

Université Abou BakrBelkaïd de Tlemcen
Faculté de Technologie
Département de Génie Biomédical
Laboratoire de Recherche de Génie Biomédical



MEMOIRE DE PROJET DE FIN D'ETUDES

Pour l'obtention du Diplôme de

MASTER en GENIE BIOMEDICAL

***Spécialité:* Instrumentation Biomédicale**

Présenté par : *TRIA KHELIF*

***Réalisation d'une Interface Dédiée au Traitement
Numérique du Signal Vocal***

Soutenu le -- juillet 2021 devant le Jury

Mme F.MEZIANI	MCA	Université de Tlemcen	Président
Mr M. KHOLKHAL	MCB	Université de Tlemcen	Examineur
Mme S. ZIANI CHERIF	MCB	Université de Tlemcen	Encadreur

Année universitaire 2020-2021

Remerciements

AU NOM DU DIEU CLÉMENT ET MISÉRICORDIEUX

Je tiens à remercier mon encadreur Mme « Ziani Cherif Souhila » pour l'orientation, la confiance, la patience, qui ont constitués un apport considérable sans lequel ce travail n'aurait pas pu être mené au bon port.

Qu'elle trouve dans ce travail un hommage vivant à sa haute personnalité.

Mes remerciements s'étendent également à Mlle BAKAAK Nour El Houda et Mr KHOLKHAL Mourad pour m'avoir fait l'honneur d'examiner ce travail, leurs bonnes recommandations et explications qui m'ont éclairé le chemin de la recherche pour l'accomplissement de ce modeste travail.

Je tiens à exprimer mes sincères remerciements à tous les professeurs qui nous ont enseigné et qui par leurs compétences m'ont soutenu dans la poursuite de mes études.

Dédicaces

Je dédie principalement et au premier lieu ce travail à *mes parents.*

Ensuite je le dédie à ma femme, puis à tous les gens qui m'ont inspiré pour suivre ce chemin.

Sommaire

Remerciements	2
Dédicaces	3
Sommaire	4
Tables des figures.....	6
Liste des tableaux	7
Résumé	8
Introduction général	10
I. Chapitre I : Généralité sur le signal vocal	11
1. Introduction	12
2. Le système anatomique	12
2.1. Configuration externe.....	13
2.2. Configuration interne	17
2.3. Innervation des cordes vocales.....	18
2.4. Les pathologies.....	21
3. Conclusion	25
II. Chapitre II : Méthodes de Traitement Numérique du Signal et détermination de la Fréquence Fondamentale	26
1. Introduction	27
2. Etat de l'art	28
3. Méthodes Temporelles utilisées	29
3.1. Autocorrélation.....	29
3.2. L'énergie	30
4. Méthodes Spectrales	30
4.1. Spectre et Densité Spectrale de puissance.....	30
4.2. Transformée de Fourier à Court Terme (STFT).....	32
5. Analyse Statistique	32
5.1. Valeur moyenne de F0	32
5.2. Ecart-Type	33
5.3. Coefficient de Variation	33
5.4. Jitter.....	34
5.5. Shimmer	34
6. Conclusion	35

III. Chapitre III : présentation de la plateforme	36
1. Introduction	37
2. Interface développée	37
3. La base de données	38
4. Traitement du Signal Vocal	39
4.1. Application des méthodes du traitement des signaux numériques pour la détermination de la fréquence fondamentale F0	39
5. Conclusion	44
IV. Chapitre 4 : Résultats et discussion	45
1. Introduction	46
2. Fréquence fondamentale	46
3. JITTER	46
4. SHIMMER	46
5. Conclusion :	51
Conclusion générale	52
Bibliographie	53

Tables des figures

FIGURE I.1. COUPE TRANSVERSALE DES CORDES VOCALES	13
FIGURE I.2.A COUPE SAGITTALE DU LARYNX: LES CARTILAGES	
B. VUE POSTERIEURE DES CARTILAGES ET LIGAMENTS LARYNGES	15
FIGURE I.3.1.LIGAMENTS ET CARTILAGES DU LARYNX	
4.2. SCHEMA MONTRANT L'ARTICULATION CRICOARYTENOÏDIENNE	15
FIGURE I.4.SCHEMA MONTRANT LES MUSCLES CRICOARYTENOÏDIENS	16
FIGURE I.5.SCHEMA MONTRANT L'ACTION DES MUSCLES INTRINSEQUES DU LARYNX	16
FIGURE I.6.CONFIGURATION INTERNE DU LARYNX.....	17
FIGURE I.7. INNERVATION MOTRICE DU LARYNX	18
FIGURE I.8.SCHEMA MONTRANT L'INNERVATION LARYNGEE	19
FIGURE I.9. LARYNGITE AIGUË[32].....	21
FIGURE I.10. LARYNGITE CHRONIQUE[32].....	22
FIGURE I.11. POLYPE DE CORDES VOCALES. [32]	23
FIGURE I.12. LES NODULES DES CORDES VOCALES[32].....	23
FIGURE I.13.TUMEURS MALIGNES DU LARYNX. [32]	24
FIGURE 14 DIAGRAMME DU TRAVAIL.....	27
FIGURE II.2. SIGNAL STATIONNAIRE ET SON SPECTRE.....	31
FIGURE II.3. SIGNAL NON STATIONNAIRE ET SON SPECTRE	31
FIGURE III.1. TRACE DU SIGNAL VOCAL 'A' POUR UN SUJET SAIN.....	40
FIGURE III.2.TRACE DU SIGNAL VOCAL 'A' ET SON SIGNAL FILTRE.	40
FIGURE III.3.TRANSFORMEE DE FOURIER DU SIGNAL VOCAL.....	41
FIGURE III.4. DETECTIONS DES PICS	42
FIGURE III.5. TRACE DE L'ENERGIE DU SIGNAL VOCAL DE LA LETTRE 'A'.....	42
FIGURE III.6.TRACE DE LA FONCTION D'AUTOCORRELATION DU SIGNAL VOCAL	43
FIGURE III.7. SPECTROGRAMME DU SIGNAL VOCAL ENREGISTRE 'A'	44
FIGURE IV.1. TRACE DE LA FFT	47
FIGURE IV.2.TRACE DE L'ENERGIE DE NOTRE SIGNAL	47
FIGURE IV.3.TRACE DE LA FONCTION D'AUTOCORRELATION DU SIGNAL VOCAL.....	48
FIGURE IV.4. TRACE DE LA DSP DU SIGNAL VOCAL.....	48
FIGURE IV.5. TRACE DU SPECTROGRAMME.	49

Liste des tableaux

TABEAU III-1. LA BASE DE DONNEES UTILISEE.....	38
TABEAU IV-1. LES RESULTATS OBTENUS DE LA FREQUENCE FONDAMENTALE.	50
TABEAU IV-2. RESULTATS DE VALEURS DE L'ENERGIE.....	51

Résume

Les symptômes les plus fréquents des troubles de la voix sont des irrégularités des cycles de la parole dont les origines sont multiples. L'endoscopie est une technique utilisée en médecine pour l'exploration visuelle afin de diagnostiquer des pathologies Laryngées. Le traitement de signal reste un outil très puissant permettant d'aider le médecin à établir un bon diagnostic à travers les différentes méthodes exploitées. Le signal vocal étant non-stationnaire contient certains paramètres acoustiques qui permettent d'identifier ces pathologies.

Ce mémoire fournit une introduction au sujet de l'analyse du signal vocal. Il examine les fondements et l'état de l'art du traitement de signal à travers trois grandes parties: l'anatomie vocale, la décomposition de la voix et le traitement numérique du signal permettant une analyse temporelle et fréquentielle en exploitant la fréquence fondamentale pour le calcul de paramètres pertinents. Nous visons à établir un point de départ pour les médecins afin de les aider à établir un bon diagnostic afin de les aider dans la procédure thérapeutique et le suivie du patient.

Notre projet de fin d'étude concerne le développement et la réalisation d'une interface dédiée au traitement numérique du signal vocal. Cette interface peut être facilement utilisée par un médecin ou un patient.

الملخص

ان جل الاضطرابات المتعلقة بالصوت تتجلى في عدم انتظام دورات الكلام المتباينة . الاوندوسكوبي (التنظير الداخلي) هي التقنية الرائدة في ميدان الطب من أجل تشخيص أمراض الحنجرة , اذ أن معالجة الاشارة تعتبر ملاذا يمهد التشخيص النموذجي الذي يحتاجه الطبيب , باعتبار أن الاشارة الصوتية غير الثابتة تشمل معلومات معينة من شأنها أن تحدد المرض.

هاته المذكرة التي بين أيديكم المبوبة بعنوان " انشاء واجهة لمعالجة الاشارة الصوتية " تتجلى في ثلاثة نقاط رئيسية : تشريح الصوت فتحليله فمعالجة الاشارة التي تسمح بالتحليل الزمني . وكل هذا المجهود ازاء انشاء محطة ذات صدق تمنح للأطباء دفعة نحو تشخيص مثالي للمرض ومتابعة حالات المرضى.

Abstract

The most common symptoms of voice disorders are irregularities in speech cycles that have multiple origins. Endoscopy is a technique used in medicine for visual exploration to diagnose laryngeal pathologies. Signal processing remains a very powerful tool to help the doctor make a good diagnosis through the different methods used. The voice signal being non-stationary contains certain acoustic parameters which make it possible to identify these pathologies.

This dissertation provides an introduction to the subject of speech signal analysis. It examines the foundations and the state of the art of signal processing through three main parts: vocal anatomy, voice decomposition and digital signal processing allowing temporal and frequency analysis by exploiting the fundamental frequency for the calculation of relevant parameters. We aim to establish a starting point for doctors to help them establish a good diagnosis in order to help them in the therapeutic procedure and the follow-up of the patient.

Our end of study project concerns the development and realization of an interface dedicated to the digital processing of the voice signal. This interface can be easily used by a doctor or a patient.

Introduction général

La voix est l'ensemble des sons produits sous l'influence de l'air des poumons qui passe dans le larynx et fait vibrer les cordes vocales. La parole est la modification de ces sons en fonction de la forme du conduit vocale. Le système de production de la voix contient trois parties: les poumons ont le rôle de générer le flux d'air, le larynx transforme ce flux d'air en énergie acoustiques et les cavités supra-laryngées pour amplifier l'énergie acoustique.

La modification anormale de la voix correspond aux troubles de la voix qui sont des conséquences des pathologies organiques ou fonctionnelles des cordes vocales. Les maladies fonctionnelles comprennent tous les troubles qui peuvent aboutir à un changement du système vocal suite à une mauvaise utilisation de la voix. Ces maladies peuvent être traitées par une thérapie vocale. Une pathologie organique est causée par une maladie structurelle (anatomique ou physiologique), soit une maladie du larynx lui-même, soit par une maladie systémique distante qui modifie la structure ou la fonction du larynx.

Les maladies laryngales peuvent être traitées via une thérapie de la voix ou par une intervention chirurgicale. Pour cette raison, une méthode robuste qui permet le diagnostic et l'évaluation du degré de sévérité de la pathologie est d'une grande importance pour le suivi de l'évolution de l'état du patient.

L'évaluation clinique de la voix est régulièrement basée sur la perception de la parole par l'auditeur. Les cliniciens évaluent sur une échelle allant de 0 à 3 le degré d'anomalie globale perçue appelée grade pour suivre l'état de la voix des patients. Cette méthode d'évaluation est subjective dans le sens que le résultat d'évaluation dépend de l'auditeur. Un inconvénient majeur de cette évaluation subjective, avec une grande variabilité intra- et inter juges [1, 2]. Les expériences ont montré que pour obtenir des évaluations reproductibles, les auditeurs (juges) doivent avoir une expérience substantielle dans l'évaluation du timbre de la voix.

Notre travail consiste en l'élaboration d'une interface capable d'enregistrer le signal vocal et d'établir son traitement numérique afin de pouvoir extraire certains paramètres pertinents qui puissent orienter le médecin à établir son diagnostic. Nous avons réparti notre mémoire en quatre chapitres, comme suit :

- Le chapitre I : Généralité sur le signal vocal.
- Le chapitre II : Méthodes de traitement numérique du signal vocal et détermination de la fréquence fondamentale.
- Chapitre III : Elaboration de notre interface de traitement du signal vocal.
- Chapitre IV : Interprétation des résultats et discussions.

I. Chapitre I : Généralité sur le signal vocal

1. Introduction

L'information portée par le signal parole peut être considéré de plusieurs façons. Nous distinguons généralement plusieurs niveaux de description non exclusifs: acoustique, phonétique et phonologique. Au niveau acoustique, nous nous sommes intéressés essentiellement au signal qui sera caractérisé par son intensité, sa fréquence, son timbre et ses propriétés statistiques, au plan phonétique.

Le traitement de la parole fait l'objet de recherches dans tous les laboratoires des grands opérateurs de télécommunications. Les travaux se sont intensifiés avec l'apparition du traitement numérique du signal. Ce vaste domaine concerne essentiellement l'enregistrement et l'analyse de la parole, le codage de la parole, la synthèse de la parole et la reconnaissance de la parole. Dans notre travail nous allons nous intéresser qu'à la première partie, à savoir, l'enregistrement et l'analyse numérique du signal vocal.

2. Le système anatomique

Les cordes vocales sont au nombre de deux et elles constituent l'élément essentiel du larynx, en effet, les fonctions sphinctériennes laryngées (respiratoires et phonatoires) dépendent de l'aspect et de l'état de ces deux éléments.

Les cordes vocales s'étendent d'avant en arrière de l'angle rentrant du cartilage thyroïde ou commissure antérieure, située en avant de l'organe laryngé jusqu'à la base du cartilage aryténoïdien en arrière ou apophyse vocale, elles se situent à 8 mm du bord inférieur du cartilage thyroïde. Leur dimension est variable surtout en fonction du sexe puisqu'elle est estimée à 22 mm chez l'homme et 18 à 20 mm chez la femme (voir la figure1). Les cordes vocales contiennent différents éléments anatomiques (muqueuse, muscles, cartilage) et font partie du larynx qui est un organe impair et médian, situé à la partie moyenne de la gaine viscérale du cou et qui occupe la partie supérieure de la région sous hyoïdienne.

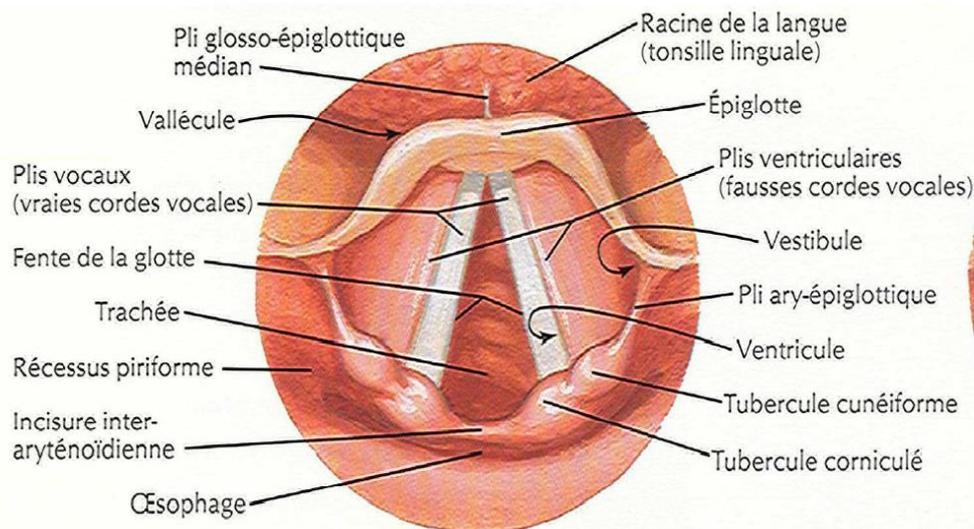


Figure I.1. Coupe transversale des cordes vocales

Il est formé par un squelette cartilagineux suspendu à l'os hyoïde, ses différents cartilages sont unis par des articulations et des membranes et l'ensemble est doublé d'un appareil musculaire et recouvert par une muqueuse de type respiratoire.

2.1. Configuration externe

2.1.1. Les cartilages

Le larynx est constitué par l'assemblage de 11 cartilages :

- ❖ Trois cartilages impairs et médians: le cartilage thyroïde, le cartilage cricoïde et l'épiglotte.
- ❖ Quatre cartilages pairs: les cartilages aryténoïdes, les cartilages cornicules, de Santorin et les cartilages cunéiformes de Wrisberg.

✓ Cartilage thyroïde

Le plus volumineux, il a la forme d'un bouclier et il est formé de deux lames latérales verticales et obliques, formant un angle dièdre ouvert en arrière et responsable en avant du relief de la proéminence laryngée ou pomme d'Adam.

✓ Cartilage cricoïde

Il a classiquement la forme d'une bague chevalière à chaton postérieur et à anneau antérieur. Le cartilage cricoïde est surmonté par les cartilages aryténoïdes en arrière, par le cartilage thyroïde latéralement et en avant, où il s'articule au niveau de ses cornes inférieures.

✓ Cartilage épi glottique

Il a la forme d'une raquette à manche inférieur incurvé dont l'extrémité forme le pied d'épiglotte qui s'insère dans l'angle rentrant du cartilage thyroïde.

✓ Cartilage aryténoïde

Il a la forme d'une pyramide triangulaire dont la base s'articule avec le cricoïde, il présente un apex supérieur, une face médiale, une face antérolatérale, une face dorsale, un angle postéro-latéral : le processus musculaire et un angle antéro-médial : le processus vocal.

2.1.2. Les articulations, membranes et ligaments

✓ Articulation cricoaryténoïdienne (voir la Figure 3)

Les articulations cricoaryténoïdiennes permettent des mouvements de translation frontale, écartant ou rapprochant les cartilages aryténoïdes, de rotation latérale ou médiale autour d'un axe vertical, écartant ou rapprochant les processus vocaux et donc les cordes vocales.

✓ Articulations crico-thyroïdiennes

Elles permettent la bascule vers l'avant ou vers l'arrière du cartilage thyroïde, avec un effet sur la tension des cordes vocales. Les différentes pièces cartilagineuses du larynx sont unies par des membranes renforcées éventuellement par des ligaments (voir les Figures 2, 3).

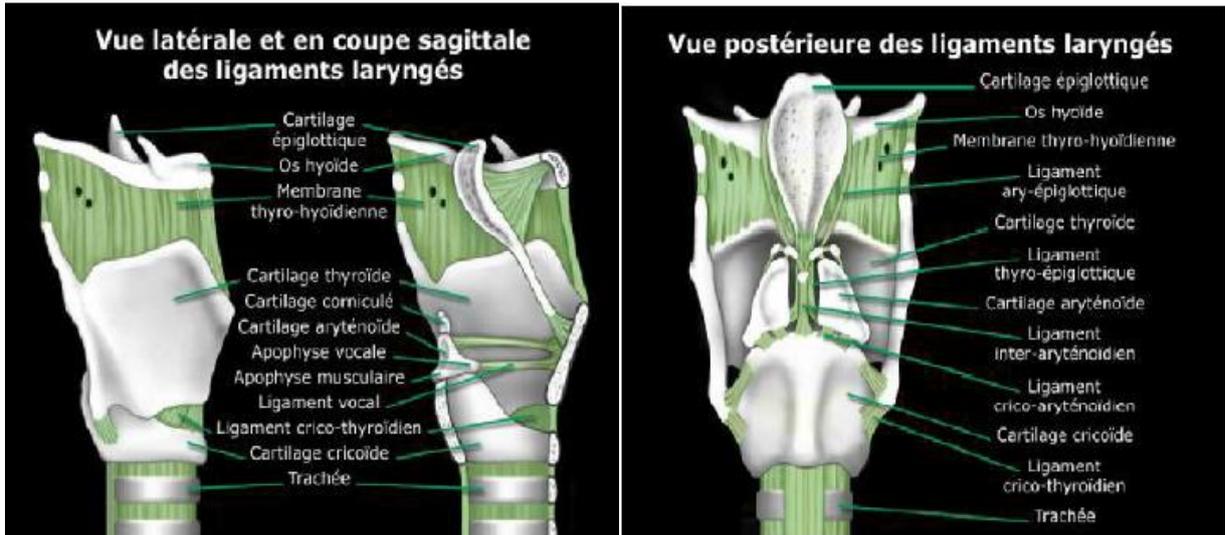


Figure 1.2.a Coupe sagittale du larynx: les cartilages

b. Vue postérieure des cartilages et ligaments laryngés

et ligaments laryngés

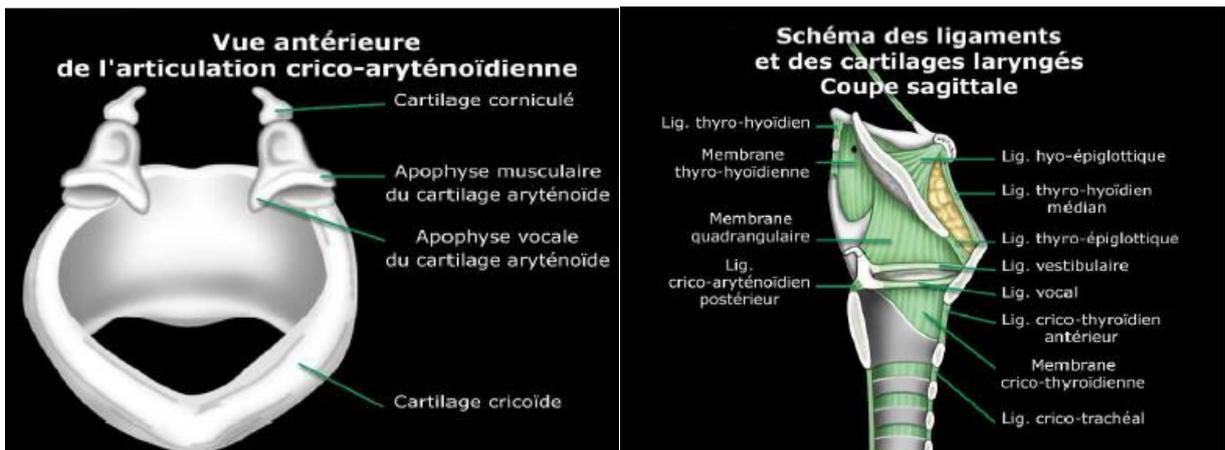


Figure 1.3.1. Ligaments et cartilages du larynx

4.2. Schéma montrant l'articulation cricoaryténoïdienne

2.1.3. Les muscles

a. Intrinsèques

Nous distinguons trois groupes musculaires :

- ❖ Les muscles tenseurs des plis vocaux qui sont les deux crico-thyroïdiens (Les seuls innervés par le nerf laryngé supérieur) (voir la Figure 5).
- ❖ Les muscles dilatateurs de la glotte qui sont, les deux cricoaryténoïdiens postérieurs.
- ❖ Les sept muscles constricteurs de la glotte qui sont :

- ✓ Les deux cricoaryténoïdiens latéraux (voir la Figure 4).
- ✓ L'inter-aryténoïdien.
- ✓ Les deux thyroïde-aryténoïdiens supérieurs.
- ✓ Les deux thyroïde-aryténoïdiens inférieurs dont la couche interne entre dans la constitution des plis vocaux (voir les Figures 5, 6).

b. Extrinsèques

Elles constituent l'appareil suspenseur du larynx, lui conférant par ailleurs une grande mobilité, trois groupes musculaires assurent cette fonction et constituent les « bretelles de suspension du larynx » :

- ❖ Les bretelles antérieures relient le larynx à la mandibule : les muscles sus-hyoïdiens.
- ❖ Les bretelles inférieures relient le larynx à l'orifice supérieur du thorax : les muscles sous-hyoïdiens.
- ❖ Les bretelles supérieures le relient à la base du crâne : muscles stylo-hyoïdien et digastrique.

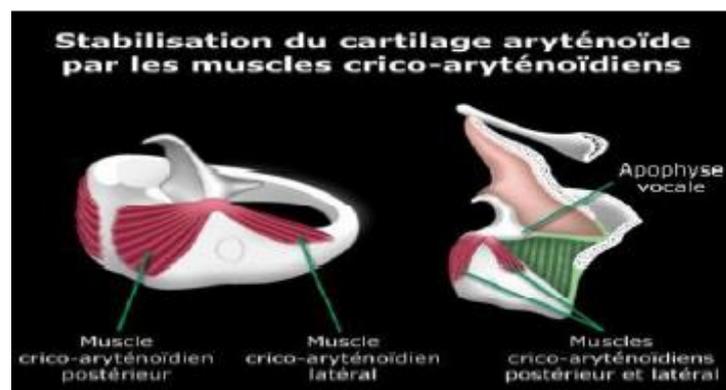


Figure 1.4. Schéma montrant les muscles cricoaryténoïdiens

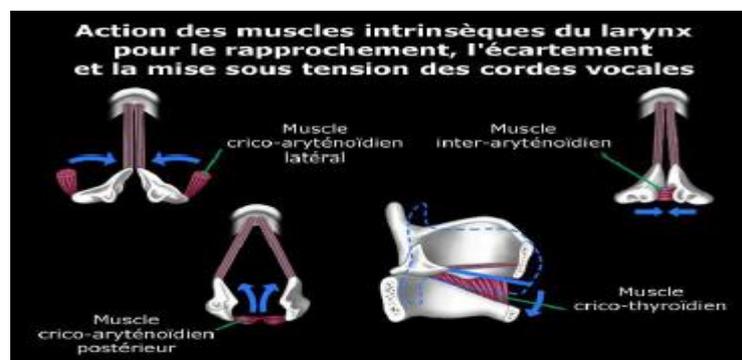


Figure 1.5. Schéma montrant l'action des muscles intrinsèques du larynx

2.2. Configuration interne

Le larynx est un tube de 5 cm de haut, 3,5 cm plus large à sa partie haute ; Il est tapissé par une muqueuse en continuité avec la muqueuse pharyngée et trachéale. Cette muqueuse est doublée par une membrane fibroblastique tendue du ligament Aryépiglottique en haut à l'arc cricoïdien en bas (voir la Figure 6). Cette membrane présente deux épaisissements :

- ❖ Le ligament vestibulaire ou thyroïde-aryténoïdien supérieur,
- ❖ Le ligament vocal ou thyroïde-aryténoïdien inférieur.
- ❖ La cavité laryngée est divisée en trois étages par ces deux replis :
 - ✓ L'étage supérieur ou vestibule laryngé.
 - ✓ L'étage moyen, limité en haut par les cordes vocales supérieures et en bas par les cordes vocales inférieures.
 - ✓ L'étage inférieur ou infra glottique, en continuité en bas avec la trachée.

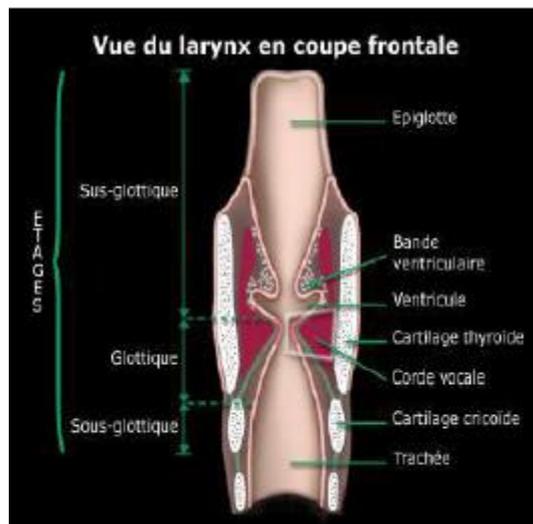


Figure I.6. Configuration interne du larynx.

2.3. Innervation des cordes vocales

L'innervation laryngée est double : centrale et périphérique. Elle est assurée par les nerfs laryngés supérieur et inférieur, branches du nerf vague ou pneumogastrique, dixième paire crânienne (voir la Figure 7).

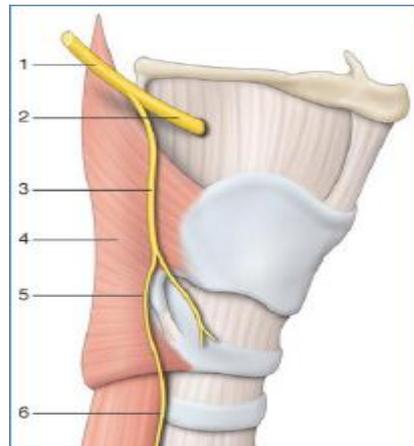


Figure 1.7. Innervation motrice du larynx

1. Nerf laryngé supérieur ; 2. Branche interne ; 3. Branche externe ;
4. Muscle constricteur inférieur ; 5. Anse de Galien ; 6. Nerf récurrent laryngé.

2.3.1. Nerf laryngé supérieur

C'est un nerf mixte, essentiellement sensitif, qui naît du nerf vague, au pôle inférieur du ganglion plexiforme (voir la Figure 8). Il descend contre la paroi pharyngée pour se diviser en arrière de la corne de l'os hyoïde en deux branches :

- ✓ Une branche médiale ou supérieure qui perfore la membrane thyroïde-hyoïdienne ; Elle donne l'innervation sensitive de la muqueuse supérieure du larynx, de la partie adjacente du pharynx et de la base de langue.
- ✓ Une branche latérale ou inférieure qui innerve le muscle crico-thyroïdien, puis perfore la membrane crico-thyroïdienne et donne l'innervation sensitive des étages moyen et inférieur du larynx et assure le tonus des muscles du larynx.

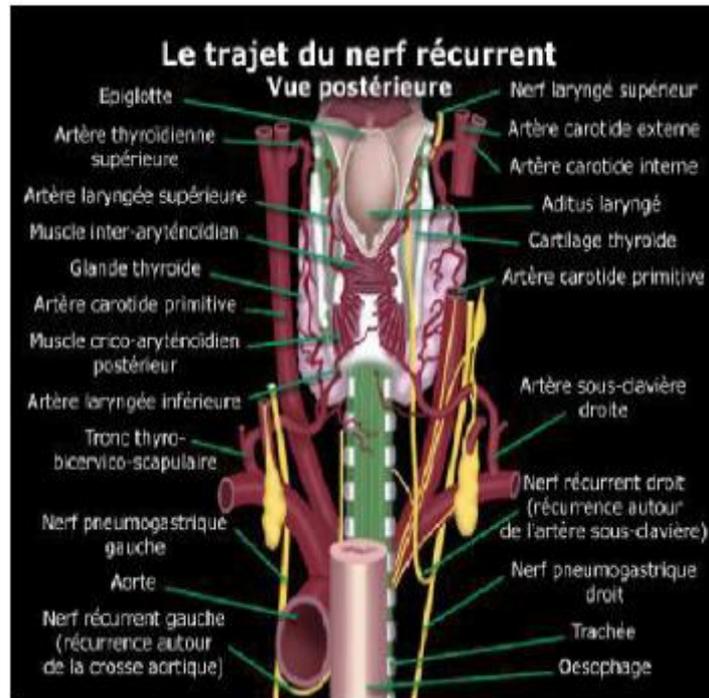


Figure I.8. Schéma montrant l'innervation laryngée

2.3.2. Nerf laryngé inférieur

C'est la branche terminale du nerf laryngé inférieur (récurrent) (figure 10) qui naît du nerf vague, à droite au-dessous de l'artère sous-clavière, à gauche au-dessous de l'arc aortique, monte vers le larynx dans l'angle osso-trachéal. Il donne des rameaux trachéaux, œsophagiens, cardiaques, et devient le nerf laryngé inférieur quand il passe sous le constricteur inférieur du pharynx. Il donne alors trois rameaux :

- ✓ Un rameau anastomosé avec le nerf laryngé supérieur constituant l'anse de Galien.
- ✓ Un rameau postérieur.
- ✓ Un rameau antérieur.

Le nerf laryngé inférieur innerve la muqueuse postérieure et tous les muscles du larynx, sauf le crico-thyroïdien.

2.3.3. Innervation périphérique

Elle est de trois ordres : sympathique, sensitive, motrice.

- ✓ L'innervation sympathique : Elle suit la vascularisation artérielle issue du plexus laryngé de Haller.

- ✓ L'innervation sensitive : Elle est assurée par le nerf laryngé supérieur ; Il faut noter que les récepteurs sensitifs sont très nombreux et surtout responsables d'une grande sensibilité de la muqueuse laryngée.

- ✓ L'innervation motrice : Elle dépend du nerf récurrent pour le muscle vocal et du nerf laryngé externe, branche du nerf laryngé supérieur, pour le muscle crico-thyroïdien.

2.3.4. d. Innervation centrale

❖ Bulbaire

Elle est constituée par les centres laryngés bulbaires qui sont au nombre de deux :

- ✓ Le centre bulbo-respiratoire qui va permettre l'ouverture glottique donc l'inspiration, parallèlement ce centre permet l'expansion de la cage thoracique.

- ✓ Le centre bulbo-phonatoire : Il envoie des fibres centrifuges responsables de l'innervation des muscles nécessaires à la phonation.

❖ Corticale

Il existe également deux centres :

- ✓ Le centre cortico-phonatoire : Le centre cortico-phonatoire semble exister au niveau de l'opercule rolandique, ceci serait confirmé par l'existence de pathogénie des laryngoplégies d'origines corticales.

- ✓ Le centre cortico-respiratoire.

2.4. Les pathologies

2.4.1. Inflammation aiguë du larynx ou laryngite aiguë

Par le fait de bactéries, un simple refroidissement peut se transformer en laryngite aiguë.

Cette irruption de bactéries s'attaque directement au larynx, ou s'introduit par le nez dans la gorge et ensuite dans le larynx. D'autre part, les laryngites peuvent aussi être la suite d'autres infections comme la rougeole, la scarlatine, la syphilis, la tuberculose ou la fièvre typhoïde. Elle peut encore avoir d'autres causes: par exemple, forcer la voix, tonitruée, crié à tue-tête, vociférer [2].

Ces excès sont souvent accompagnés de rauçité brulante, d'accès de toux violente et parfois d'une légère hausse de température. Un laryngoscope révélera que la muqueuse est rouge et que les cordes vocales, normalement d'un blanc clair, sont également rouges et gonflées. Chez les enfants, le gonflement des cordes vocales peut rendre la respiration difficile, allant jusqu'à l'oppression (Figure 9)



Figure I.9. *Laryngite aiguë*[32]

2.4.2. Laryngite chronique :

La laryngite chronique a les mêmes symptômes que celle de la forme aiguë, elle est souvent un prolongement de fautes de soins et à un retard à s'attaquer à la cause. Elle peut aussi être due à un forçage continu de la voix, par exemple chez des orateurs, des marchands de quatre saisons. Ou encore à l'inhalation des poussières, par exemple dans les manufactures de tabac, les matelassiers, chez les meuniers, les mineurs, les tailleurs de pierre, etc. Ou tout simplement par l'abus du tabac et de l'alcool. La muqueuse du larynx est alors d'un gris rouge, gonflée par endroits et couverte d'excrétions purulentes et glaireuses [3] [2]

(Figure 10).

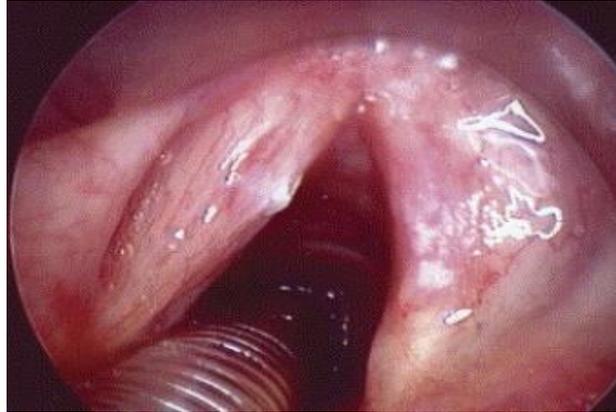


Figure I.10. Laryngite chronique[32]

2.4.3. Les tumeurs du larynx

Les tumeurs du larynx peuvent être bénignes ou malignes. La première catégorie comprend les polypes, les enflures granuleuses et nodulaires (par exemple, les chanteurs sont souvent affectés par ces enflures nodulaires, saillies sur les bords des cordes vocales).

2.4.4. Les tumeurs bénignes du larynx

- ❖ Polype de cordes vocales : Souvent secondaire à un malmenage vocal, le polype peut-être favorisé par une intoxication alcoolo-tabagique ou une exposition aux poussières. Le rôle de l'inflammation joue peu, il s'agit d'une tuméfaction bénigne de la corde vocale du plan glottique. Il est de forme lisse ou mamelonnée. Son implantation sessile ou pédiculée. Il est localisé au 1/3 moyen ou à la commissure antérieure. Il se présente sous une forme inflammatoire, œdémateuse ou hémorragique. Il est responsable d'une dysphonie: La voix est grave, voire bitonale, le diagnostic repose sur la laryngoscopie indirecte et la fibroscopie [7] (Figure11).

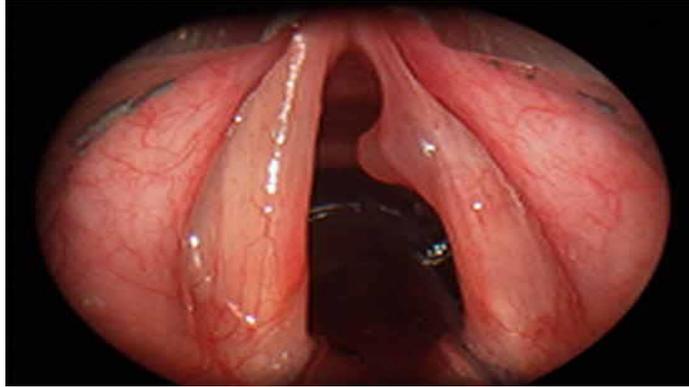


Figure I.11. *Polype de cordes vocales.* [32]

- ❖ Les nodules des cordes vocales : Ils correspondent à un épaississement localisé, généralement bilatéral (des deux côtés) du rebord libre de la corde vocale. Cet épaississement se situe à la jonction du premier et du deuxième tiers des cordes vocales et empêche l'approximation de celles-ci durant la phonation.

Les nodules sont la conséquence du choc répété des cordes vocales entre elles à cause d'une surutilisation ou d'une mauvaise utilisation de la voix (p. ex. cris excessifs, vocalisations forcées, toux forte, éclaircissements répétés de la voix pour dégager des sécrétions, etc.)

Certains facteurs peuvent aussi favoriser leur apparition: infections fréquentes des voies respiratoires supérieures (rhinite, sinusite), bronchites, asthme, allergies, exposition à la fumée de cigarette, reflux laryngopharyngé, présence de produits irritants dans l'atmosphère (poussières, allergènes) [4] [8] (Figure 12).



Figure I.12. *Les nodules des cordes vocales*[32]

- ❖ Les granulomes : sont des formations inflammatoires qui ne donnent vraiment des troubles dysphoniques que lorsqu'ils sont volumineux et gênent l'occlusion cordiale. Le plus souvent ils se manifestent par une simple sensation de gêne pharyngée et sont découverts par l'examen laryngé systématique. Les granulomes sont unis ou bilatéraux, en général développés dans la partie postérieure de l'orifice glottique. Parfois, il n'y a pas de cause évidente, si ce n'est le surmenage vocal ou un reflux gastro-œsophagien qu'il faudra toujours soigneusement rechercher [9].

2.4.5. Les tumeurs malignes du larynx

Quant aux tumeurs malignes, le cas le plus grave est celui du cancer du larynx. Cette affection commence généralement par une insignifiante tumeur sur une corde vocale, tumeur qui se développe petit à petit, altérant la voix jusqu'à la raucité, rendant la déglutition de plus en plus pénible et provoquant de l'oppression allant jusqu'à l'étouffement. Il en résulte un amaigrissement et un affaiblissement généralisés et finalement une paralysie des cordes vocales [5].

Les tumeurs bénignes aussi bien que les malignes ne peuvent être traitées avec succès que chirurgicalement. Toutefois, l'opération variera suivant la nature de la laryngite.

L'examen des tumeurs bénignes se fait au moyen du laryngoscope de Clar, composé d'un petit miroir et d'un foyer lumineux. Le médecin introduit le miroir pour vu d'un manche assez long au fond de la bouche, entre le palais et la luette, au-dessus du larynx dont l'image est ainsi nettement réfléchi dans le miroir grâce au foyer lumineux braqué sur la bouche. Cette méthode permet de voir les tumeurs et de les éloigner au moyen de longues pinces [6] (Figure 13).



Figure 1.13. Tumeurs malignes du larynx. [32]

3. Conclusion

Nous avons abordé dans ce chapitre des généralités sur le signal vocal, commençant par le système anatomique puis les différentes pathologies pouvant affecter le larynx, qu'elles soient bénignes ou malignes.

Le chapitre suivant est consacré aux différentes méthodes de traitement numérique du signal et la détermination de la fréquence fondamentale.

II. Chapitre II : Méthodes de Traitement Numérique du Signal et détermination de la Fréquence Fondamentale

1. Introduction

Parmi les symptômes les plus fréquents des troubles de la voix nous retrouvons les irrégularités des cycles de la parole dont les origines sont multiples entre autres une malformation congénitale des plis vocaux ou organique acquise, de l'articulation laryngée ou d'une déficience respiratoire.

Afin de diagnostiquer ces pathologies laryngées, des techniques utilisées en médecine sont souvent envahissantes comme l'endoscopie. Le signal vocal est un signal non-stationnaire contenant certains paramètres acoustiques permettant d'identifier ces pathologies. Différentes techniques ou méthodes de traitement du signal sont utilisées pour l'identification et l'analyse du signal vocal afin d'extraire les paramètres dits acoustiques importants. Ces paramètres sont : la fréquence fondamentale (pitch), les formants (fréquences de résonances du conduit vocal notées $F_1, F_2, F_3, F_4, \dots$), Coefficient de Variation de F_0 , le Jitter, le Shimmer qui seront définis en fin de ce chapitre.

Nous nous sommes intéressés plus précisément au signal vocal de la voyelle « a », qui est quasi-stationnaire. Pour faciliter le traitement, nous segmentons ce signal en trames de durée fixe supposées stationnaires. Ce chapitre présente les différents outils de traitement de signal utilisé pour le traitement du signal vocal objet de notre étude.

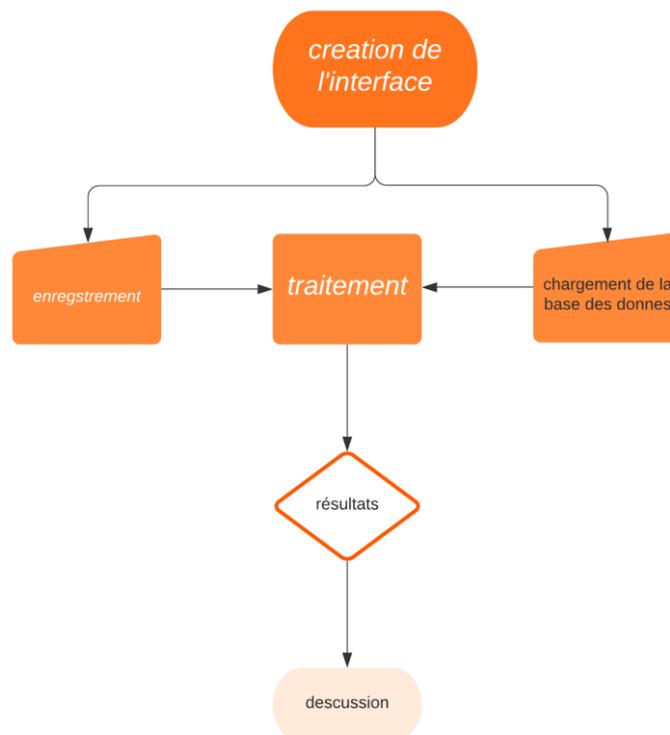


Figure 14 Diagramme du travail

2. *Etat de l'art*

Parmi les travaux qui ont été réalisées, on trouve :

1-Mémoire de master intitulé : « Analyse spectro-temporelle du signal vocal en vue de dépistage et du suivie des dysphonies chroniques d'origine laryngées », juin 2012, présenté par Mme Abdelouaheb Sara. Elle a développé un système dédié à la caractérisation objective des dysphonies chroniques d'origine laryngées, utilisant VISUAL BASIC, d'un algorithme permettant de réaliser l'analyse spectro-temporelle d'un signal vocal acoustique, elle a appliqué l'algorithme sur trois sujet sains et trois autres pathologiques. Les résultats obtenus montrent une variabilité des caractéristiques spectro-temporelles entre sujets sains et sujets pathologiques notamment en ce qui concerne les contenus spectraux évalués par TFD-CT et la fréquence fondamentale F0 moyennée sur plusieurs trames du signal voisé. [24]

2- La thèse de doctorat intitulée «Détection du fondamental de la parole en temps réel: application aux voix pathologiques », juillet 2013, présenté par Fadoua BAHJA. Ce travail s'inscrit dans le cadre de travaux de recherche qui visent la détermination de la fréquence fondamentale du signal de parole proposant un algorithme fondé sur la transformation en ondelettes, utilisant deux bases de données : Bagshaw et Keele. [26]

3- La publication de MireaFarrus et Janvier Hernando : « Jitter and Shimmer Measurements for Speaker Recognition » janvier 2007. Dans cet article, plusieurs types de mesures de jitter et de Shimmer ont été analysés. Les expériences réalisées avec la base de données de parole conversationnelle Switchboard-I montrent que les mesures de jitter et de shimmer donnent d'excellents résultats dans la vérification du locuteur en tant que caractéristiques complémentaires des paramètres spectraux et prosodiques. [22]

4- Mémoire de maitrise en électronique, sous le thème : « dé-bruitage d'un signal de la parole corrompu par un bruit coloré en utilisant la transformée en ondelettes et implantation sur un processeur de traitement numérique des signaux » par L'allouai Abdelkrim (2004). Il a utilisé une combinaison de deux types de seuillage à la fois (seuillage doux et de loi μ), la méthode utilisée consiste à décomposer le signal bruité en paquets d'ondelettes de Daubechies du 11^{ème} ordre jusqu'au 7^{ème} niveau, puis à appliquer les deux seuillages. Les résultats obtenus par cette approche démontrent une amélioration par rapport à l'utilisation individuelle de chacune des deux méthodes de seuillage.[31]

5- Mémoire de master intitulé «Analyse multi variée de la parole pour l'évaluation objective de la qualité vocale » par Melle Amel FENIZA et Melle Hassiba LAISSANI. L'objectif de l'analyse multi variée est de combiner plusieurs indices acoustiques pour l'estimation des Dy périodicités vocales en utilisant la décomposition en modes empiriques du spectre d'amplitude du signal de la parole. L'analyse est effectuée trame par trame et l'algorithme de décomposition modale empirique est appliquée au spectre d'amplitude de chaque trame afin de le décomposer en fonctions de modes

intrinsèques (IMFs). Un rapport signal à dyspériodicité vocale est ensuite calculé pour chaque IMF. La méthode proposée a été appliquée à un corpus de voyelles [a] produites par 223 locuteurs normo phoniques et 28 dysphoniques. Une analyse de régression linéaire multiple combinant les rapports signal à dyspériodicité des différentes IMFs a été utilisée pour prédire les scores de l'évaluation auditive et ses performances en termes de corrélation ont été comparées à celle de la méthode de référence basée sur le variogramme généralisé. Les résultats expérimentaux ont montré que l'analyse multi variée basée sur la décomposition modale empirique présente une meilleure performance comparée à la méthode de référence basée sur le vario gramme généralisé. [28]

6- Mémoire de magister en électronique «Techniques de détection de la période du pitch par les méthodes temps fréquence et temps échelle» présenté par AJGOU Riad. Dans ce mémoire le chercheur a étudié les différents types de représentations temps fréquence et temps échelle, et a essayé de les adapter afin de mieux détecter la période du pitch du signal vocal.

Les représentations temps-fréquence et temps-échelle s'avèrent les plus commodes à l'analyse du signal de la parole.

Les méthodes temps-échelle en générale sont les plus commodes à l'analyse de signal parole, la méthode utilisé basée sur la détection des maximums des coefficients d'ondelette qui correspondent physiquement à la variation du signal parole, l'estimation du pitch est basée sur l'énergie calculée sur les coefficients d'ondelette, le nombre de moments nuls d'ondelette conditionne la capacité d'extraire la fréquence fondamentale.

L'inconvénient essentielle de cette méthodes et de tomber dans les erreurs de doublement ou inversement dans l'estimation de la fréquence fondamental qui demeure la plus significative. [29]

7- La thèse de Damien Lolive, sous le thème : «Transformation de l'intonation: application à la synthèse de la parole et à la transformation de voix », Les travaux de cette thèse se situent dans le cadre de la transformation de la prosodie en se focalisant sur la fréquence fondamentale, F0, facteur jugé proéminent dans le traitement de la prosodie. En particulier, les recherches s'intéressent aux différentes étapes nécessaires à la construction d'un tel système: la stylisation, la classification et la transformation des contours mélodiques. Pour chaque étape, il propose une méthodologie qui tient compte des problèmes qui se sont posés à l'étape précédente. Tout d'abord, un modèle B-spline est proposé pour la stylisation des contours mélodiques. Ensuite, pour représenter l'espace mélodique du locuteur, une approche par modèles de Markov est introduite. Enfin, une méthodologie de transformation de la prosodie à partir de corpus non parallèles par une technique d'adaptation au locuteur est présentée. Les résultats obtenus tendent à montrer qu'il est nécessaire de traiter la dynamique du F0 et de piloter la transformation par des informations d'ordre morphosyntaxique.[30]

3. Méthodes Temporelles utilisées

3.1. Autocorrélation

L'autocorrélation mesure la similitude du signal $x(t)$ avec sa version décalée. Pour un signal déterministe, elle se calcule selon la formule suivante :

$$R_{xx}(\tau) = \int_{-\infty}^{+\infty} x(t)x^*(t - \tau)dt, (2.1)$$

Si le signal est aléatoire, alors la formule devient :

$$R_{xx}(\tau) = E[x(t)x^*(t - \tau)], (2.2)$$

où x^* désigne le conjugué du signal $x(t)$, τ le décalage et $E[]$ l'espérance mathématique.

Pour un signal quasi-périodique ou périodique, le signal résultant de l'autocorrélation est un signal également périodique de même période. De ce fait, l'autocorrélation est un bon outil d'estimation de la périodicité dans les signaux perturbés par du bruit en détectant des régularités, ou des profils répétés. Ainsi, la période d'un signal vocal et par conséquent sa fréquence fondamentale peut être évaluée aisément.

3.2. L'énergie

L'énergie de signale est définie par : $E = \int s(t)^2 dt$, (2.3)

4. Méthodes Spectrales

4.1. Spectre et Densité Spectrale de puissance

Le Spectre et la Densité Spectrale de Puissance, se réalisent après avoir effectué l'opération mathématique appelée Transformée de Fourier qui permet le passage du domaine temporel au domaine fréquentiel du signal.

Pour un signal numérique $x(n)$, cette transformée est exprimée comme suit :

$$X(k, f) = \sum_{n=0}^{N-1} x(n)e^{-2\pi j \frac{nk}{N}}, (2.8)$$

Avec :

N : le nombre d'échantillons et $f = \frac{f_e}{N}$ (où : f_e représente la fréquence d'échantillonnage),

$n, k = 0, 1, \dots, N - 1$ Et $X(k, f)$ est le spectre du signal.

La Densité Spectrale de Puissance d'un signal (noté DSP) correspond à la Transformée de Fourier de la fonction l'autocorrélation de ce signal :

$$DSP[x(t)] = TF(R_{xx}), \quad (2.9)$$

Cette densité de puissance représente le spectre d'énergie dont l'allure de son enveloppe est susceptible de nous indiquer l'existence et la localisation des formants et des harmoniques de la fondamentale. En effet, pour les signaux non stationnaires ou quasi-stationnaires comme le signal vocal, la Transformée de Fourier ne permet pas une bonne analyse de leur évolution temps fréquence. Comme il est si bien montré dans l'exemple ci-dessous. (Figure 2.1).

Cet exemple montre que les spectres des deux signaux sont semblables malgré que les signaux soient différents, l'un stationnaire et l'autre non-stationnaire.

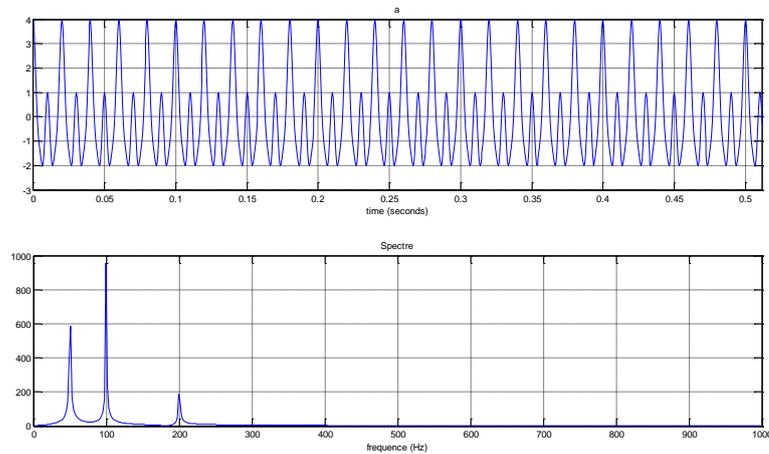


Figure II.15. *Signal stationnaire et son spectre.*

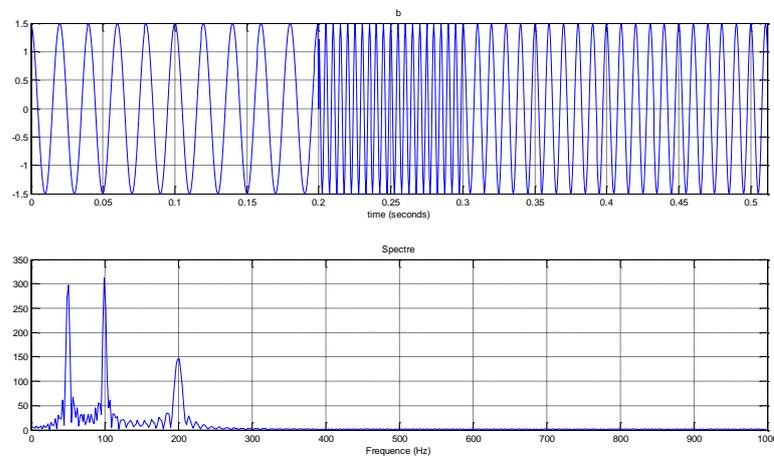


Figure II.16. *Signal non stationnaire et son Spectre*

Par conséquent, le spectre ou la densité spectrale de puissance ne nous montre pas l'évolution fréquentielle dans le temps et l'instant d'apparition des fréquences contenues dans un signal. La solution préconisée pour remédier à cela serait de découper le signal en trames (quasi-stationnaires) à l'aide d'une fenêtre de durée déterminée et d'appliquer la Transformée de Fourier à chacune d'elle afin de suivre l'évolution fréquentielle dans le temps du signal vocal, d'où l'appellation de Short Time Fourier Transformer (STFT) ou Transformée de Fourier à Court Terme et de tracer son spectrogramme.

4.2. Transformée de Fourier à Court Terme (STFT)

La Transformée de Fourier à Court Terme noté STFT (Short Time Fourier Transformer) permet d'analyser un signal à travers des trames de longueur constante. Elle s'exprime comme suit:

$$STFT_x^w(\tau, f) = \int_{-\infty}^{\infty} x(t) \cdot w^*(t - \tau) \cdot e^{-j2\pi ft} dt \quad (2.10)$$

Où $w(t)$ est la fenêtre de largeur T centrée en τ et w^* sa conjuguée. Le résultat peut être représenté sous forme de spectrogramme (temps, fréquence, amplitude à l'aide d'une palette de couleur). Ce spectrogramme est un diagramme représentant le spectre d'une séquence associant à chaque fréquence une puissance.

5. Analyse Statistique

Le Signal vocal étant non stationnaire, afin de l'analyser par ces différentes méthodes de traitement du signal, nous devons le segmenter en trame quasi-stationnaire de 23 ms ou 1024 points chacune. Ces trames se chevauchent successivement entre elles d'une durée de 2.3 ms ($\cong 100$ points). Ainsi, pour pouvoir interpréter les résultats, nous déterminons la Fréquence Fondamentale (F_{0i}) pour chaque trame (i).

Ainsi pour se prononcer sur une éventuelle pathologie Laryngée nous devons effectuer certains calculs statistiques développés ci-dessous.

5.1. Valeur moyenne de F_0

La valeur moyenne de F_0 notée $\overline{F_0}$ indique une mesure globale de la hauteur, aiguë ou grave, de la voix d'un sujet:

$$\overline{F_0} = \frac{1}{N} \sum_{i=1}^N F_{0i}, \quad (2.12)$$

Où N dans notre cas représentent le nombre de trames et F_{0i} la Fréquence Fondamentale dans la trame (i)

5.2. Ecart-Type

L'Ecart type σ de F_0 correspond à la dispersion en Hz de F_{0i} autour de la valeur moyenne $\overline{F_0}$.

$$\sigma = \sqrt{\frac{1}{N} \sum_{i=1}^N (F_{0i} - \overline{F_0})^2}, \quad (2.13)$$

5.3. Coefficient de Variation

Le coefficient de variation noté $C.V.F_0$ est exprimé par l'équation 2.14.

$$C.V.F_0 [\%] = \frac{\sigma}{\overline{F_0}}, \quad (2.14)$$

Ce coefficient est un bon indicateur de l'instabilité vibratoire des cordes vocales et devient important dans le cas d'un sujet pathologique.

5.4. Jitter

Le Jitter moyen traduit la variation de fréquence entre des cycles d'oscillation successifs :

$$\overline{Jitter} = \frac{1}{N-1} \sum_{i=1}^{N-1} |F_{0t} - F_{0t+1}| \quad (2.15)$$

$$Jitter_{facteur} [\%] = \frac{\overline{Jitter}}{F_0} \quad (2.16)$$

Pour un sujet normal le jitter facteur est inférieur à 1 % et augmente en cas de dégradation de l'appareil phonatoire^[22].

5.5. Shimmer

Le Shimmer est une mesure de la variation d'amplitude A_i de cycle à cycle.

$$\overline{Shimmer} (dB) = \frac{1}{N-1} \sum_{i=1}^{N-1} 20 \log \left(\frac{A_i}{A_{i+1}} \right) \quad (2.17)$$

Ce paramètre est inférieur ou égal à 0.4 dB pour un sujet normal et dépasse cette valeur en cas d'une pathologie Laryngée^[23].

6. Conclusion

Dans ce chapitre nous avons abordé des méthodes de traitement numérique temporel et spectral permettant de mettre en évidence certains paramètres pertinents pour établir un bon diagnostic.

Le chapitre suivant est consacré à l'élaboration de notre interface d'enregistrement et de traitement numérique du signal physiologique.

III. Chapitre III : présentation de la plateforme

1. Introduction

Dans ce chapitre nous présenterons notre plateforme d'acquisition du signal vocal. Nous développons une interface graphique sous Matlab R2015, afin de déterminer les paramètres pertinents, à savoir : la fréquence fondamentale de chaque signal enregistré pour des sujets test afin d'extraire la fréquence fondamentale F_0 du signal, puis nous appliquons les diverses méthodes afin d'extraire les paramètres statistiques, tel l'écart Type σ , les coefficients Jitter et Shimmer. Nous nous sommes aussi intéressé au calcul de l'énergie de ce signal pour voir sa variation.

Nous avons exploité ces méthodes d'analyse pour des signaux de sujets pathologiques de la base de données Physionet. Les résultats ont été analysés et interprétés et seront présentés au chapitre suivant.

2. Interface développée

L'interface graphique développée permet d'enregistrer le signal vocal, ou nous sommes intéressés au signal laryngé de la voyelle « a ». Cette interface contient deux menus (le premier pour les différentes commandes puis les tracés, et le second pour le calcul des paramètres).

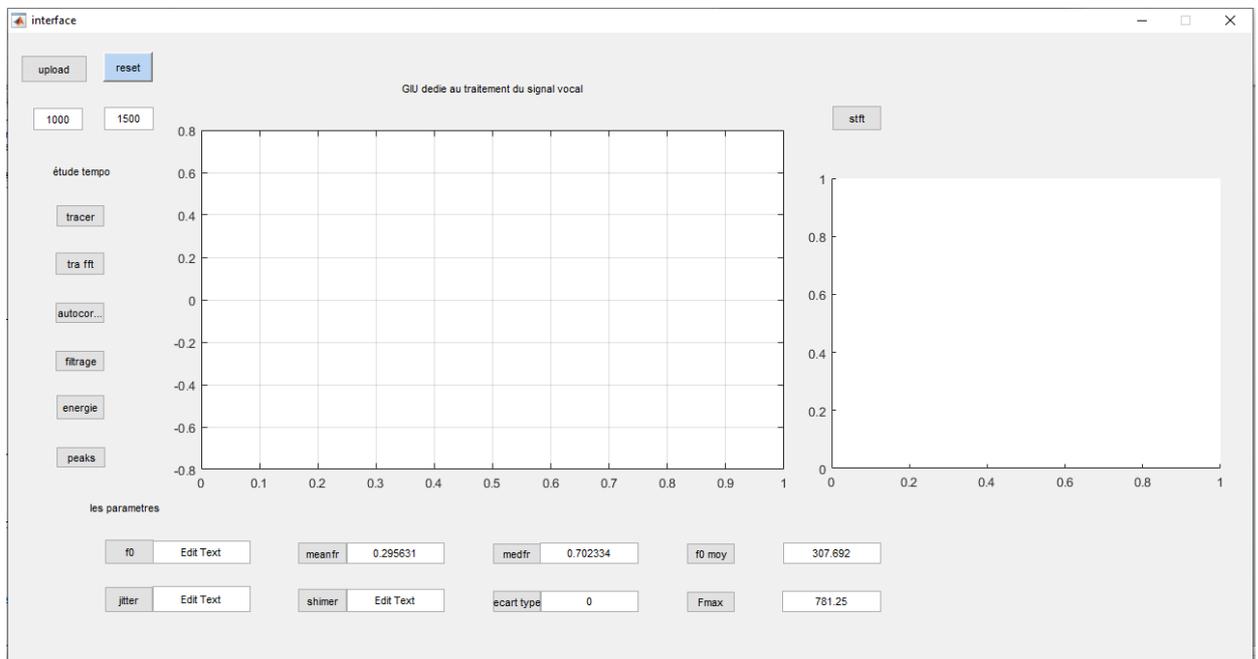


Figure III.1. L'interface graphique.

Le bouton (upload) permet de charger le signal vocal et d'établir le traitement précis.

Nous pouvons sélectionner la page qui nous permet de sélectionner le signal à traiter. D'autres boutons ont été créés. Chacun a une fonction de calcul et de tracé bien précise comme par exemple : la FFT et l'autocorrélation.

D'autres boutons tels que « énergie » ; nous permet de calculer et d'afficher l'énergie du signal. Le bouton « peak » nous permet de détecter les pics des différentes trames sélectionnées au début.

D'autres boutons nous permettent de calculer : la fréquence moyenne et la fréquence médiane.

Enfin, un menu qui contient un bouton (STFT) permet de calculer le spectrogramme de transformée de Fourier du signal et de voir la variation de la fréquence en fonction du temps.

3. La base de données

La disponibilité des bases de données des signaux vocaux accessibles au public est un facteur clé. La base utilisée dans ce travail contient 15 signaux des différents sujets sains et pathologiques.

Tableau III-1. La base de données utilisée

ID	Age	Sexe	Diagnostique	Situation	Fumeur ou non
voice001	32	M	hyperkineticdysphonia	Researcher	No
voice002	55	M	Healthy	Employee	Yes
voice003	34	M	hyperkineticdysphonia (nodule)	Researcher	No
voice004	28	F	Hypokineticdysphonia	Researcher	No
voice005	54	F	hypokineticdysphonia	Researcher	No
voice008	41	M	hyperkineticdysphonia (nodule)	Researcher	Yes
voice009	51	F	reflux laryngitis	Researcher	No
voice010	53	F	hyperkineticdysphonia (nodule)	Researcher	Y/20/D
voice011	47	F	hypokineticdysphonia	Researcher	No
voice012	55	M	reflux laryngitis	Employee	No
voice013	56	M	reflux laryngitis	Researcher	No
voice014	42	F	reflux laryngitis	Employee	No
voice015	31	F	reflux laryngitis	Researcher	No

4. Traitement du Signal Vocal

Notre démarche du traitement numérique des signaux vocaux à analyser comporte quatre étapes principales à savoir :

1. Chargement et Prétraitement du signal vocal,
2. Application des méthodes du traitement des signaux numériques à la détermination de la fréquence fondamentale F_0 ,
3. Extraction des paramètres statistiques.
4. Interprétation des résultats.

Les signaux vocaux n'étant pas périodiques la fréquence fondamentale change au cours du temps.

4.1. Application des méthodes du traitement des signaux numériques pour la détermination de la fréquence fondamentale F_0

Nous illustrons ci-dessous le traitement d'un signal importé de la base des données. Ce signal est d'un sujet sain, masculin de 55 ans, employé et fumeur noté : S1_SM55_F.

4.1.1. Tracé du signal

Le signal S1_M55_F est représenté sur la figure suivante :

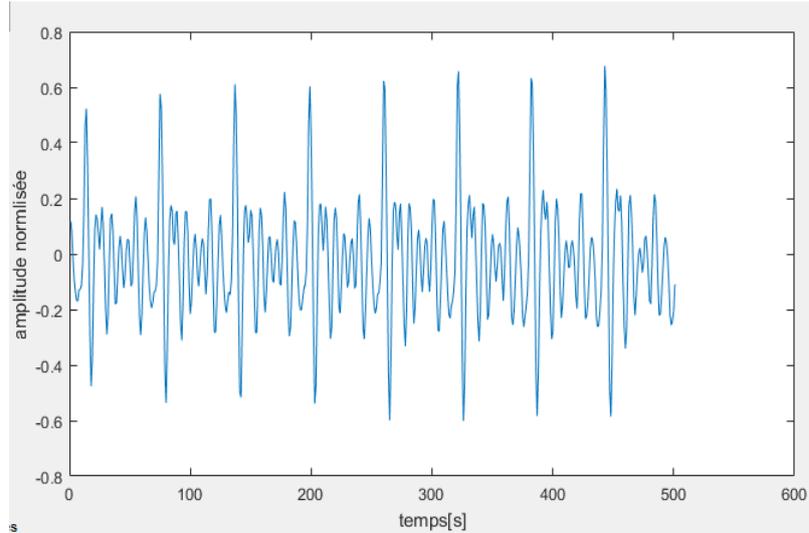


Figure III.1. Tracé du Signal vocal 'a' pour un sujet sain.

4.1.2. Filtrage

On a appliqué un filtre RII (fenêtre hamming) d'ordre 8

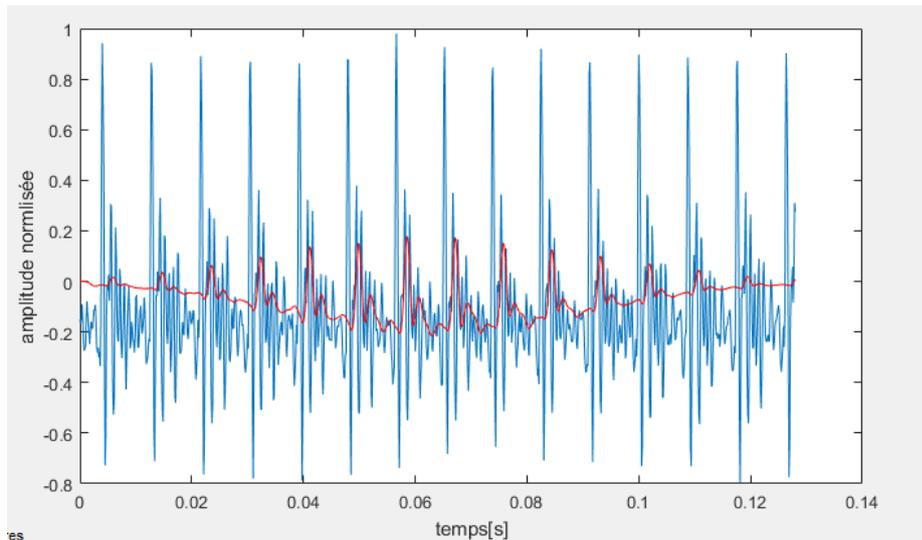


Figure III.2. Tracé du Signal vocal 'a' et son signal filtré.

4.1.3. FFT :

La transformation de Fourier, occupe une place privilégiée dans la théorie et le traitement des signaux. Néanmoins, la nature même de cette transformation ne peut donner d'un signal temporel qu'une information fréquentielle globale de laquelle toute notion de chronologie a disparu : son champ d'application naturel est l'analyse des signaux stationnaires. Ainsi, dès

lors que l'on considère des signaux modulés ou des processus non stationnaires, une analyse spectrale classique fournit une information moyennée sur la durée d'observation. Une solution partielle à ce problème, la plus intuitive, consiste à associer à un signal non stationnaire une suite de transformées de Fourier à court terme. Une deuxième solution, plus satisfaisante, consiste à chercher directement un outil adapté à l'étude de phénomènes non stationnaires c'est la représentation temps - fréquence

La figure (19) illustre le résultat de l'application de la Transformée de Fourier sur le signal choisis.

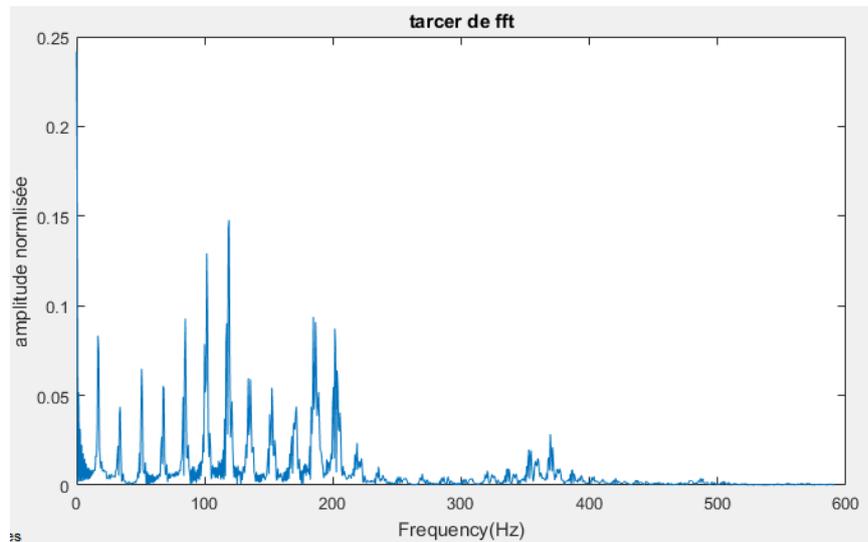


Figure III.3. Transformée de Fourier du signal vocal.

4.1.4. Détection des pics :

Nous avons utilisé un algorithme pour détecter le maximum des pics, en fixant un seuil au 2/3 de la valeur maximale.

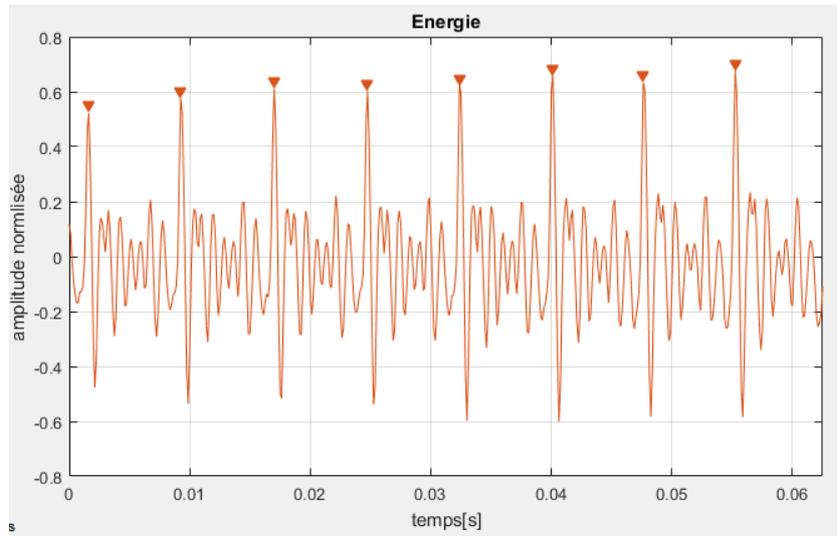


Figure III.4. Détections des pics

4.1.5. Calcul de l'énergie :

Le tracé de l'énergie du signal 'a' est illustré sur la figure ci-dessous :

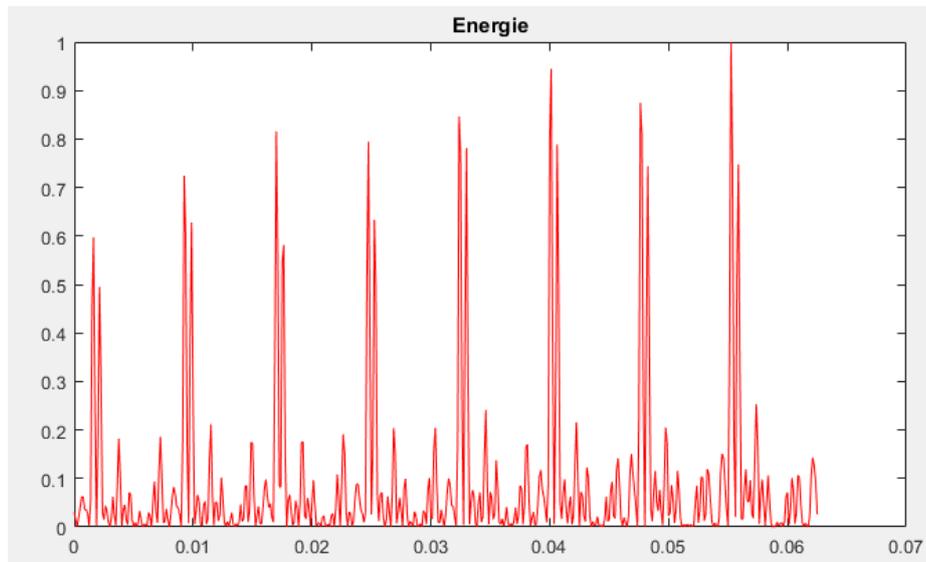


Figure III.5. Tracé de l'énergie du signal vocal de la lettre 'a'.

4.1.6. Autocorrélation :

La fonction d'autocorrélation nous permet d'estimer la Densité spectral de puissance moyenne de notre signal, afin d'extraire la fréquence fondamentale f_0 .

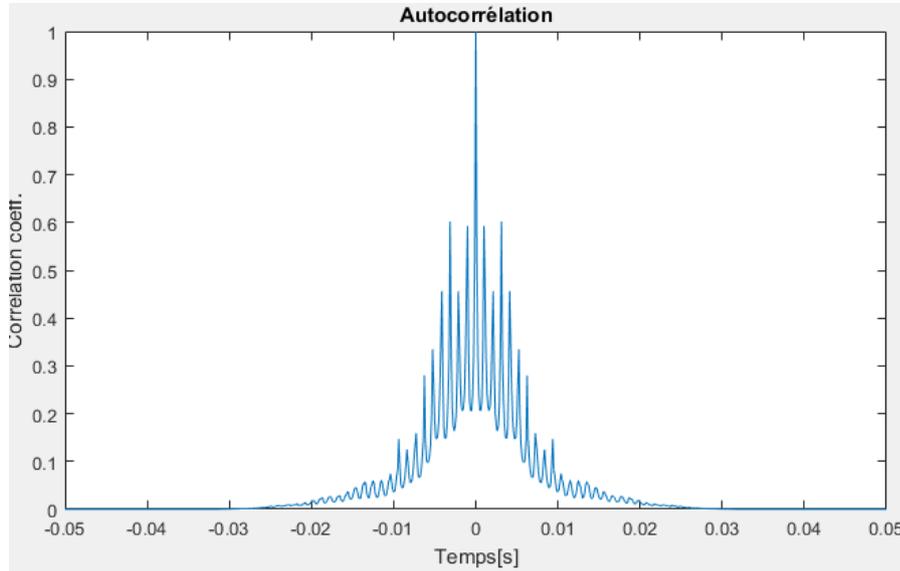


Figure III.6. Tracé de la fonction d'autocorrélation du signal vocal

4.1.7. Spectrogramme :

Nous avons appliqué l'algorithme de STFT pour représenter un signal parole 'a', dont la fréquence d'échantillonnage est : $F_s = 8\text{KHz}$. La Figure (23) représente la partie temps fréquence de ce signal vocal où nous pouvons comparer les résultats avec ceux de la FFT directement.

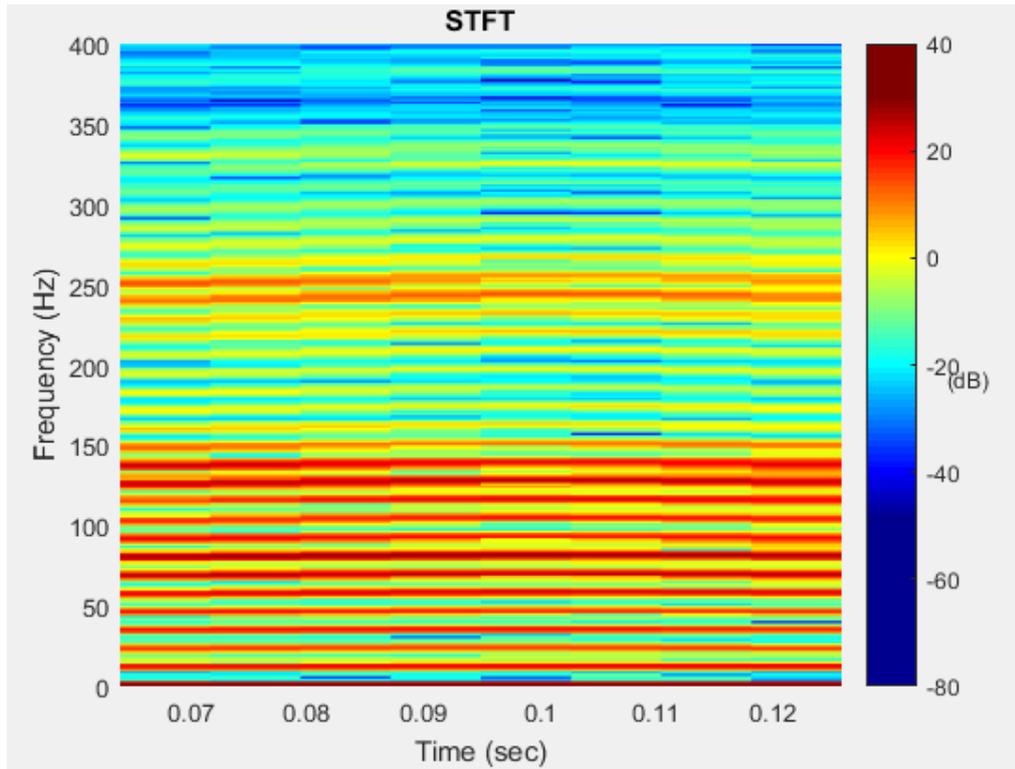


Figure III.7. Spectrogramme du signal vocal enregistré 'a'

5. Conclusion

Dans ce chapitre nous avons abordé notre interface développée permettant l'enregistrement et l'analyse spectro-temporelle du signal vocal 'a'.

Ce traitement numérique nous a permis de détecter les différents paramètres fréquentiels.

Dans le chapitre suivant, nous représenterons nos résultats pour les sujets pathologiques afin de les comparer.

IV. Chapitre 4 : Résultats et discussion

1. Introduction

Dans ce chapitre, nous présentons l'application de notre interface sur des signaux vocaux des sujets sains et malades (féminins et masculins) de différents âges basée principalement sur la détermination de la fréquence fondamentale, utilisant trois méthodes à savoir : la FFT de notre signal, la DSP et extraction de la fréquence dans le domaine temporelle. Cette fréquence est primordiale afin d'extraire d'autres paramètres qui sont cités par la suite. Nous présentons l'analyse et l'interprétation des résultats obtenus sous forme de tableaux ainsi que les performances de chaque méthode.

2. Fréquence fondamentale

La fréquence fondamentale est définie par le nombre de cycles vibratoires (ouverture-fermeture) par secondes. Elle s'exprime en Hertz et correspond à la hauteur de la voix et permet de produire des sons graves ou aigus. Elle varie en fonction de la longueur des cordes vocales, de leur épaisseur et de leur tension. Cette fréquence augmente lorsque la pression sous glottique augmente, et que le larynx remonte dans le cou en raccourcissant les dimensions du pharynx, produisant une augmentation de la tension et de la longueur des cordes vocales. La fréquence fondamentale diminue avec l'âge, en moyenne, elle est de 90 à 110 Hz en voix parlée chez l'homme et de 200 à 220 Hz chez la femme.

3. JITTER

Le JITTER est une mesure des perturbations de la fréquence fondamentale du signal sonore. On se place ici au niveau de la période du son, donc du cycle vibratoire, et on observe les différences de durée entre une ou plusieurs périodes. Le JITTER décrit un désordre vocal, causé selon les auteurs par un dysfonctionnement neurologique, aérodynamique ou bio mécanique.

Il est intéressant de noter que les perturbations de F_0 sont majorées à l'attaque et en final du son. Plusieurs auteurs préconisent donc d'analyser le JITTER à partir de la partie médiane de la voyelle.

4. SHIMMER

Le SHIMMER mesure la perturbation à court terme de l'amplitude de la fréquence fondamentale.

Comme pour le JITTER, il y a de multiples manières de calculer le SHIMMER. Nous décrivons ici les méthodes choisies pour calculer le SHIMMER : pour mesurer les perturbations à court terme de l'amplitude de signal sonore, on divise la moyenne des différences entre l'amplitude maximale de deux périodes successives, par la moyenne maximale de chaque période, en d'autres termes le SHIMMER permet d'estimer l'irrégularité de l'intensité d'un ensemble des cycles vibratoires. La valeur normale est de 0.4 [dB] et la limite Supérieure Pathologique est de 1.5 [dB]. [21]

Nous avons exploité notre interface afin d'étudier quatre signaux correspondant à l'enregistrement vocal de la lettre 'a' de quatre sujets, le premier est un homme sain, les trois autres ont différentes pathologies atteignant les cordes vocales, les résultats de chaque signal sont représentés et groupés dans des tableaux par la suite.

Cas : Homme sain, âgé de 55 ans

La figures suivantes représentent le tracé de la FFT de notre signal vocal, nous pouvons remarquer que la fréquence prédominante est de : $f_0=119\text{Hz}$,

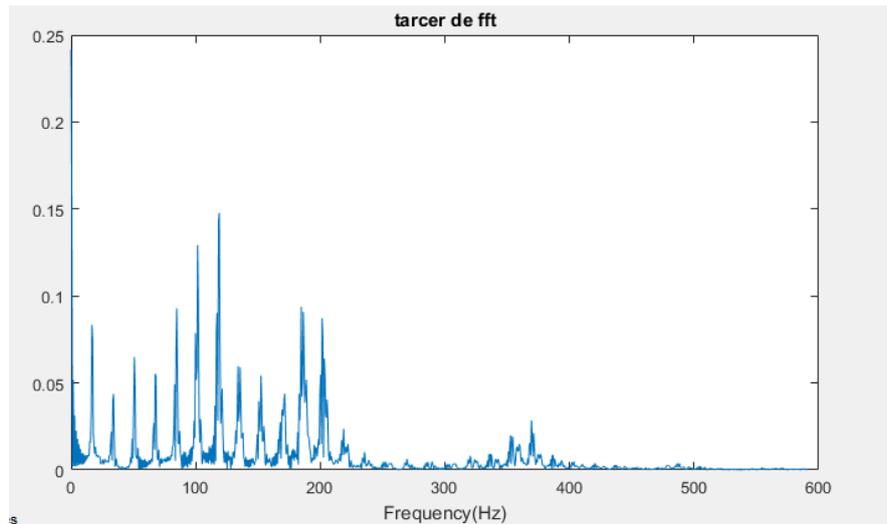


Figure IV.1. Tracé de la FFT

Le tracé suivant représente le tracé de l'énergie de notre signal.

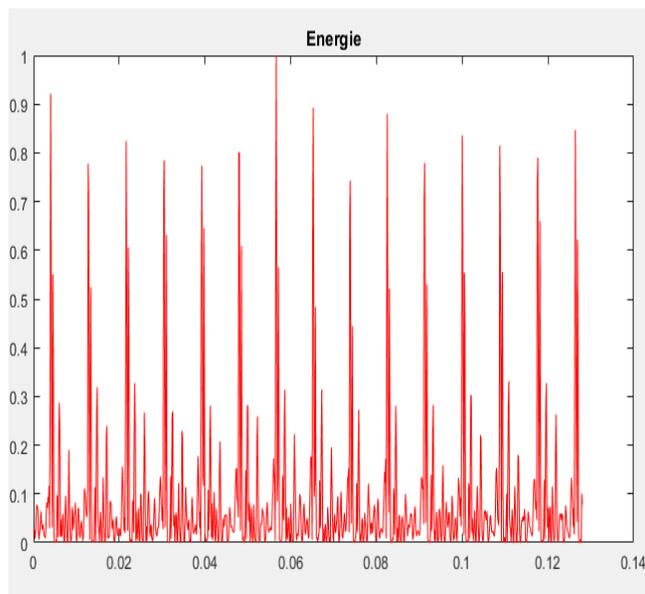


Figure IV.2. Tracé de l'énergie de notre signal.

La figure suivante représente le tracé de la fonction d'autocorrélation de notre signal, ce qui va nous permettre de calculer la densité spectrale de puissance moyenne (DSP).

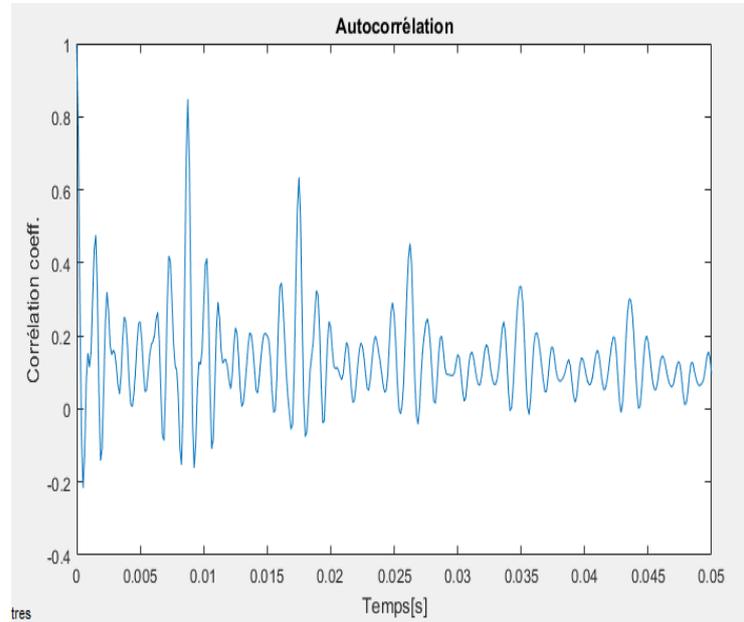


Figure IV.3. Tracé de la fonction d'autocorrélation du signal vocal.

la DSP de notre signal est représentée sur la figure suivante, où la fréquence prédominante est : $f_0=117\text{Hz}$.

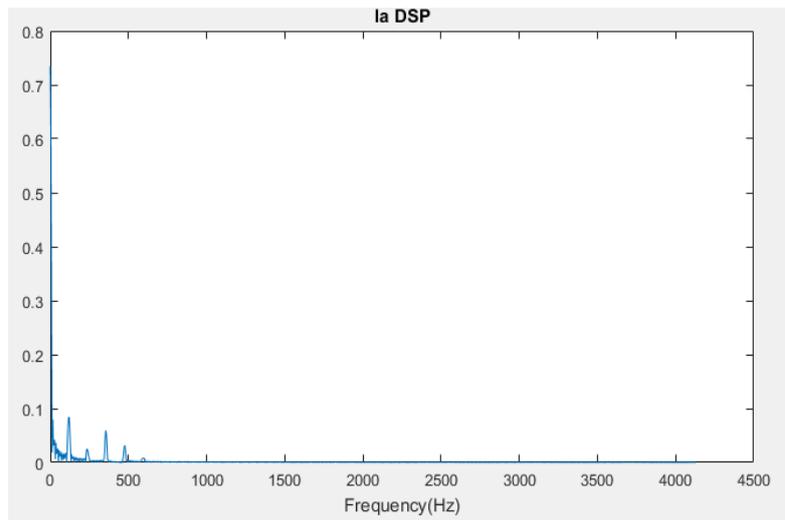


Figure IV.4. Tracé de la DSP du signal vocal.

La figure suivante illustre le tracé du spectrogramme, la STFT de notre signal où nous remarquons la présence d'un contenu spectral allant de 20 à 40dB pour des fréquences au voisinage de 100Hz.

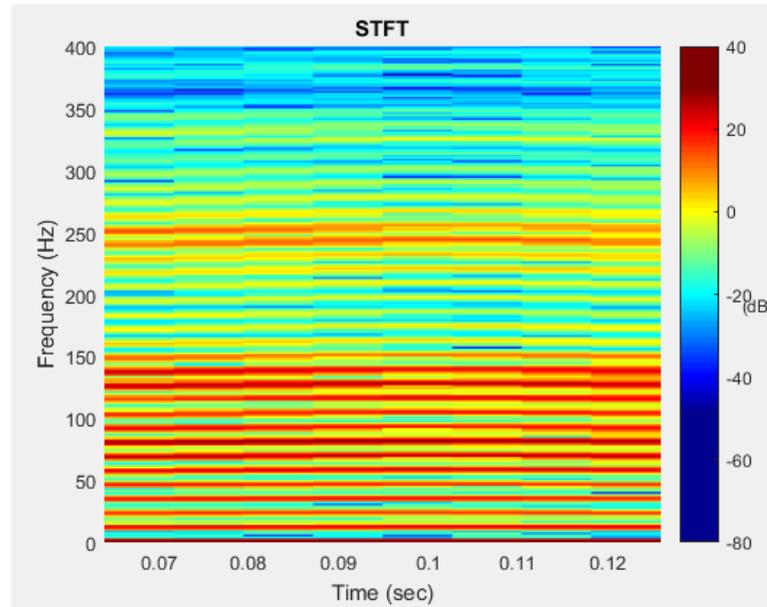


Figure IV.5. Tracé du spectrogramme.

Les résultats obtenus de la fréquence fondamentale f_0 et d'autres paramètres acoustiques sont représentés sur le tableau suivant :

Tableau IV-1. les résultats obtenus de la fréquence fondamentale.

Tableau des résultats				
Sujets		DSP	signal	FFT
Sujet1 : homme sain, 55 ans	f_0	117.09	123.03	119.25
	Ecart_Type	12.88	12.18	12.01
	Jitter_moy[Hz]	6.61	1.64	4.33
	Jitter_facteur (%)	5.65	1.33	2.11
	Shimmer_moy [dB]	3.13	0.57	3.21
Sujet2 : homme 34 ans, nodule	f_0	97.31	96.52	95.2
	Ecart_Type	1.13	1.09	1.1
	Jitter_moy[Hz]	0.59	0.48	0.55
	Jitter_facteur (%)	0.61	0.50	0.58
	Shimmer_moy [dB]	0.57	1.03	1
Sujet3 : femme 47 ans, hypokineticdysphonia	f_0	220.05	218.22	196
	Ecart_Type	4.62	4.50	4.30
	Jitter_moy [Hz]	2.35	2.31	2.20
	Jitter_facteur (%)	1.07	1.06	0.95
	Shimmer_moy [dB]	0.30	0.90	0.25
Sujet4 : femme 51 ans, reflux laryngites	f_0	213.20	211.35	209.5
	Ecart_Type	1.73	1.84	1.66
	Jitter_moy [Hz]	1.58	1.59	1.53
	Jitter_facteur (%)	0.74	0.75	0.68
	Shimmer_moy [dB]	0.97	0.93	0.84

Tableau IV-2. Résultats de valeurs de l'énergie.

Energie	Domaine temporel	Domain fréquentiel
Sujet1	105.863	105.974
Sujet2	104.812	109.849
Sujet3	99.0165	100.321
Sujet4	86.9989	89.4226

Ces tableaux représentent les résultats de la fréquence fondamentale, du Shimmer, du Jitter et de l'énergie des 4 signaux pour les différents sujets.

5. Conclusion :

Dans ce chapitre nous avons exploité certains signaux de la base donnée physionet afin de voir la variation de certains paramètres pertinent tels que la fréquence fondamentale, le shimmer, le jitter et l'énergie, nous estimons que les résultats obtenus sont satisfaisants, encourageants et comparable aux résultats des méthodes existant citées dans la littérature.

Conclusion générale

L'objectif principal de notre travail est de réaliser une interface graphique (un GUI) permettant de faire l'acquisition du signal vocal et d'établir son traitement numérique afin d'extraire des paramètres pertinents pouvant aider le médecin à établir un bon diagnostic d'abord par extraire les fréquences fondamentales et certains paramètres acoustiques importants dans le diagnostic des pathologies laryngées pour chaque sujet. Ces procédés reposent sur l'application d'outils de traitement du signal, à savoir la densité spectrale de puissance, fonction d'autocorrélation et transformée de Fourier de signal vocal.

Selon les résultats obtenus, nous pouvons affirmer que cette interface développée peut être utilisée par des cliniciens pour l'aide au diagnostic.

Plusieurs perspectives peuvent être envisagées comme l'augmentation du nombre de cas pathologiques à étudier, la conception d'une base de données et l'application d'autres méthodes pour calculer d'autres paramètres.

Bibliographie

- [1] Dr Hajj.A Place de la rééducation orthophonique dans la prise en charge des dysphonies (Etude prospective à propos de 42 cas).2011.
- [2] <http://www.ort-parc.com/Activite-ORL-Medicale-Pathologie-Voix-Cordes-Vocales>
- [3] <http://lapc.free.fr/tutorials/Troubles%20de%20la%20voix.pdf>
- [4] S.D.Courbiere médecin ORL, F.Resuel phoniatre, « la voix : la corde vocal et sa pathologie» .collège international de médecine et chirurgie de l'hôpital américain de paris, ISBN 2-905238-23-2.
- [5] A.Grovanni « la voix : la corde vocal et sa pathologie». Orientation diagnostique devant : trouble aigu de la parole. Professeur à la faculté de médecine de Marseille, Juin 2005.
- [6] A.Reyt « la voix : la corde vocal et sa pathologie. Cancers du larynx : Diagnostique, principes de traitement» .Professeur à la faculté de médecine de Grenoble, Mars 2005.
- [7] http://www.edimark.fr/phototheque/galerie_detail.php?id_galerie=998.
- [8] <http://blogopedie.canalblog.com/archives/2007/04/05/4535001.html>.
- [9] I.Atallah « la voix : la corde vocal et sa pathologie. Examen du larynx : de laryngoscopie a la stroboscopie », Université Joseph Fourier Science Technologie Médecine, CHU Grenoble.
- [10] A. Gilloire, M. Vetterli: "Adaptive filtering in sub-bands with application to Acoustic Echo Cancellation". IEEE TRANS ASSP, pp. 45-70.
- [11] A. Benallal : "Etude des algorithmes des moindres carrés transversaux rapides et application à l'identification de réponses impulsionnelles acoustiques". Thèse CNET/Université Rennes 1.
- [12] B. Hidayat : "Traitements fréquentiels à la prise de son multi microphones en vue de dé réverbération du signal de parole". Thèse CNET/Université de Rennes 1.
- [13] S. Dixon, M. Goto et M. Mauch, « Pourquoi la voix est-elle intéressante ? » dans Pro c. 16 ème Confi. Int. Society of Music Information Retrieval, 2015
- [14] O. Ait Amrane : "Identification des systèmes à réponse impulsionnelle longue par filtrage adaptatif en fréquence : application à l'annulation d'écho acoustique". These Telecom Paris.
- [15] G. Faucon et al: "Study and comparison of three structures for enhancement of noisy speech". ICASSP 89 Glasgow.
- [16] D. DECLERCQ, A. QUINQUIS, Le signal déterministe, collection signaux et systèmes en question.
- [17] D. DECLERCQ, A. QUINQUIS, Le signal aléatoire, collection signaux et systèmes en question.
- [18] D. DECLERCQ, A. QUINQUIS, Le filtrage des signaux, collection signaux et systèmes en question.
- [19] D. DECLERCQ, A. QUINQUIS, Détection et estimation des signaux, collection signaux et systèmes en question.
- [20] A. QUINQUIS, Le traitement du signal sous Matlab : pratique et applications, HERMES 2000
- [21] Institut numérique : L'EVALUATION SUBJECTIVE ET OBJECTIVE DE LA VOIX16/05/2013
- [22]MireaFarrus, Janvier Hernando, PascualEjarque, « Jitter and Shimmer Measurements for Speaker Recognition »,Universities de Catalunya, Espagne.
- [23] R.Cmejla , J.Rusz ,L. Vokral , « Laboratory tasks from Voice Analysis in the Study of Biomedical Engineering using Matlab »,Czech Technical University in Prague
- [24]Abdelouahed Sara, « Analyse Spectro-Temporelle du Signal Vocal En Vue De Dépistage et du suivie des Dysphonies Chroniques d'origine Laryngées »Mémoire de Master, **Université** Abou BakrBelkacem, Tlemcen, Juin2012.

[25] Abdelouahed Sara, « Etude et Réalisation d'une Plateforme Télé médecine dédiée à l'évaluation objective et au Suivi des dysphonies chroniques laryngée par analyse spectro-temporelle du signal vocal », Thèse de Doctorat, Université Abou Bakr Belkacem, Tlemcen.

[26] Fadoua BAHJA, Détection du fondamental de la parole en temps réel : application aux voix pathologiques, thèse de doctorat, université mohammed v – agdal, faculté des sciences rabat. Soutenue le 09/07/2013

[28] Melle Amel FENIZA, Melle Hassiba LAISSANI « Analyse multivariée de la parole pour l'évaluation objective de la qualité vocale », Université Mohammed Seddik BENYAHIA – Jijel, Juillet 2019

[29] AJGOU Riad, « Techniques de détection de la période du pitch par les méthodes temps fréquence et temps échelle », Université de Biskra, 11 /03 /2010.

[30] [Damien Lolive](#), « Transformation de l'intonation : application à la synthèse de la parole et à la transformation de voix », IRISA, INRIA Rennes, ENSSAT - École Nationale Supérieure des Sciences Appliquées et de Technologie.

[31] L'allouai Abdelkrim, « débruitage d'un signal de la parole corrompu par un bruit coloré en utilisant la transformée en ondelettes et implantation sur un processeur de traitement numérique des signaux », Montréal, École de technologie supérieure, 2004.

[32] <http://fr.ap-hm.fr/site/orl-pediatrique/soins>