



Représentation automatique du rythme de la parole pathologique via le spectre de modulations d'amplitude

Robin Vaysse, Corine Astésano, Jérôme Farinas

17 juin 2023

Comment caractériser automatiquement le rythme de la parole pathologique ?

- Quels outils automatiques pour caractériser le rythme de la parole ?
- Comment se comporte le rythme de personnes atteintes de pathologies impactant la voix ?
 - Quels impacts sur le rythme des patients ?
 - **Compensent-ils leurs déficits** par une prosodie particulière ?

Structuration de la parole via :

- Matérialité **phonétique**
 - Fréquence fondamentale (F0)
 - Intensité
 - Durée
- Matérialité **phonologique**
 - Structuration en **unités hiérarchiques**

Structuration prosodique



Niveaux de structuration hiérarchique :

- Le syntagme intonatif (IP, *Intonational phrase*)
-
-
-



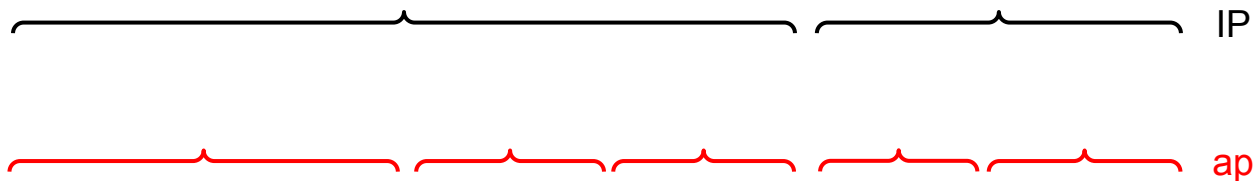
L'ordinateur portable de Gabriel est cassé. Il faudra le changer.

Structuration prosodique



Niveaux de structuration hiérarchique :

- Le syntagme intonatif (IP, *Intonational phrase*)
-
- **Le syntagme accentuel (ap, *accentual phrase*)** (Jun & Fougeron, 2000)
-



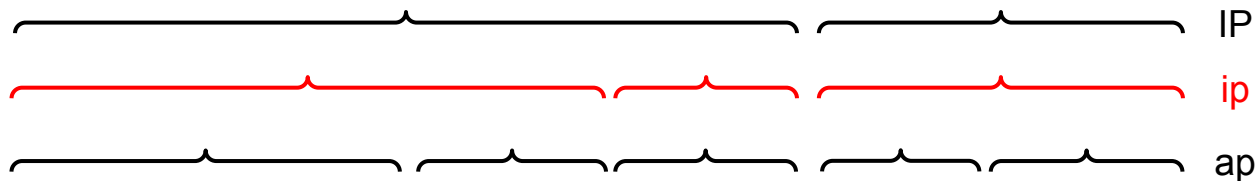
L'ordinateur portable de Gabriel est cassé. Il faudra le changer.

Structuration prosodique



Niveaux de structuration hiérarchique :

- Le syntagme intonatif (IP, *Intonational phrase*)
- **Le syntagme intermédiaire (ip, *intermediate phrase*)** (Michélas & D'Imperio, 2010)
- Le syntagme accentuel (ap, *accentual phrase*) (Jun & Fougeron, 2000)
-



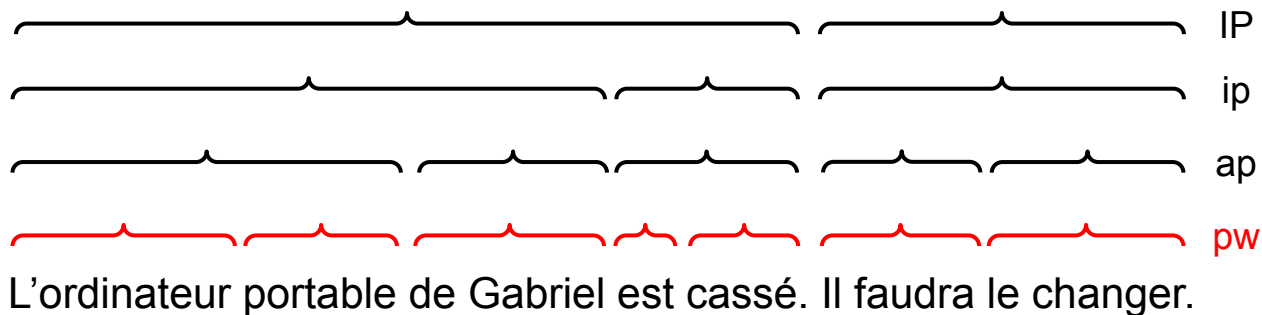
L'ordinateur portable de Gabriel est cassé. Il faudra le changer.

Structuration prosodique



Niveaux de structuration hiérarchique :

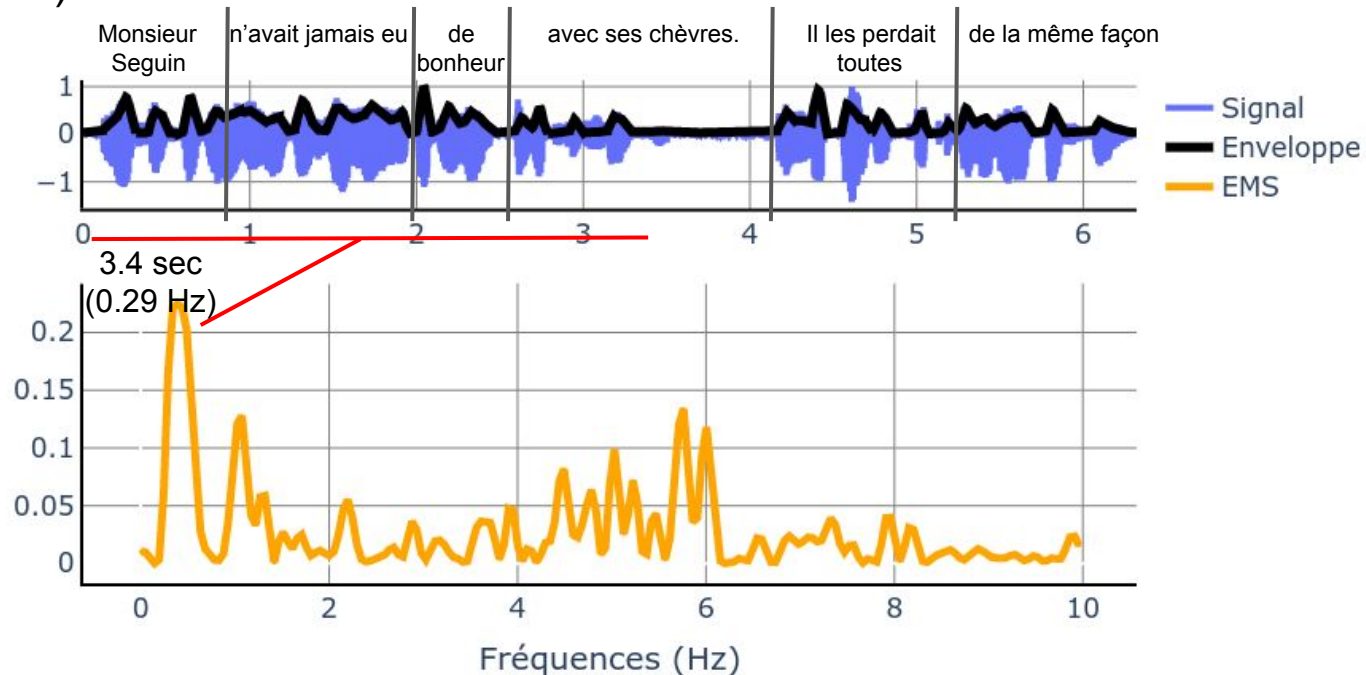
- Le syntagme intonatif (IP, *Intonational phrase*)
- Le syntagme intermédiaire (ip, *intermediate phrase*) (Michelas & D'Imperio, 2010)
- Le syntagme accentuel (ap, *accentual phrase*) (Jun & Fougeron, 2000)
- Le mot prosodique (pw, *prosodic word*) (Astésano 2019)



- Historiquement pour l'identification des langues
 - Durées des **intervalles vocaliques** (%V, ΔV) et consonantiques (Ramus et al. 1999)
 - Variabilité des **durées relatives** entre ces intervalles (Ling et al. 2000)
 - rPVI (*raw Pairwise Variability Index*)
 - nPVI (*normalised Pairwise Variability Index*)
- Analyse automatique **linéaire** du rythme
 - Ne représente pas la **hiérarchie**

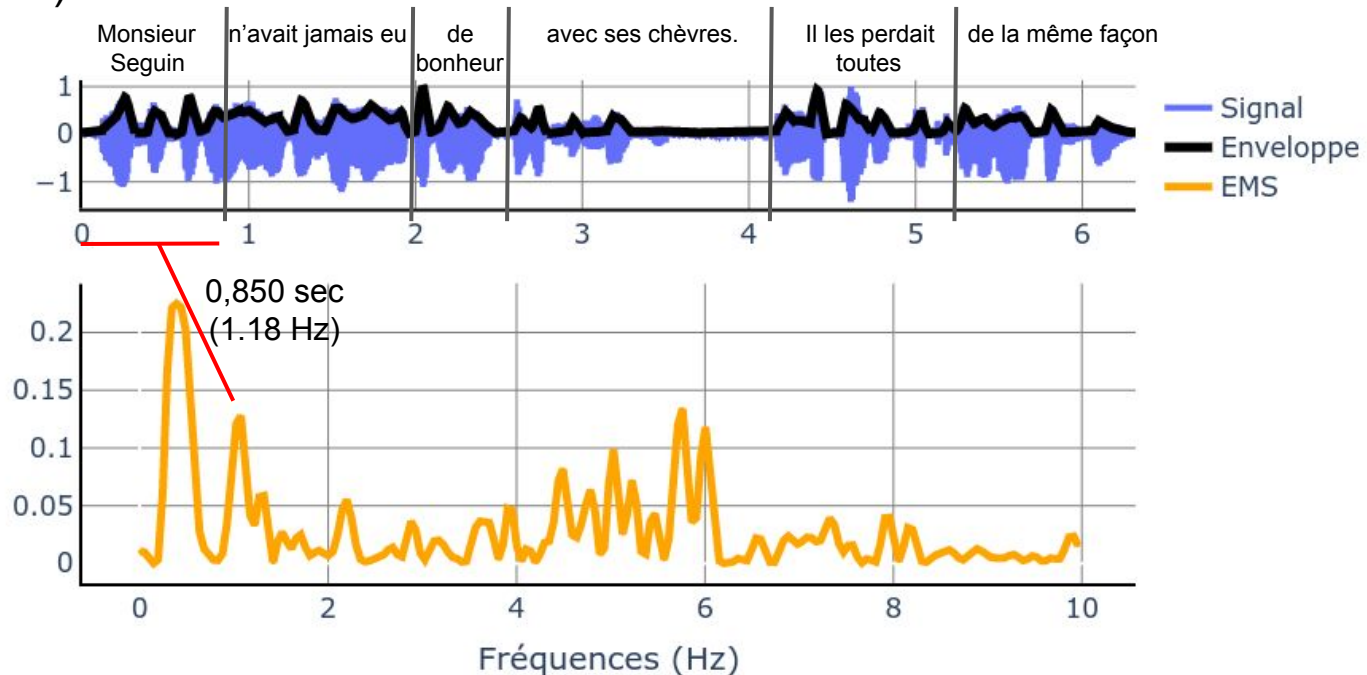
Spectre de modulations d'amplitude

- Extraction de l'**enveloppe d'amplitude** du signal (Tilsen & Johnson 2008)
- Calcul de la **transformée de Fourier** de l'enveloppe (**EMS, Envelope Modulation Spectrum**)

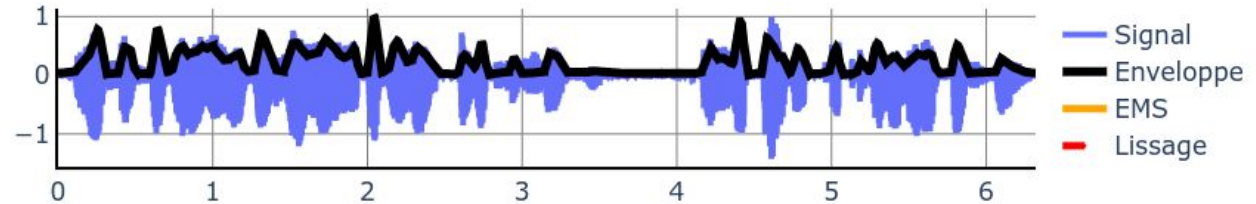


Spectre de modulations d'amplitude

- Extraction de l'**enveloppe d'amplitude** du signal (Tilsen & Johnson 2008)
- Calcul de la **transformée de Fourier** de l'enveloppe (**EMS, Envelope Modulation Spectrum**)

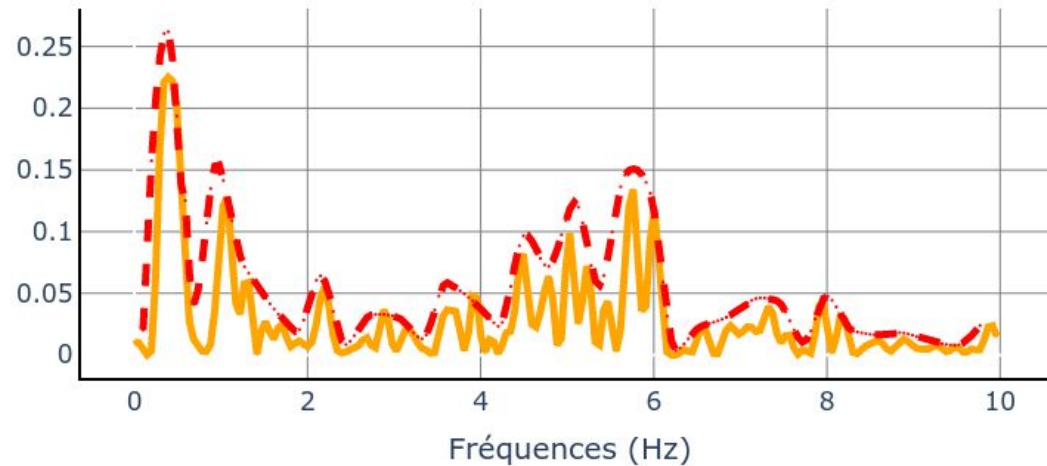


Lissage du spectre



Lissage du spectre

- Focalisation sur les zones de fréquences importantes



- **Projet ANR RUGBI** (*“Relevant linguistic Units to improve the intelligibility measurement of speech production disorders”*)
 - **Complémenter et expliciter la mesure du déficit d’intelligibilité** de la parole pathologique
 - Recherche **automatique d’unités prosodiques** pertinentes
- Corpus de parole de patients atteints de cancers des Voies Aéro Digestives Supérieures (VADS)
 - **87 patients** atteints de cancers VADS
 - **26 sujets contrôles**
 - Annotations cliniques perceptives de l’intelligibilité et de la sévérité du trouble de la parole

Annotation de la prosodie

Sélection de 10 sujets contrôles et 10 patients VADS

- Annotation manuelle des structures prosodiques



m@	sj2	s@	ge~	na	vE	Za	mE	zy	d@	bO	n9	Ra	vEk	se	SEvR
monsieur		seguin		n'avait		jamais		eu	de	bonheur			avec	ses	chèvres
pw		pw		pw		pw		pw		pw			pw		pw
ap			ap					ap			ap				
ip			ip					ip			ip				
IP															
Monsieur seguin n'avait jamais eu de bonheur avec ses chèvres															

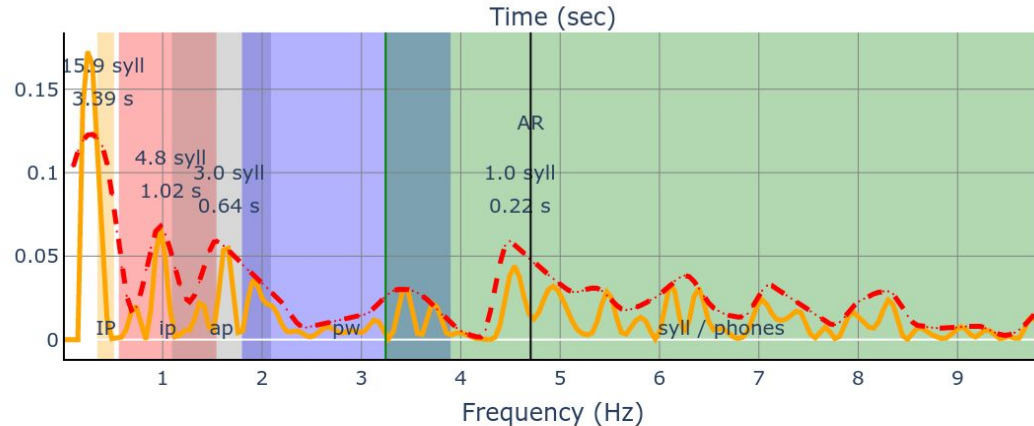
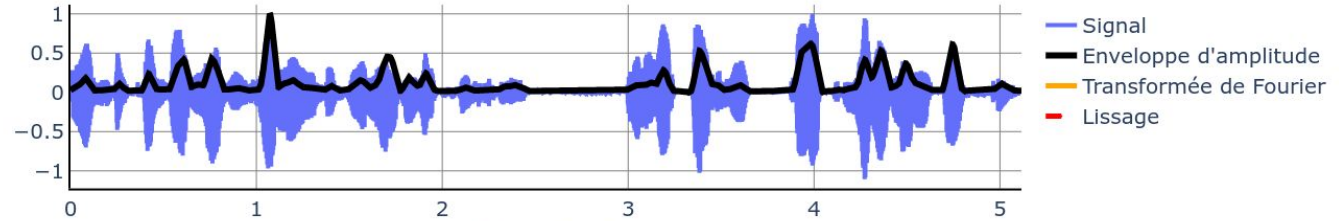
Visible part 3.150349 seconds

EMS de la parole lue

- Superposition de l'EMS aux durées des structures prosodiques annotées

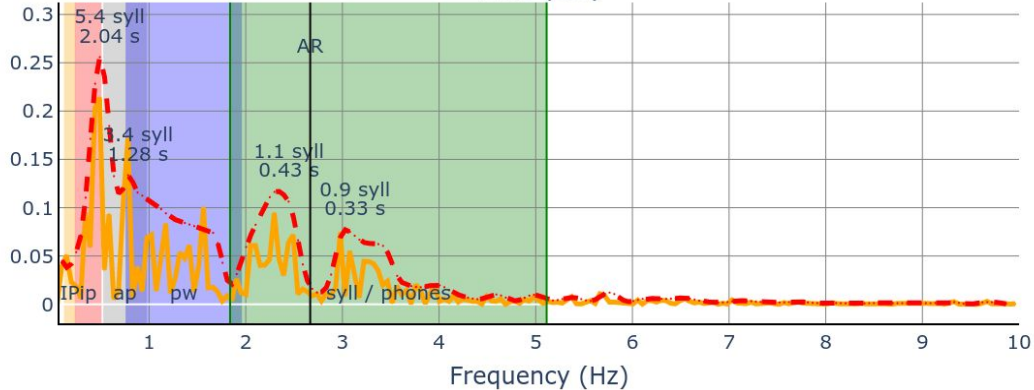
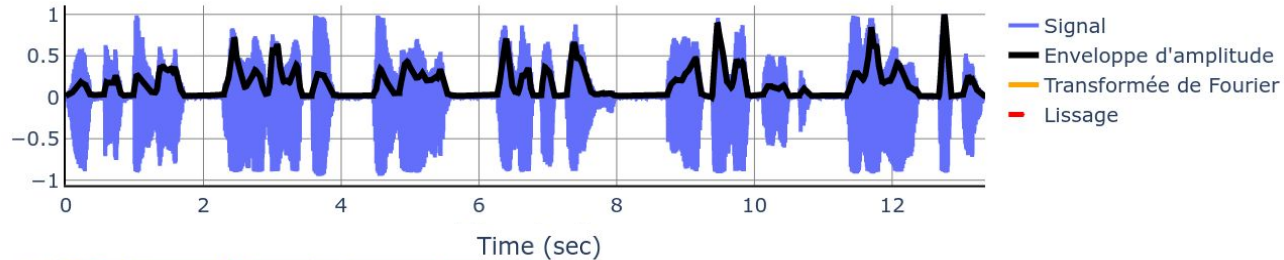
- **Sujet contrôle**

- Sévérité : **10**



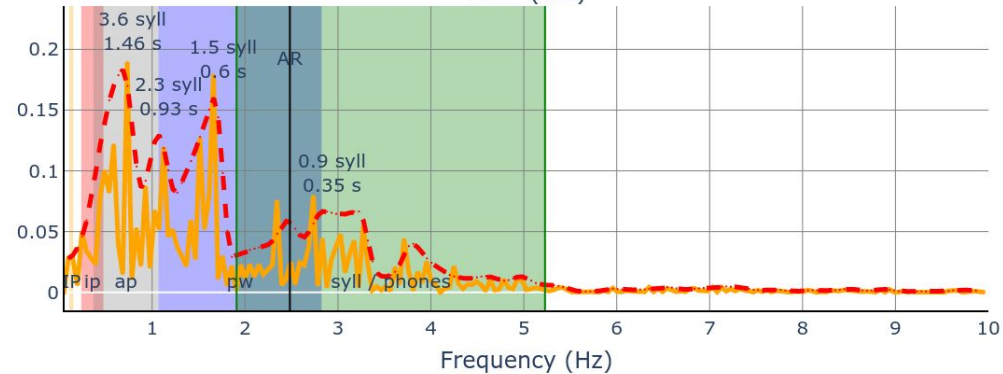
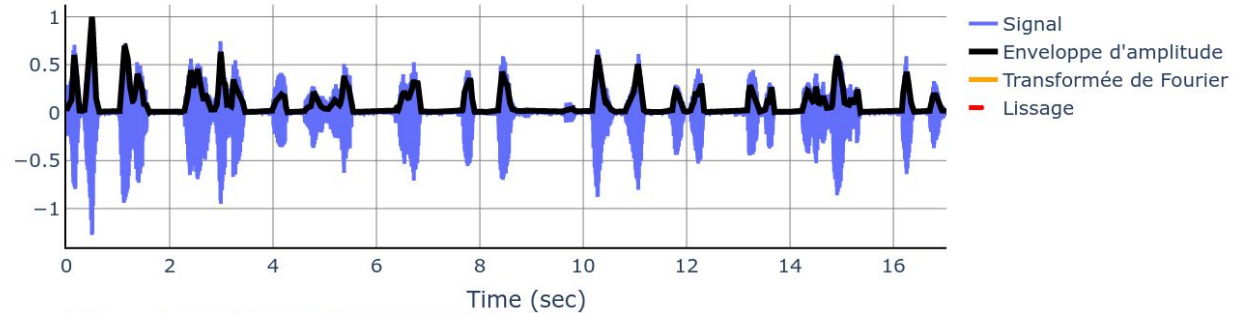
EMS de la parole lue

- Application à la parole pathologique
- Patient **cancer VADS**
 - Sévérité : **1.33**



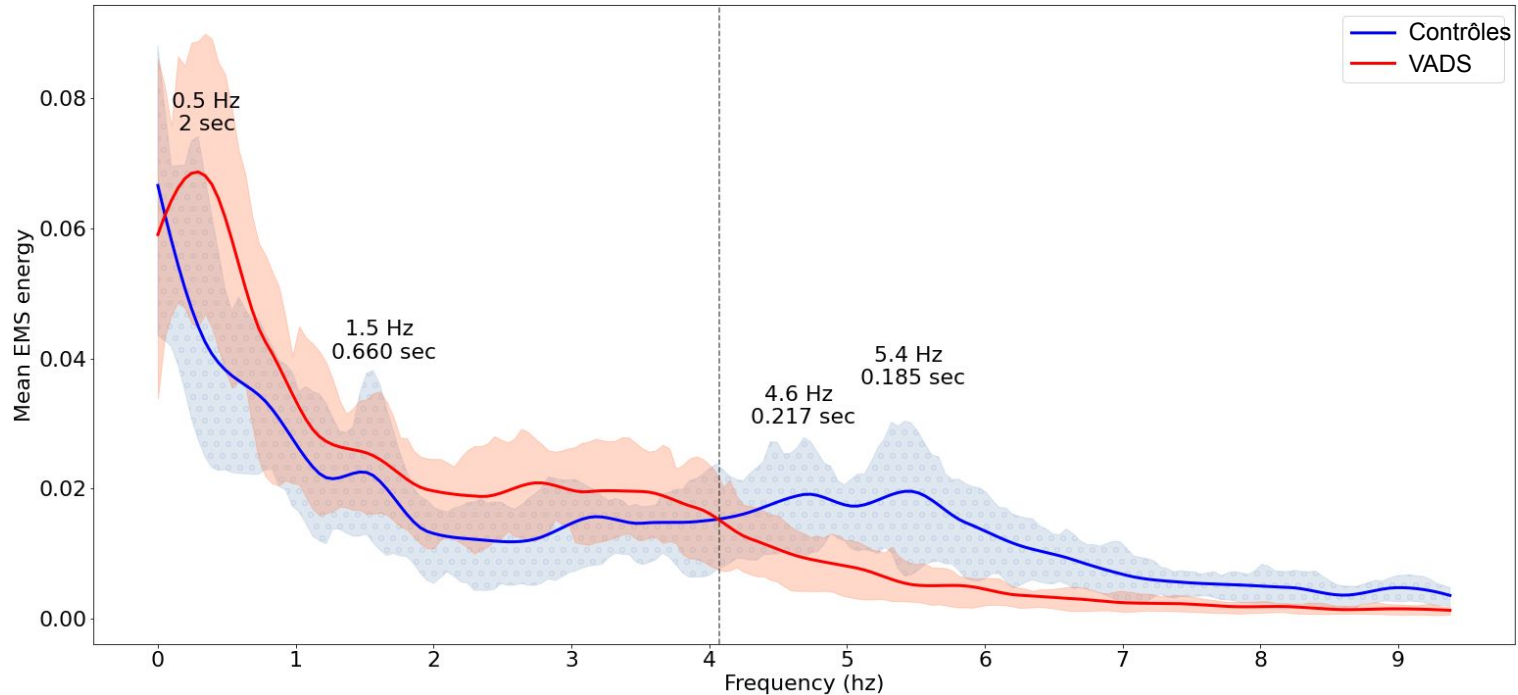
EMS de la parole lue

- Application à la parole pathologique
- Patient **cancer VADS**
 - Sévérité : **1.88**



EMS de la parole lue

- Moyenne des EMS sur l'ensemble du corpus



- Modélisation automatique du rythme par le spectre de modulations d'amplitude
 - Prise en compte de la **hiérarchie prosodique**
 - Visualisation de la **richesse spectrale** des unités de la parole (phonèmes /.../ IP)
- Compensation des déficits articulatoires par une structuration prosodique régulière des énoncés

- Prise en compte de **l'évolution temporelle** du rythme
 - Représentation sous forme de **spectrogramme du rythme**
- Analyse du **rôle des pauses** dans l'équilibrage rythmique
- **Détermination automatique des unités hiérarchiques** liées aux pics de l'EMS
- Extraction de paramètres stables de l'EMS pour la caractérisation du rythme

Bibliographie

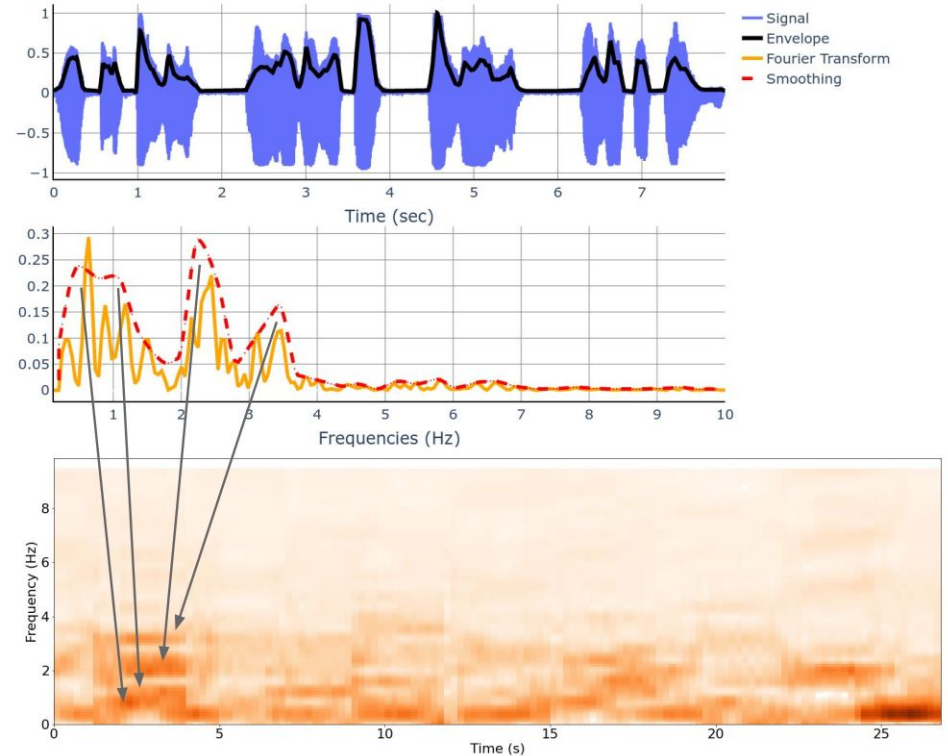


- Astésano, C. 2019, «The prosodic word as the domain of french accentuation empirical evidence», In *Phonetics and Phonology in Europe (PaPE) 2019*, p.170–171.
- Gibbon, D. (2021). The rhythms of rhythm. *Journal of the International Phonetic Association*, 1-33.
- Jun, S.-A. & C. Fougeron. 2000, «A phonological model of french intonation», In *Intonation*, Springer, p. 209–242.
- Le Coz, M. 2014, *Spectre de rythme et sources multiples : au cœur des contenus ethnomusicologiques et sonores*, thèse de doctorat, Université Toulouse 3, Paul Sabatier.
- Ling, L. E., E. Grabe et F. Nolan. 2000, «Quantitative characterizations of speech rhythm: Syllable-timing in singapore english», *Language and speech*, vol. 43, n o 4, p. 377–401
- Liss, J. M., S. LeGendre et A. J. Lotto. 2010, «Discriminating dysarthria type from envelope modulation spectra», *Journal of Speech, Language, and Hearing Research*, vol. 53, no 5, p. 1246–1255.
- Michelas, A. & M. D’Imperio. 2010, «Accentual phrase boundaries and lexical access in french», In *Speech Prosody*, p.4
- Ramus, F. et J. Mehler. 1999, «Language identification with suprasegmental cues: A study based on speech resynthesis», *The Journal of the Acoustical Society of America*, vol. 105, n o 1, p. 512–521.
- Tilsen, Sam & Johnson, Keith. (2008). Low-frequency Fourier analysis of speech rhythm. *The Journal of the Acoustical Society of America*. 124. EL34-9. 10.1121/1.2947626.
- Vaysse, R., J. Farinas, C. Astésano et R. André-Obrecht. 2021, «Automatic Extraction of Speech Rhythm Descriptors for Speech Intelligibility Assessment in the Context of Head and Neck Cancers», dans *Proc. Interspeech 2021*, p. 1912–1916

Spectrogramme du rythme

Représentation en spectrogramme

- Calcul de l'EMS sur des fenêtres glissantes
- Évolution du rythme dans le temps



EMS de la parole spontanée

Application à la parole spontanée

