



# DÉTECTION PRÉCOCE DE LA MALADIE DE PARKINSON PAR L'ANALYSE DE LA VOIX, ET CORRÉLATIONS AVEC LA NEUROIMAGERIE

Présenté par **Laetitia JEANCOLAS**

Sous l'encadrement de :

- **Dijana Petrovska-Delacrétaz**, Télécom SudParis, France
- **Badr-Eddine Benkelfat**, Télécom SudParis, France
- **Habib Benali**, Université Concordia, Montréal, Canada



# SOMMAIRE

1. Introduction
2. Constitution de nos bases de données
3. Méthodes proposées pour la détection précoce de MP
  - 3.1 Analyse MFCC-GMM
  - 3.2 Analyse x-vecteurs
  - 3.3 Analyse paramètres globaux
4. Corrélations avec la neuroimagerie
5. Conclusions et perspectives

# INTRODUCTION

- Maladie neurodégénérative : 2<sup>e</sup> plus répandue après Alzheimer (1,5% > 65 ans)
- Diagnostic clinique : médecin recherche la présence de symptômes moteurs
- Apparition quand > 60 % des neurones dopaminergiques de la substance noire détruits

➡ **Trouver une méthode de diagnostic plus précoce**

- **Perturbations vocales:** un des premiers symptômes
- Domaines touchés:
  - Phonation
  - Articulation
  - Prosodie
  - Rythme



- **Détection de MP avec la voix:**

Taux de bonnes classifications (Acc)  $\in$  [70 - 99%]

- **Détection de MP débutante** (< 4 ans après diagnostic)

Peu d'études (Novotny 2014, Ruzs 2015, Orozco 2016)

- **Stade préclinique** (avant diagnostic clinique):

Présence de perturbations vocales (Postuma 2012, Ruzs 2015, Hlavnicka 2017, Benba 2019)

- **Télédiagnostic**

Applications smartphone ou tablette (Zhang 2018, Ruzs 2018)

Simulation du réseau téléphonique (Tsanas 2012, Vasquez-Correa 2017)

**Nouveau:** Appels téléphoniques réels (Jeancolas 2019)



- Constitution de **grandes bases de données** (> 200 sujets)



- Caractéristiques de la voix au **stade débutant** et **prodromique**

- Détection **automatique** de MP

- Effet du **genre**
- Effet du **microphone**
- Comparaison des **tâches vocales**



- Suivi de l'évolution et **Corrélations avec la neuroimagerie**

- **Objectif à long terme:**
  - outil de diagnostic précoce de MP utilisable par les médecins
  - outil d'aide au diagnostic précoce (et suivi) par téléphone

# CONSTITUTION DE NOS BASES DE DONNEES

- **Protocole ICEBERG : Suivi sur 5 ans** avec examens annuels à l'hôpital Pitié Salpêtrière

- **Participants: 252** sujets français
  - 115 MP débutants (74 ♂ and 41 ♀)
  - 46 iRBD (Rapid eye movement Behaviour Disorder) (41 ♂ et 5 ♀)
  - 91 sains (48 ♂ and 43 ♀)

- **Enregistrement vocaux annuels** (1 session par sujet)



microphone serre tête **professionnel**  
carte son externe - préamplification  
5cm de la bouche



microphone de l'**ordinateur**  
préamplification - débruitage actif  
50cm de la bouche

(96 kHz, 24 bits)  
étendue spectrale = [50Hz, 20kHz]

- **Enregistrements téléphoniques mensuels** (~ 5 sessions par sujet)



appels avec les téléphones des participants → répondeur interactif

(8 kHz, 16 bits)  
étendue spectrale = [300Hz, 3400Hz]

- **Tâches** (~10min) : voyelles soutenues, répétitions rapides ou lentes de syllabes, monologue, (lecture et répétitions de phrases)



# METHODES PROPOSEES POUR LA DETECTION PRECOCE DE MP

## ■ Méthode MFCC-GMM, classique en reconnaissance du locuteur :

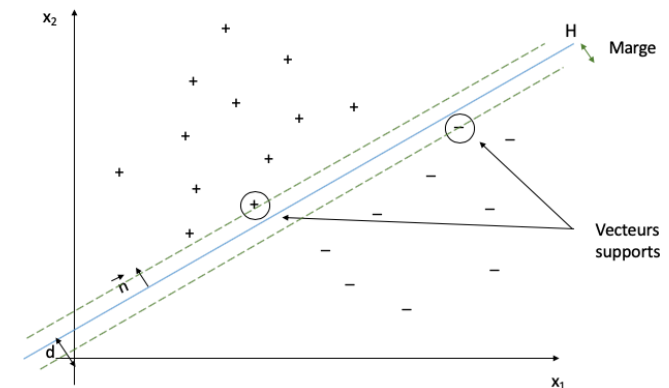
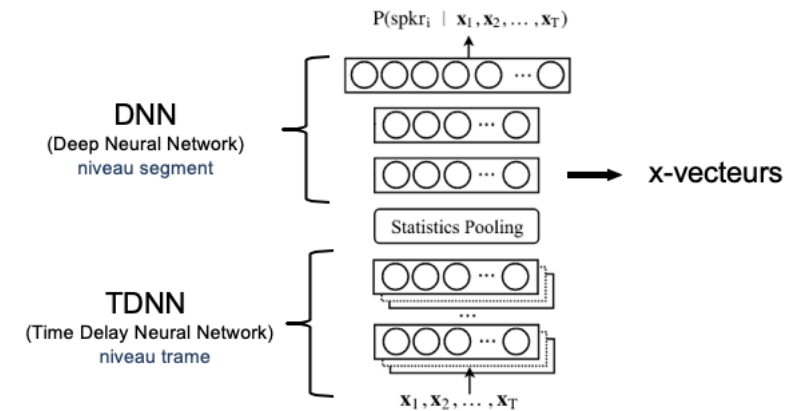
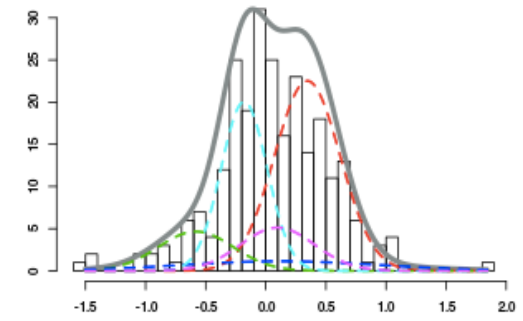
- extraction de paramètres court-termes : MFCC (Mel-Frequency Cepstral Coefficient)
- modélisation GMM (Modèle de Mélange Gaussien) pour chaque groupe
- classification à l'échelle de la trame (20ms)

## ■ Méthode x-vecteurs, nouveau en reconnaissance du locuteur :

- extraction des MFCC
- entraînement DNN (Deep Neural Network) et extraction des x-vecteurs
- classification à l'échelle du segment (3s)

## ■ Méthode des paramètres globaux :

- extraction de paramètres « globaux »
- analyses de variances
- classification SVM (Support Vector Machine) à l'échelle de la tâche (~30s)



## Détection MP débutant vs sain

Taux d'égale erreur (EER)  $EER \leftrightarrow \text{Faux positifs} = \text{Faux négatifs}$



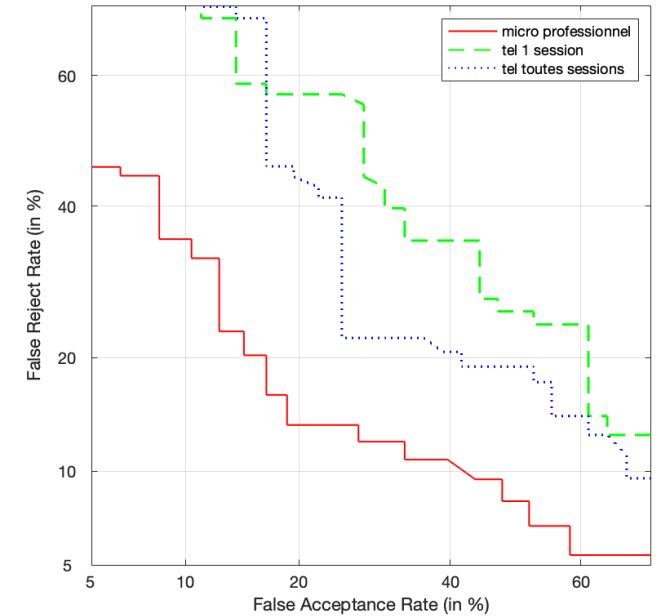
- EER = **17 %** avec 1min30 lect + repet + /pataka/ 
- EER = **25 %** avec 1min10 lect + repet  et « cross micro »
- EER = **25 %** avec 5min DDK 



- EER ~ 40% → *interprétations* : + grande variabilité des MFCC (Fraile 2009)  
atteinte neuronale moins marquée (Haaxma 2007)  
circuits neuronaux parole différents (Jung 2019)



Détection iRBD vs sain : EER = 37 %  
Détection iRBD moteur<sup>+</sup> vs sain : → **EER = 28 %**



- Influence de la durée des segments
- Influence de l'augmentation des données
- EER distance cos > (LDA + distance cos) ~ PLDA
- Comparaison avec MFCC-GMM



- Légère amélioration pour monologue



- **amélioration ~ 10 %** pour monologue

→ EER=30%

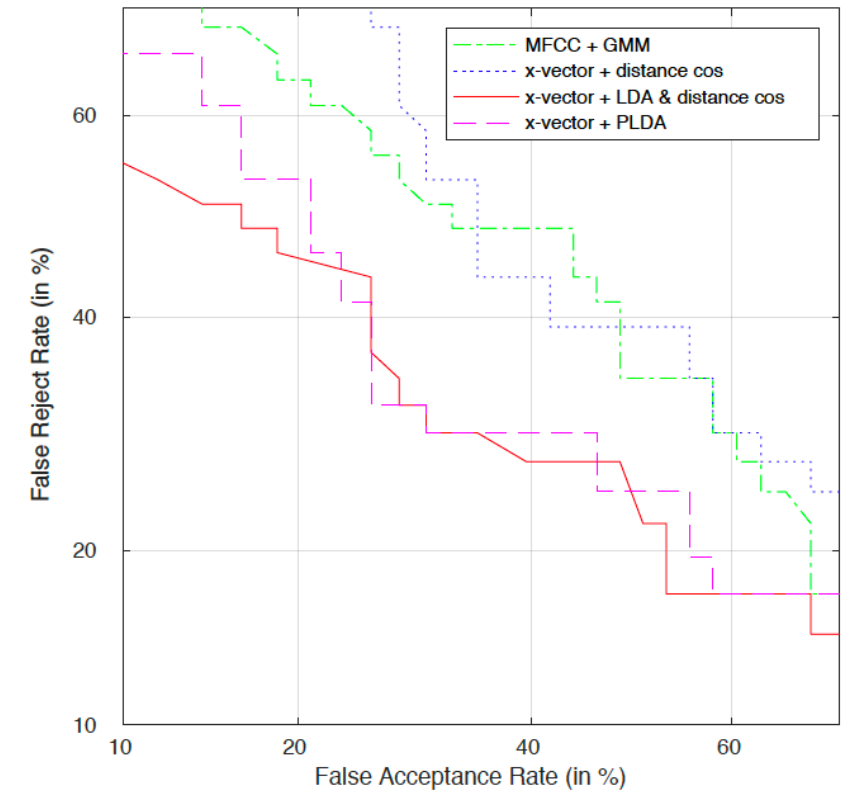


Fig: DET lors du monologue



- Paramètres les plus discriminants:
  - **SD log fo** → monotonie
  - **Nb pauses courtes et durée médiane des pauses** → débit saccadé
  - **RSD et Ecart absolu rythme** → difficulté à suivre un rythme constant et imposé
- Atteinte dès le stade préclinique (iRBD)
- Hommes plus atteints que femmes
- Classification avec SVM

 EER = 22%  
 EER = 31%



EER = 27%  
EER = 33%



Fusion par vote majoritaire (aux seuils EER) des 3 méthodes de classifications

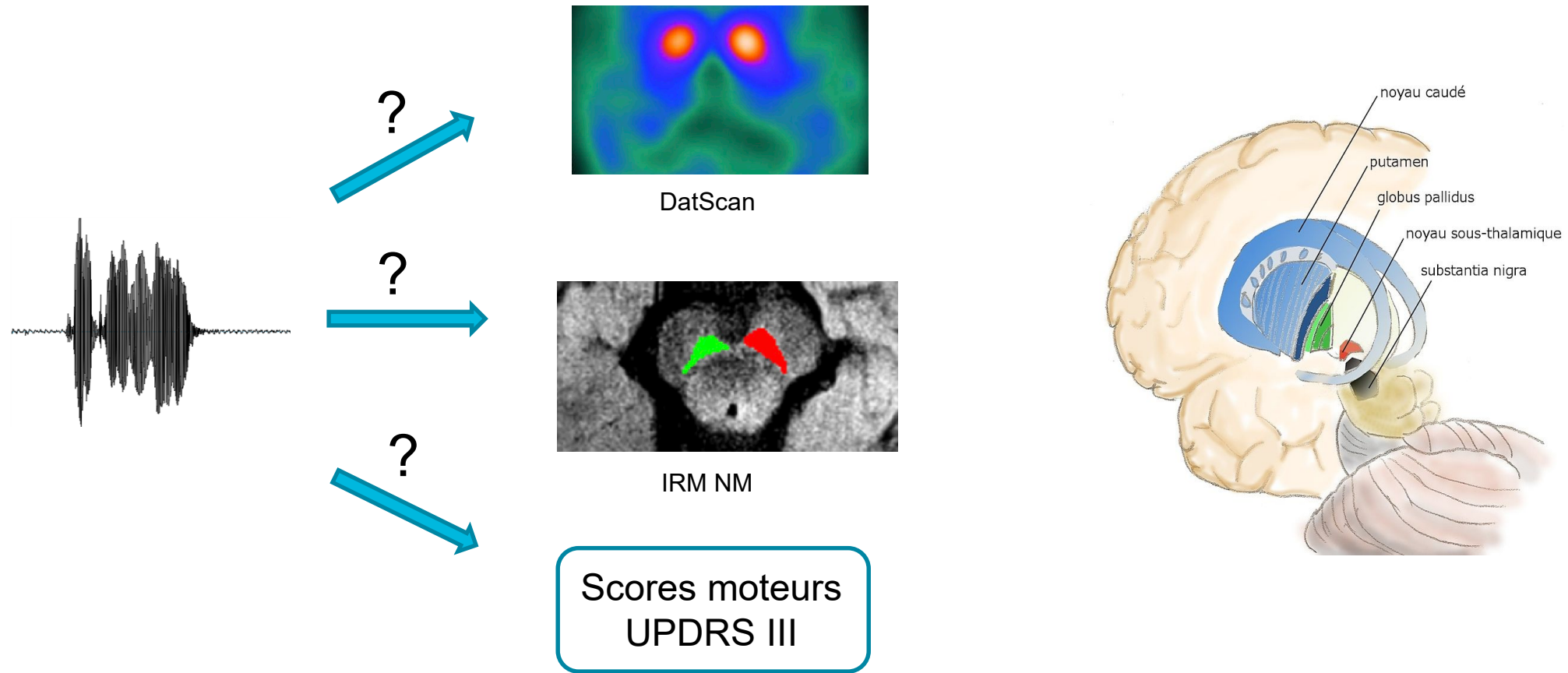


| MFCC-GMM         | X-vecteurs<br>LDA + dist cos | Param globaux<br>SVM | Fusion                            |
|------------------|------------------------------|----------------------|-----------------------------------|
| Acc = 83%        | Acc = 78%                    | Acc = 78%            | <b>Acc = 89%</b><br>Sp=92% Se=87% |
| <b>Acc = 75%</b> | Acc = 71%                    | Acc = 73%            | Acc = 75%<br>Sp=78% Se=73%        |
| Acc = 60%        | <b>Acc = 70%</b>             | Acc = 68%            | Acc = 69%<br>Sp=67% Se=71%        |

$Sp = 1 - \text{Faux positif} = \text{taux de bonnes classifications parmi les sujets sains}$

$Se = 1 - \text{Faux négatif} = \text{taux de bonnes classifications parmi les sujets MP}$

# CORRELATIONS AVEC LA NEUROIMAGERIE



**Modèle de régression linéaire multiple:**  $Y_i = a_0 + a_1X_{i1} + a_2X_{i2} + \dots + a_pX_{ip} + \varepsilon_i$



| Variables résultats           | Examen        | RMSE   | R2    | F    | valeur-p |
|-------------------------------|---------------|--------|-------|------|----------|
| UPDRS III OFF                 | tests moteurs | 0.215  | 0.215 | 4.42 | 9.56E-5  |
| Transporteurs dopaminergiques | DatScan       | 1.34   | 0.422 | 4.47 | 4.06E-4  |
| Volume substance noire        | IRM NM        | 5.2E-5 | 0.191 | 3.11 | 3.32E-3  |

*R2 : coefficient de détermination (proportion de la variance prédite)*

*Valeur-p. : mesure de significativité (significatif quand  $p < 0.05$ )*

- 42 % de la variance du DatScan peut être expliquée par une combinaison linéaire des 7 paramètres voix et de l'information du genre
- **Ces paramètres voix prédisent linéairement de manière significatives les données de DatScan, de IRM NM et l'UPDRS III**

# CONCLUSIONS ET PERSPECTIVES

- Constitution de **3 grandes bases de données** (> 200 sujets)



- Identification de **marqueurs de détection précoce**, dès le stade **préclinique**:

- Articulation
- Monotonie
- Débit saccadé
- Difficulté à suivre un rythme (constant et imposé)

- **Détection de MP débutant** par 3 méthodes:
  - MFCC-GMM
  - **x-vecteurs**
  - paramètres globaux

} fusion

|   |  |               |         |            |  |                          |            |     |  |        |          |            |
|---|------------------------------------------------------------------------------------|---------------|---------|------------|--------------------------------------------------------------------------------------|--------------------------|------------|-----|--------------------------------------------------------------------------------------|--------|----------|------------|
|   | durée                                                                              | tâches        | méthode | Acc        | durée                                                                                | tâches                   | méthode    | Acc | durée                                                                                | tâches | méthode  | Acc        |
| ♂ | 6min                                                                               | tout sauf /a/ | fusion  | <b>89%</b> | 4min                                                                                 | monol-lec-rythme globale | <b>78%</b> |     | 5min                                                                                 | DDK    | MFCC-GMM | <b>75%</b> |
| ♀ | 1min                                                                               | monol         | x-vect  | <b>70%</b> | 4min                                                                                 | monol-lec-rythme globale | <b>70%</b> |     | 5min                                                                                 | monol  | x-vect   | <b>67%</b> |

- Performances   meilleures que l'état de l'art
- **Nouveau:** détection de MP débutant par  → **télédiagnostic de MP**
- Langue française
- Effet du genre:  MP voix plus altérée que 
- Modèle prédictif des données de neuroimagerie et de clinique → **suivi de l'évolution de MP**

- ❖ Phase de développement (optimisation hyperparamètres, seuil EER)  
→ validation sur nouvel ensemble de sujets
- ❖ Patients ON : atténuation de certaines modifications vocales  
→ ajustements en OFF
- ❖ ♀ et ♂ appariés en âge → tester avec UPDRS apparié
- ❖ **Pistes d'améliorations :**
  - autres types de fusion
  - extractions d'autres paramètres vocaux
  - classifieurs spécialisés dans tâches dépendantes du texte
  - corrélations :
    - ajouts paramètres globaux supplémentaires
    - différentes régions du cerveau
    - modèles non linéaires

***Merci de votre attention !***

Contact : [laetitia.jeancolas@telecom-sudparis.eu](mailto:laetitia.jeancolas@telecom-sudparis.eu)