LABORATOIRE DE PHONETIQUE ET PHONOLOGIE

FNSNF
FONDS NATIONAL SUISSE
DE LA RECHERCHE SCIENTIFIQUE

Using spectral features to distinguish speakers with Parkinson's disease from typical speakers.

Slis, A.[1], Fougeron, C.[1], Pernon, M.[1,2] & Lancia, L.[1]

[1]LPP, UMR 7018, CNRS/U. Sorbonne-Nouvelle, Paris, France.
[2]Department of Clinical Neurosciences, Geneva University Hospital, Switzerland.

Background

In most cases, speakers with Parkinson's disease develop Hypokinetic dysarthria[1]:

- slower and smaller articulatory movements with reduced peak velocities of articulators[2]
- reduced F2 transitions[2]
- Smaller vowel space[3]

Challenge acoustic studies: Difficulty separating mild from moderate dysarthric speakers. Isolating formant curves is time consuming, susceptible to error, and ignoring important acoustic information caused by changes in vocal tract shape.

Our study: complete shape of the spectrum envelope and its patterns of change over time by extracting Mel-Frequency Cepstrum Coefficients (MFCC), which represent the changes in vocal tract shape[4].

Hypothesis: analyzing the acoustic signal in terms of Mel Frequency Cepstral Components provides sufficient spectral information to separate the speech productions of people diagnosed with Parkinson's Disease (PD) from typical speakers.

Participants: speakers with mild and moderate dysarthria and control speakers produced vowel glide and glide vowel sequences (/ajajaj/, /ujujuj/ and /wiwiwi/ 4 times in separate trials and recorded.

MFCC

- power spectra were calculated from the acoustic waveform on 25 ms windows, with 2 ms steps.
- a MEL filter bank was applied to these power spectra and the summed spectral energy in each band was calculated.
- Log-transform filterbank energies.
- Discrete Cosine Trans- formation (DCT).
- DCT coefficients 2-13 were considered.
- squared differences of cepstral energy values in consecutive frames were computed and summed at each time step (ASC_MFCC; see figure), resulting in a contour.

METRICS

MASC1: peak value of ASC_MFCCs from minimum ASC_MFCC (the steady state during the vowel) to the next maximum ASC_MFCC (the point of highest spectral energy change, i.e., during the transition to the glide), the peak value of ASC_MFCCs (MASC1) was computed.

MASC2: maximum to the next minimum (MASC2).

DASC: duration of the vowel-glide gesture from minimum to minimum (i.e., from steady state vowel to steady state glide

AASC: the derivative from the ASC_MFCC curve (acceleration of the average squared change of the MFCCs) was calculated. From this curve, the time from a minimum ASC_MFCC to a maximum acceleration peak (AASC1) and time from the maximum acceleration to a minimum ASC_MFCC valley (AASC2) were computed.

Results

MASC1:
/wiwiwi/ showed decreasing MASC1 values in the order: healthy>PD_MILD > PD_moderate.
/ajajaj/ and /ujujuj/: moderate dysarthric PD differed from healthy speakers.

DASC:
No significant differences

AARC1:
No difference between the typical and the dysarthric speakers.

Conclusions

1. Using ASC_MFFCs as a tool to distinguishing typical from atypical PD speech is promising.
2. The lower values for MASC1s characterizing the PD speakers suggest that the vocal tract changes are slower, especially for /wiwiwi/, which is consistent with the slower and smaller articulatory movements and reduced peak velocities of articulators, found in earlier kinematic studies[2] as well as the reduced F2 slopes and vowel space[2,3].

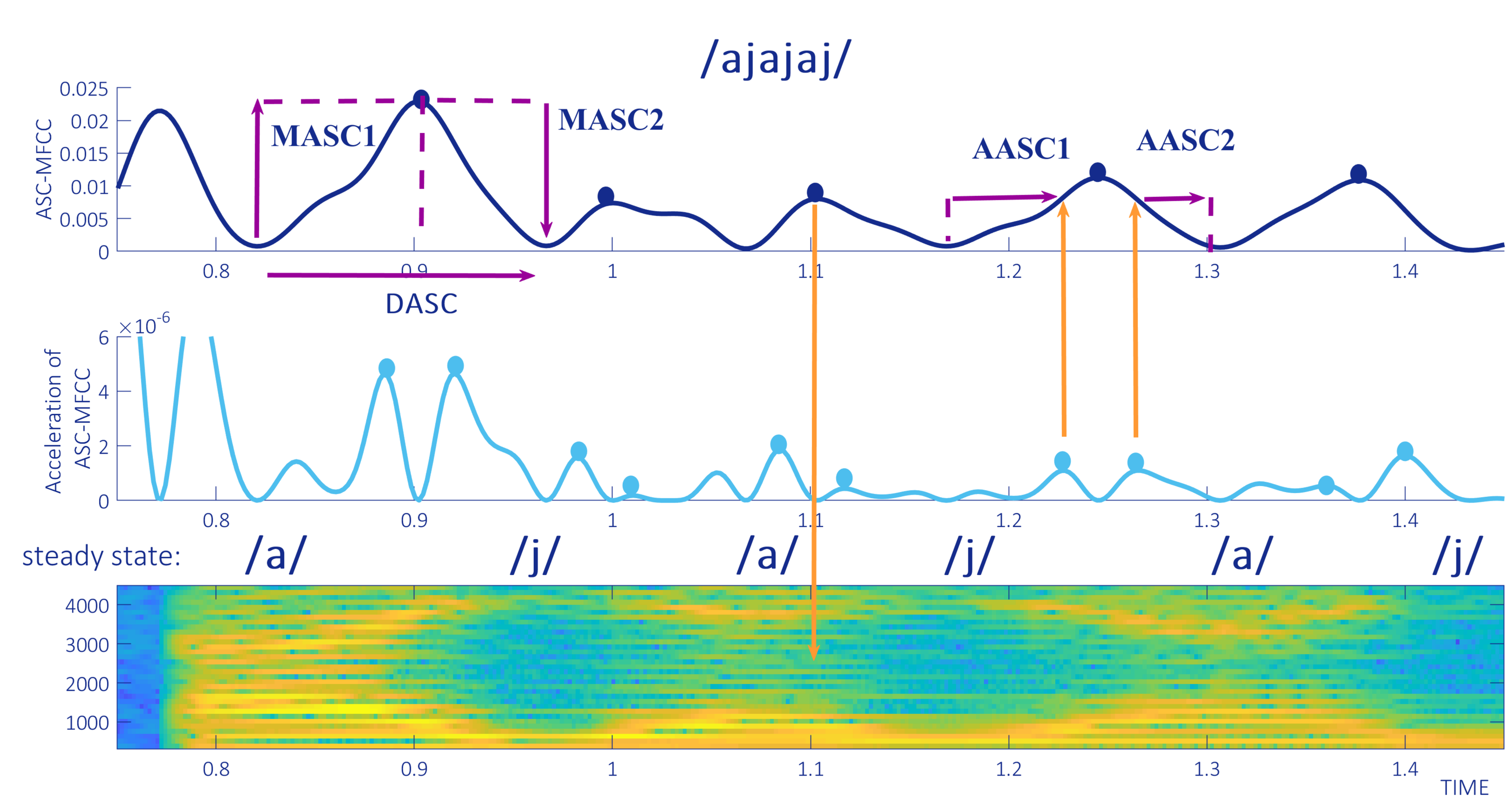
It is speculated that PD speakers likely face more difficulties in changing the shape of the vocal tract to produce the contrasts of /wiwiwi/, which showed differences in the mild and moderate cases. Individuals with PD have been shown to experience perioral stiffness[5], which potentially explains these results.

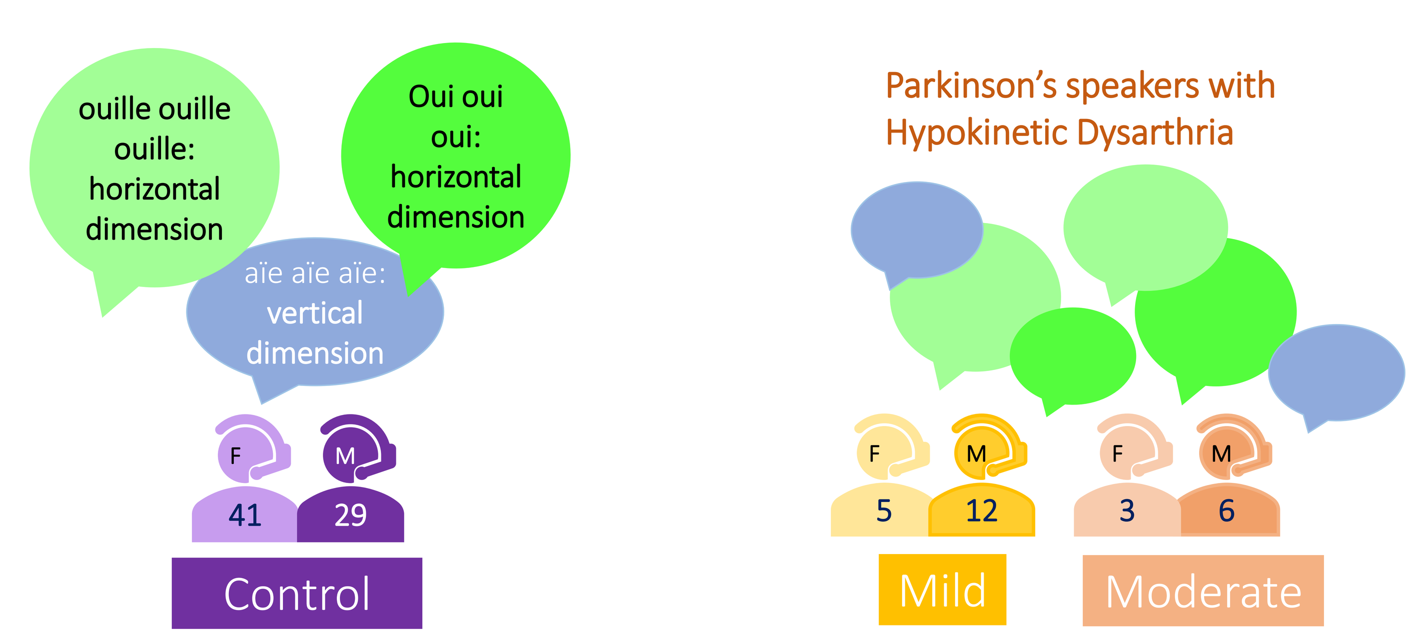
Question

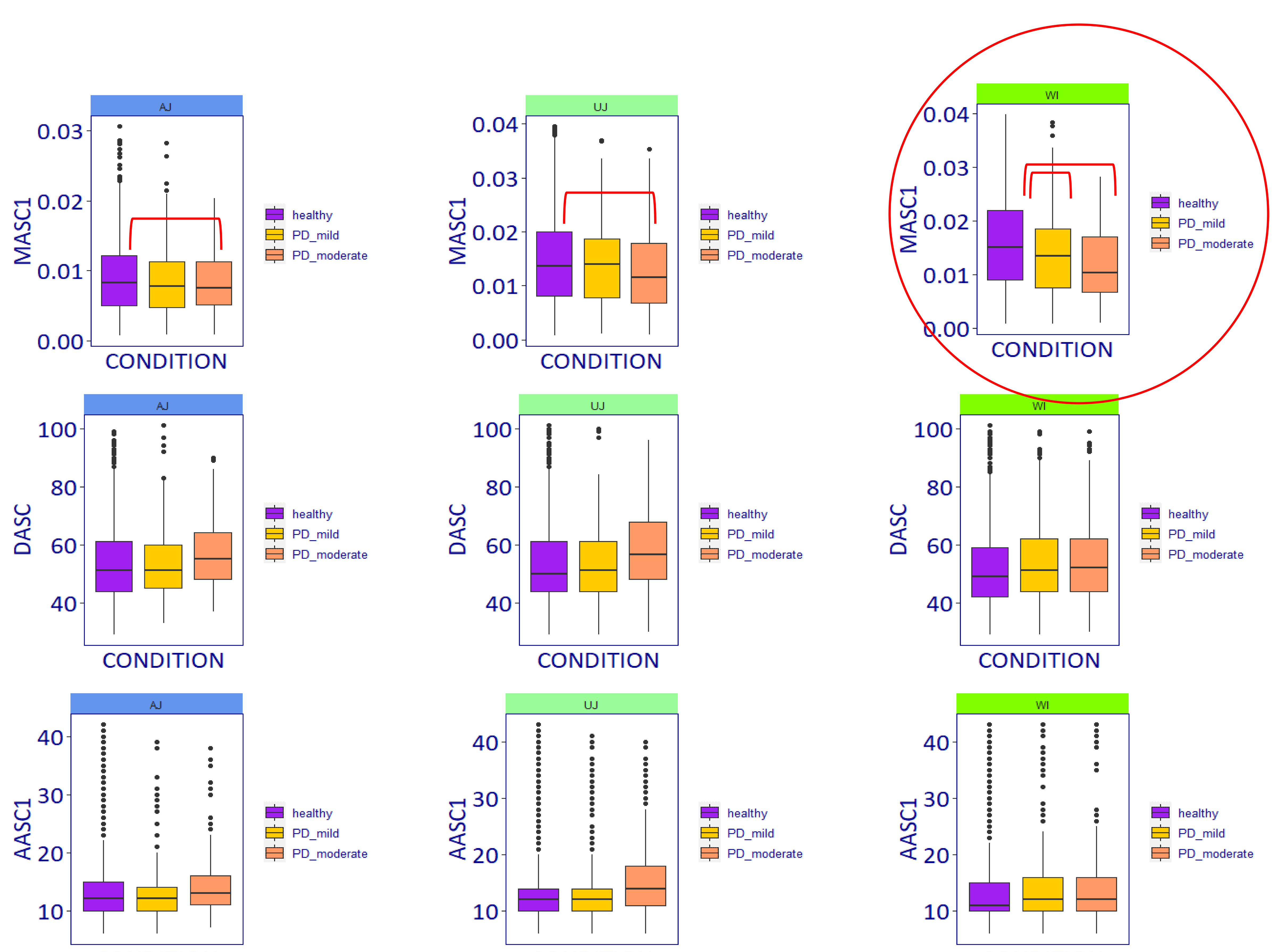
Can we separate dysarthric speakers with Parkinson's disease (PD) from typical speakers with an acoustic measure?

Time course of Mel-Frequency Cepstral Coefficients (MFCC)

$$\text{MFCC}_k = \sum_{i=2}^{13}\left(f(i,k+1) - f(i,k)\right)^2$$







1) Duffy, J. R. (2005). *Motor speech disorders: Substrates, differential diagnosis, and management.* St. Louis, Mo: Elsevier Mosby.
2) Forrest, K., Weismer, G., & Turner, G. S. (1989). Kinematic, acoustic, and perceptual analyses of connected speech produced by Parkinsonian and normal geriatric adults. The Journal of the Acoustical Society of America, 85(6), 2608–2622. https://doi.org/10.1121/1.397755
3) Whitfield, J. A., & Goberman, A. M. (2014). Articulatory–acoustic vowel space: Application to clear speech in individuals with Parkinson's disease. Journal of Communication Disorders, 51, 19–28. https://doi.org/10.1016/j.jcomdis.2014.06.005
4) Goldstein, L. (2019). The role of temporal modulation in sensorimotor interaction. Frontiers in Psychology, 10, 2608. https://doi.org/10.3389/fpsyg.2019.02608
5) Caligiuri, M. P. (1987). Labial kinematics during speech in patients with parkinsonian rigidity. Brain: A Journal of Neurology, 110(4), 1033–1044. https://doi.org/10.1093/brain/110.4.1033