 25 women.
- Age of PD: mean 61.14 ± 9.6 ; Age of HC: mean 60.9 ± 9.5
- Hoehn & Yahr: mean 2.29 ± 0.8 ; UPDRS-III: 36.7 ± 18.7
- Recorded in ON-State, $F_s=44.1\text{KHz}$, 16 bits.

DB 3 – Czech

- 42 Czech native speakers, 21 PD patients and 21 HC.
- PD: 17 men and 4 women; HC: 14 men and 7 women.
- Age of PD: mean 62.2 ± 11 ; Age of HC: mean 57.2 ± 13
- Hoehn & Yahr: mean 2.2 ± 0.5 ; UPDRS-III: mean 17.9 ± 7.4
- Recorded without previous medication, $F_s=48\text{KHz}$, 16 bits.

Experiments and results

German: 4 words, 4 sentences and rapid repetition of /pa/-/ta/-/ka/

Utterance	Acc V --- Acc UV	AUROC: V --- UV
Bahnhofsvorsteher	73% --- 97%	0.72 --- 0.97
Rettungsschwimmer	69% --- 96%	0.66 --- 0.96
Toilettenpapier	71% --- 95%	0.70 --- 0.95
Bundesgerichtshof	62% --- 94%	0.65 --- 0.95
S1	73% --- 97%	0.77 --- 0.97
S2	77% --- 96%	0.75 --- 0.96
S3	78% --- 94%	0.77 --- 0.95
S4	72% --- 93%	0.70 --- 0.94
/pa/-/ta/-/ka/	70% --- 98%	0.68 --- 0.98

S1: *Im Inhaltsverzeichnis stand nichts über Lindenblütentee*

S2: *Seit seiner Hochzeit hat er sich sehr verändert*

S3: *Der Kerzenständer fiel gemeinsam mit der Blumenvase auf den Plattenspieler*

S4: *Peter und Paul essen gerne Pudding*

Spanish: 5 words, 1 sentence and rapid repetition of /pa/-/ta/-/ka/

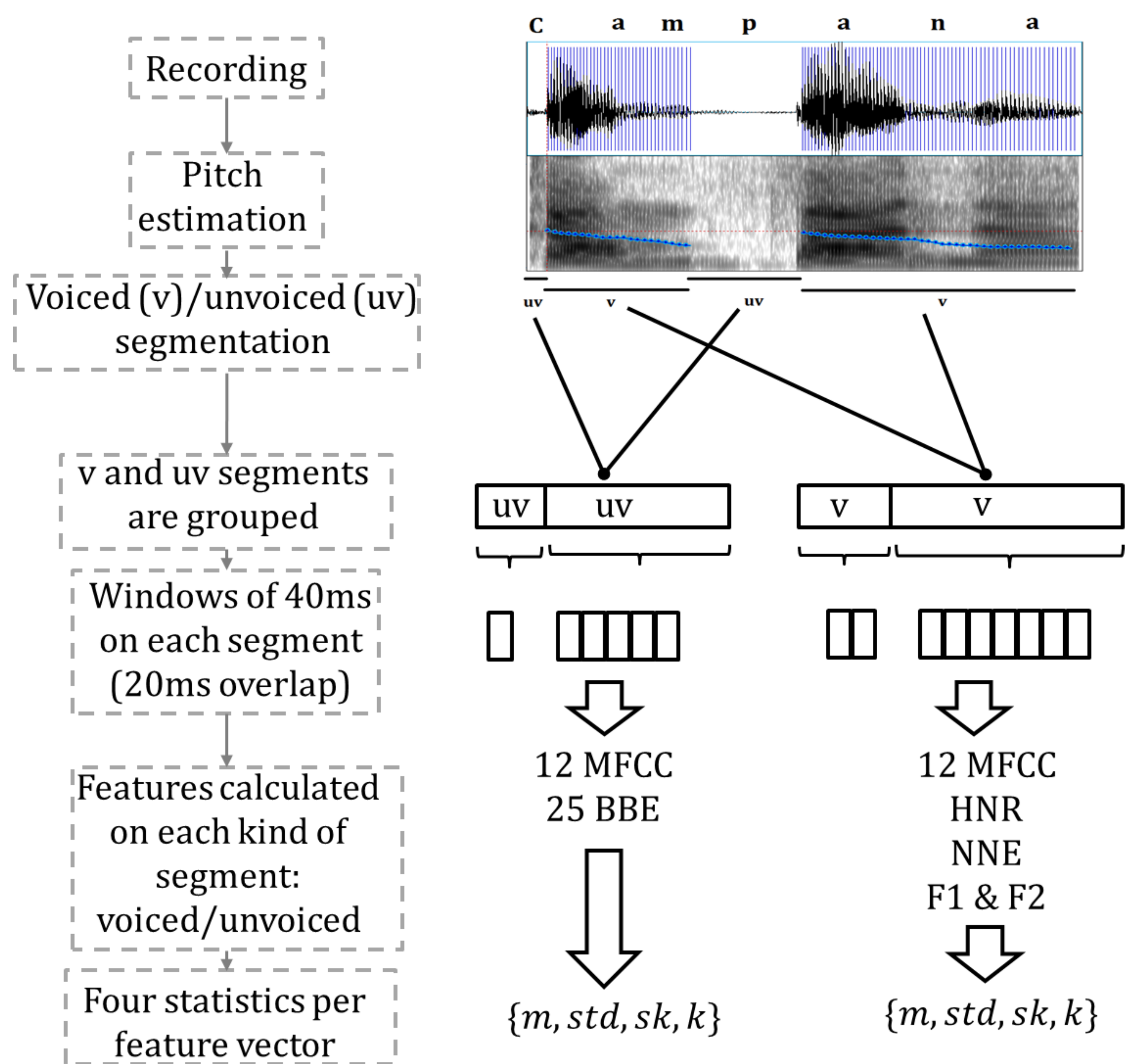
Utterance	Acc V --- Acc UV	AUROC: V --- UV
Atleta	82% --- 99%	0.79 --- 0.99
Campana	73% --- 99%	0.76 --- 0.99
Gato	76% --- 98%	0.76 --- 0.98
Petaka	84% --- 97%	0.82 --- 0.98
Braso	75% --- 96%	0.74 --- 0.98
Sentence	78% --- 90%	0.82 --- 0.93
/pa/-/ta/-/ka/	80% --- 99%	0.82 --- 0.99

Sentence: *Mi casa tiene tres cuartos*

Czech: rapid repetition of /pa/-/ta/-/ka/

Utterance	Acc V --- Acc UV	AUROC: V --- UV
/pa/-/ta/-/ka/	90% --- 97%	0.94 --- 0.99

Preprocessing and Characterization



Discussion

- The information from unvoiced segments had been used previously only to detect onset and offset of syllables and words.

- Unvoiced segments provide information related to the operation of the abductor in vocal folds and they are affected when the people has spasms like in PD.

- The information provided by unvoiced segments can be used to classify between people with PD and healthy controls in different languages with databases recorded in different conditions

- Although voiced segments provide useful information about the speech production process, according to the method presented in this work, it is better to use unvoiced segments to discriminate between the speech of people with PD and healthy controls

Acknowledgments

Juan Rafael Orozco Arroyave is under grants of COLCIENCIAS through the call N° 528 «Doctorados Nacionales»