



جامعة ابوبكر بلقايد - تلمسان

Université Abou Bakr Belkaïd de Tlemcen

Faculté de Technologie

Département 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 : Maati Nacer

Soutenu le 28 septembre 2022

Analyse Cepstrale du signal vocal

M.	Taouli Sid Ahmed	<i>MCA</i>	Université de Tlemcen	Président
Mme	Ziani Cherif Souhila	<i>MCA</i>	Université de Tlemcen	Encadreur
Melle	Bekhti Fadia	<i>MCA</i>	Université de Tlemcen	Examineur

Année universitaire 2021-2022

Sommaire

Abréviations	5
Résumé.....	6
Introduction.....	8
Chapitre 1 : Etat de l'art sur le signal vocal.....	9
Introduction.....	10
1 Système de production de la parole	10
1.1 Partie sub-glottique (l'appareil respiratoire)	10
1.1.1 Le diaphragme	10
1.1.2 Les poumons.....	10
1.1.3 La trachée	10
1.1.4 Mécanisme	11
1.2 Partie glottique (Le larynx)	11
1.2.1 Les cartilages du larynx.....	11
1.2.2 Les muscles du larynx	13
1.2.3 Les cordes vocales	13
1.3 Partie supra-glottique	14
2 Mécanisme de la production de la parole	14
3 Caractéristique de la production de la parole	15
4 Les pathologies	16
4.1 Inflammation aiguë du larynx ou laryngite aiguë	16
4.2 Laryngite chronique	16
4.3 Tumeurs du larynx.....	17
4.3.1 Tumeurs bénignes	17
4.4 Tumeurs malignes du larynx :.....	18
Conclusion	19
Chapitre 2 : Analyse objective de la voix	21
Introduction.....	22
5 La transformée de Fourier	23
6 L'analyse cepstrale	23
6.1 Préaccentuation :	24
6.2 Segmentation :.....	24

Sommaire

6.3	Chevauchement :	24
6.4	Fenêtrage :	25
6.5	Le Cepstre :	28
6.6	La prominence du pic cepstral (CPP).....	28
	Conclusion	30
	Chapitre 3 : Résultats obtenus et discussions.....	31
	Introduction.....	32
1	Présentation du Programme	32
2	Traitement du signal numérique par exécution de la fréquence fondamentale par la méthode cepstral	32
2.1	L'affichage du signal d'entrée :	32
2.2	Segmentation et filtrage :	33
2.3	Formation du cepstre :.....	33
2.4	Calcul du cepstre en dB et détection du pic cepstral	35
	Conclusion :	37
	Conclusion générale.....	38
	Références.....	39

Table des figures

FIGURE 1:SCHEMA DE L'APPAREIL VOCAL [30].	11
FIGURE 2: EXEMPLE D'UN SIGNAL DE PAROLE VOISEE ET NON VOISEE [33].	15
FIGURE 3: LARYNX [35].	12
FIGURE 4: COUPE TRANSVERSALE DES CORDES VOCALES [30].	14
FIGURE 5: ANATOMIE DE LA PARTIE SUPERIEURE DE L'APPAREILLE PHONATOIRE [30]	14
FIGURE 6:LARYNGITE AIGU [30]	16
FIGURE 7: LARYNGITE CHRONIQUE [30]	16
FIGURE 8:POLYPE DE CORDES VOCALES [30]	17
FIGURE 9:LES NODULES DES CORDES VOCALES [30]	18
FIGURE 10: GRANULOMES [30]	18
FIGURE 11: CANCER DU LARYNX [30]	19
FIGURE 12: SIGNAL AUDIO [34].	22
FIGURE 13:REPONSE D'UN FILTRE DE PREACCENTUATION POUR DES DEFERENTES VALEURS D'ALPHA [29]	24
FIGURE 14: ANALYSE SPECTRALE A COURT TERME [29]	25
FIGURE 15: FENETRES LES PLUS UTILISEES [35].	27
FIGURE 16: SCHEMA ILLUSTRANT ALLURES FREQUENTIELLES DE QUELQUES FENETRES [35].	27
FIGURE 17: SCHEMA ILLUSTRANT LES COMPOSANTES DU CEPSTRE ET DE SA LIGNE DE REGRESSION POUR DETERMINER LA CPP DANS UN SIGNAL QUASI-PERIODIQUE.[31]	28
FIGURE 18:SIGNALE VOCAL	32
FIGURE 19:SIGNAL VOCAL PRES TRAITE	33
FIGURE 20: CEPSTRE DU SIGNAL VOCAL 'A'	34
FIGURE 21:TRACEE DU CEPSTRE AVEC LA LIGNE DE REGRESSION	35

Abréviations

FFT : La Transformée de Fourier Rapide (Fast Fourier Transforme)

DFT : La Transformée de Fourier Discrète (Discret Fourier Transforme)

Fs : La Fréquence d'échantillonnage (Sampling Frequency)

CPP : La Prominence du Pic Cepstrale (Cepstral Peak Prominence)

F0 : La fréquence fondamentale

Résumé

L'appareille phonatoire est exposé à de nombreuses maladies pouvant atteindre le cancer, les maladies les plus fréquentes sont les maladies d'origine laryngée. Pour arriver à faire un diagnostic objectif qui va avoir l'avantage d'être non invasif le traitement de signal vocal présente un outil préférable pour limiter l'incidence de cette maladie. Le signal vocal porte des paramètres acoustiques relatifs à l'état morphologique de l'appareil phonatoire, avec lequel nous pouvons distinguer un larynx sain d'un autre pathologique.

Ce travail est dédié à faire une analyse cepstrale du signal vocal afin de réaliser un algorithme ayant pour but d'identifier les maladies laryngées.

Ce mémoire fournit une introduction au sujet de l'analyse du signal vocal. Il examine les fondements et l'état de l'art du signal vocal et son traitement à travers trois grandes parties: l'anatomie vocale, la décomposition de la voix et le traitement numérique du signal permettant à travers l'analyse cepstrale en exploitant la fréquence fondamentale pour le calcul de certains paramètres pertinents.

Nous visons à établir un point de départ pour les médecins afin de les aider à établir un bon diagnostic ainsi que dans la procédure thérapeutique et le suivie du patient.

Mots clés : Signal vocal, TFR, TFInverse, analyse cepstrale, fréquence fondamentale, Qfrence.

ملخص

الجهاز الصوتي البشري معرض للإصابة بالعديد من الأمراض التي يمكن أن تصل إلى السرطان، ومن أكثر الأمراض شيوعاً أمراض الحنجرة.

للوصول إلى تشخيص موضوعي يتمتع بميزة كونها غير جراحية، تقدم معالجة الإشارات الصوتية أداة مفضلة للحد من حدوث هذه المرض.

تحمل الإشارة الصوتية معلومات صوتية تتعلق بالحالة المورفولوجية للجهاز الصوتي، والتي يمكننا بواسطتها التمييز بين الحنجرة السليمة والحنجرة المصابة بالمرض.

هذا العمل مخصص لإجراء تحليل طيفي زمني للإشارة الصوتية من أجل إنشاء خوارزمية تهدف إلى تحديد أمراض الحنجرة.

توفر هذه المدكرة مقدمة لموضوع تحليل الإشارات الصوتية. يفحص أسس وطريقة معالجتها من خلال ثلاثة أجزاء رئيسية: تشريح الصوت، وتحليل ومعالجة الإشارات الرقمية بإجراء تحليل طيفي وحساب بعض العناصر ذات الصلة.

نهدف الى إنشاء نقطة انطلاق للأطباء تساعدهم على القيام بالتشخيص بدون الحاجة الطرق الجراحية.

Summary

The phonatory apparatus is exposed to many diseases that can reach cancer, the most frequent diseases are those of laryngeal origin. In order to make an objective diagnosis, which has the advantage of being non-invasive, voice signal processing is a preferable tool for limiting the incidence of this disease. The vocal signal carries acoustic parameters related to the morphological state of the phonatory apparatus, with which we can distinguish a healthy larynx from a pathological one.

This work is dedicated to make a spectro-temporal analysis of the vocal signal in order to realize an algorithm aiming at identifying laryngeal diseases.

This thesis provides an introduction to the subject of speech signal analysis. It examines the basics and the state of the art of voice signal and processing through three main parts : voice anatomy, voice decomposition and digital signal processing allowing cepstral analysis by exploiting the fundamental frequency for the calculation of some relevant parameters.

We aim to establish starting point for doctors to help them in the diagnosis, treatment and follow-up of the patient.

Keywords: Speech Signal, FFT, Inverse FT, Cepstral Analysis, Fundamental Frequency, Qfrence

Introduction

Parler est le meilleur moyen de s'exprimer dans les cercles sociaux. Nous pouvons sortir nos sentiments et montrer nos attitudes, ce qui présente un véritable profil de notre identité, c'est l'outil d'affirmation de soi. L'être humain risque de perdre cette fonctionnalité, plusieurs maladies liées au larynx, jusqu'aux cancers, qui peut finir par l'ablation de ce dernier. [1][2]

Les maladies laryngées sont nombreuses, ceux qui ont un risque limite telle qu'une laryngite aigue qui peut être causée par un simple forçage de la voix. D'autres maladies graves, sont à l'origine de dommages plus importants tels que les tumeurs malignes.

Ces cancers sont fréquents, directement liés au tabagisme chronique et souvent associés à une consommation excessive d'alcool, touchant l'homme dans l'immense majorité des cas. [3]

Le diagnostic subjectif à travers l'oreille du clinicien basée sur la dégradation de la voix. La recherche des causes utilise des méthodes objectives où l'inconvénient majeur est son invasivité.

Le signal de la parole porte beaucoup d'information, données relatives à la personne qui parle. Le traitement numérique de ce signal peut être utilisée par les phoniâtres afin d'améliorer la qualité de la voix, Et de connaître les causes des dysphonies

Le travail présenté s'articule autour du traitement du signal vocal humain afin de réaliser une analyse numérique sur ce signal et de distinguer les sons pathologiques, les résultats de cette analyse vont aider à faire un diagnostic initial non invasif.

Nous avons réparti notre mémoire en trois chapitres, comme suit :

- Le chapitre I : Etat de l'art sur le signal vocal.
- Le chapitre II : Analyse objective de la voie.
- Chapitre III : Résultats obtenus et discussions.

Chapitre 1 : Etat de l'art sur le signal vocal

Introduction

Un son est généré par un objet vibrant, dans le cas de l'être humain ce sont les cordes vocales qui produisent ces vibrations.

Ces vibrations provoquent des déplacements et des oscillations de molécules d'air, ce qui entraîne des régions locales de compression et de raréfaction.

La pression alternative se propage dans l'air comme une onde, de sa source à un auditeur ou à un microphone. Dans ce chapitre nous allons nous intéresser principalement l'état de l'art du signal vocal.

1 Système de production de la parole

La production de la parole se fait par combinaison de plusieurs organes, qui sont indispensables pour la phonation, et ainsi pour d'autres usages dans les systèmes digestif et respiratoire.

La phonation est un acte volontaire, résultat de transformation de souffle d'air expirée le long de la conduite vocale. Après expiration, l'air des poumons passe par deux phases, la phase de modification en vibration au niveau de la partie glottique (Le larynx), et la phase de filtration au niveau de la partie supra-glottique. [3]

1.1 Partie sub-glottique (l'appareil respiratoire)

Le moteur de l'appareil phonatoire assure l'énergie nécessaire pour la production du son, comporte comme une souffleuse [4][3]. Cette partie sub-glottique comprend trois organes : le diaphragme, les poumons et la trachée.

1.1.1 Le diaphragme

Il se trouve en dessous des poumons, constitue une séparation entre l'abdomen et le thorax. Le diaphragme commande le phénomène d'aspiration et expiration d'air, c'est le moteur principal avec sa contraction il crée une pression négative dans le thorax, par conséquent il assure une pénétration d'air via la trachée puis son relâchement fait l'inverse afin d'alimenter l'appareil résonateur.

1.1.2 Les poumons

Organes respiratoires situés de chaque côté du Médiastin participe dans la production de la parole comme réservoir d'air nécessaire pour la soufflerie qui va faire vibrer les cordes vocales.

1.1.3 La trachée

Formée de 16 à 20 anneaux cartilagineux fermés en arrière par une séparation postérieure en membrane contient également le muscle trachéal. La trachée c'est le conduit aérien qui assure le passage d'air des poumons jusqu'aux cordes vocales.

1.1.4 Mécanisme

La compression des poumons fait expirée l'air a une pression nécessaire et fait vibrer les cordes vocales, c'est les poumons qui alimentent l'appareil phonatoire par l'énergie pour la phonation. Il constitue l'appareil biomécanique où le diaphragme exerce une force pour comprimer les poumons, et qui va par conséquence libérer un souffle permettant de créer des sons.

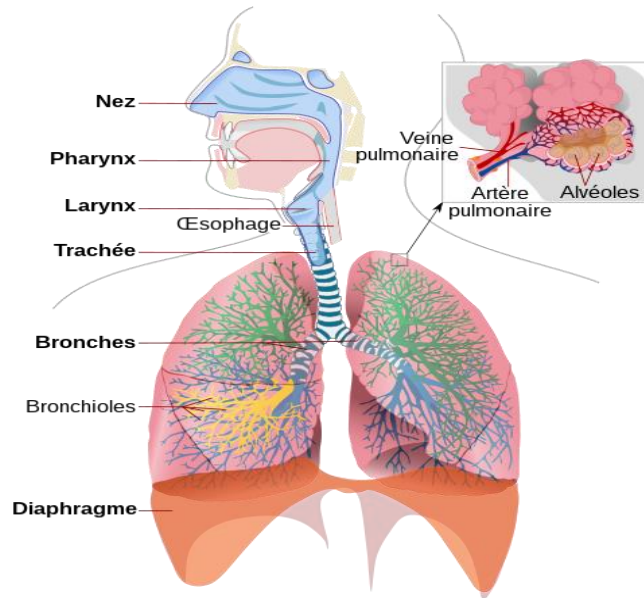


Figure 1:Schéma de l'appareil vocal [30].

1.2 Partie glottique (Le larynx)

Structure cartilagineuse contient 11 cartilages : trois grands cartilages impaires (cricoïde, thyroïde et épiglote) et 4 paires (aryténoïde, cornicule et cunéiforme), des muscles et articulations qui garantissent la mobilité de ces parties. Les cordes vocales en nombre de deux constituent l'organe important pour transformer l'air expirée par les poumons en des sons voisés ou non voisés.

Le larynx aussi joue un rôle de protection des voix respiratoires, il empêche le passage des aliments par l'épiglotte qui se ferme lors de déglutition. [4][5]

1.2.1 Les cartilages du larynx

1.2.1.1 Cartilages impairs

1.2.1.1.1 Cartilage thyroïdien

Le cartilage thyroïdien est le plus grand des cartilages laryngés composé de cartilage hyalin. Il forme les parties antérieures et latérales du larynx et n'a pas de composante postérieure. Les larges moitiés plates droite et gauche (lamines) du cartilage fusionnent antérieurement sur la ligne médiane pour former une projection antérieure en forme de V appelée proéminence laryngée (communément appelée "pomme d'Adam").

1.2.1.1.2 Cartilage cricoïde

Le cartilage cricoïde est un cartilage hyalin beaucoup plus petit, en forme d'anneau de bague, situé directement sous le cartilage thyroïdien. Il forme la face inférieure du larynx et est relié à la trachée par la face inférieure. C'est le seul anneau complet de cartilage qui entoure les voies respiratoires.

1.2.1.1.3 L'épiglotte

L'épiglotte, est un grand cartilage élastique en forme de feuille qui est recouvert d'une muqueuse. Elle est attachée par son pédoncule à la face interne de l'angle formé par les lamelles du cartilage thyroïdien via le ligament thyroépiglottique sur la ligne médiane.

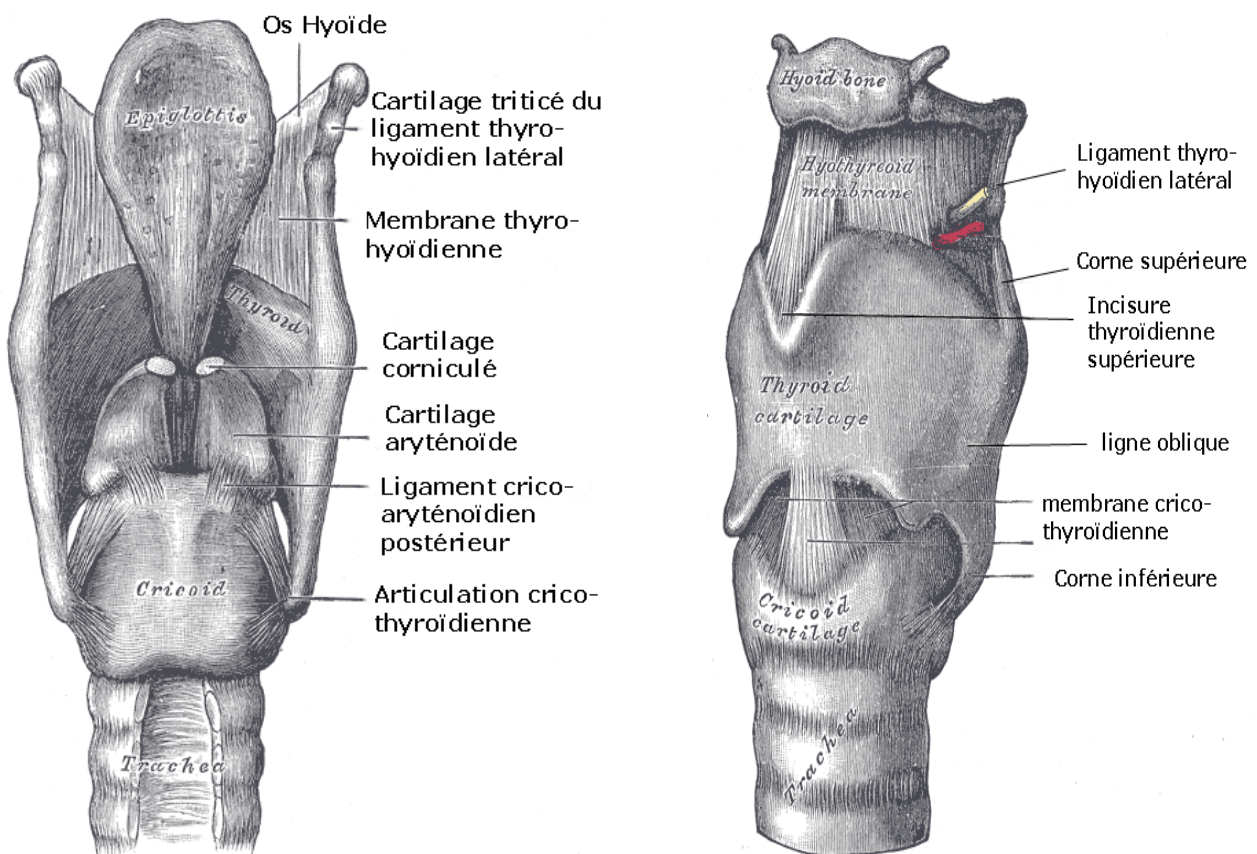


Figure 2: Larynx [35].

Chapitre 1

1.2.2 Les muscles du larynx

Ces muscles produisent les mouvements du larynx et de ses cartilages, permettant ainsi la bonne conduction de l'air, la parole, les mouvements de l'épiglotte et la protection des voies respiratoires. Les muscles du larynx sont divisés en deux catégories :

1.2.2.1 Muscles intrinsèques

Les muscles intrinsèques déplacent les cordes vocales afin de produire les sons de la parole, fonctionnellement, ils sont divisés en :

➤ Adducteurs

Crico-aryténoïdiens latéral, aryténoïde transversal.

➤ Abducteurs

Crico-aryténoïde postérieur.

➤ Sphincters

Aryténoïde transversal, aryténoïde oblique, aryépiglottique.

➤ Muscles qui tendent les cordes vocales

Crico-thyroïde.

➤ Muscles qui détendent les cordes vocales

Thyroaryténoïde, vocalis.

1.2.2.2 Les muscles extrinsèques

Les muscles extrinsèques du larynx sont ceux qui sont attachés d'une manière ou d'une autre à l'os hyoïde, que ce soit par leur origine ou leur insertion, et qui déplacent donc le cartilage thyroïdien. Les muscles infrahyoïdiens font partie de la partie inférieure du larynx et s'y attachent, ainsi qu'à la face inférieure de l'os hyoïde. Ce groupe de muscles comprend le muscle sternohyoïdien, le muscle omohyoïdien, le muscle sternothyroïdien et le muscle thyrohyoïdien qui travaillent à abaisser le larynx et l'os hyoïde.

1.2.3 Les cordes vocales

Les cordes vocales sont formées par des ligaments vocaux, délimités par des muscles sur lesquels on peut intervenir, le tout est recouvert d'une muqueuse qui forme la partie vibrante qui produit le son. Elles forment un clapet qui peut être ouvert ou fermé. Elles sont tendues de l'angle rentrant du cartilage thyroïde à l'apophyse vocale des cartilages aryténoïdes. La glotte est l'espace plus ou moins grand entre les deux cordes vocales.

- Pour produire des sons voisés les cordes se ferment pour provoquer une élévation de la pression dans la partie sub-glottique, l'exertion d'une pression supplémentaire sur ces deux cordes par l'expiration d'air d'une façon continue les oblige à s'ouvrir et à se fermer de manière répétitive, les faisant vibrer. Plus la pression exercée est supérieure plus l'intensité du son produit est élevée.

Il contrôle le passage d'air au moment de la respiration ou la phonation, Lors de la respiration, ils sont complètement ouverts formant la forme d'une glotte.

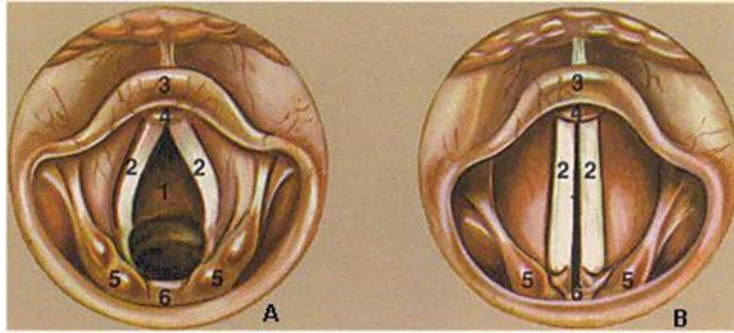


Figure 3: Coupe transversale des cordes vocales [30]

1.3 Partie supra-glottique

Formé par deux cavités : une cavité buccale ou sa forme géométrique dépend de la position de la langue et de positionnement de la mâchoire et des lèvres, et l'autre nasale, indéformable. Ces deux cavités constituent le conduit vocal. [6][4]

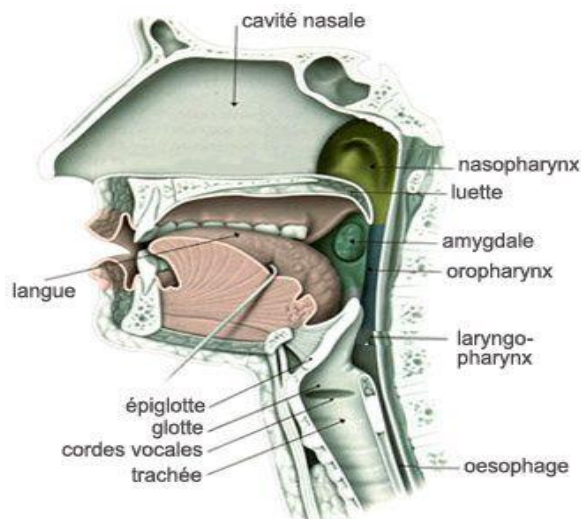


Figure 4: Anatomie de la partie supérieure de l'appareil phonatoire [30]

2 Mécanisme de la production de la parole

Physiquement la voix humaine est une variation de pression de l'air expiré qui passe par un système d'articulation volontairement comprimé par l'action du diaphragme.

L'air par la suite rencontre les cordes vocales où on distingue deux situations possibles, soit que les cordes vocales sont écartées, ce qui permet de faire passer l'air et de produire du bruit.

La deuxième situation pour des cordes vocales fermées, elles sont soumises à une force de pression qui les met en vibration d'où la création d'un son quasi périodique.

Le conduit vocal constitue le dernier passage pour l'air provenant des poumons avant de se propager dans l'atmosphère, et par rapport à sa forme géométrique le timbre du son est ainsi défini. [7][8]

3 Caractéristique de la production de la parole

Malgré les nombreuses différences entre les individus, et l'existence de nombreuses langues, la parole suit des modèles qui ont des caractéristiques bien définies en termes d'intensité sonore, de distribution de fréquence, de taux de hauteur et de taux syllabique.

Les trois principaux attributs des communications vocales, à savoir la génération, la propagation et la réception, correspondent bien aux caractéristiques de la parole, et la réception, correspondent bien aux caractéristiques physiques de la parole elle-même.

La génération fait référence à l'appareil de production de la parole décrit précédemment. Propagation fait référence à l'environnement physique et au fait que les fréquences vocales se propagent dans l'air avec des niveaux d'absorption et d'amortissement assez faibles, l'air avec des niveaux d'absorption et de distorsion assez faibles.

Comme nous l'avons vu plus haut, les caractéristiques de la parole s'adaptent en fonction de l'environnement, des limites de l'audition et de la production vocale et du contexte (situation). A court terme, nous réagissons à la distance, au bruit de fond, à l'émotion, à l'urgence, au contenu du message et aux indices de communication non verbale. Alors que la plupart de nos communications verbales utilisent la normale, les cris et les chuchotements sont relativement courants. [9][10].

La figure suivante représente l'exemple d'un signal de parole :

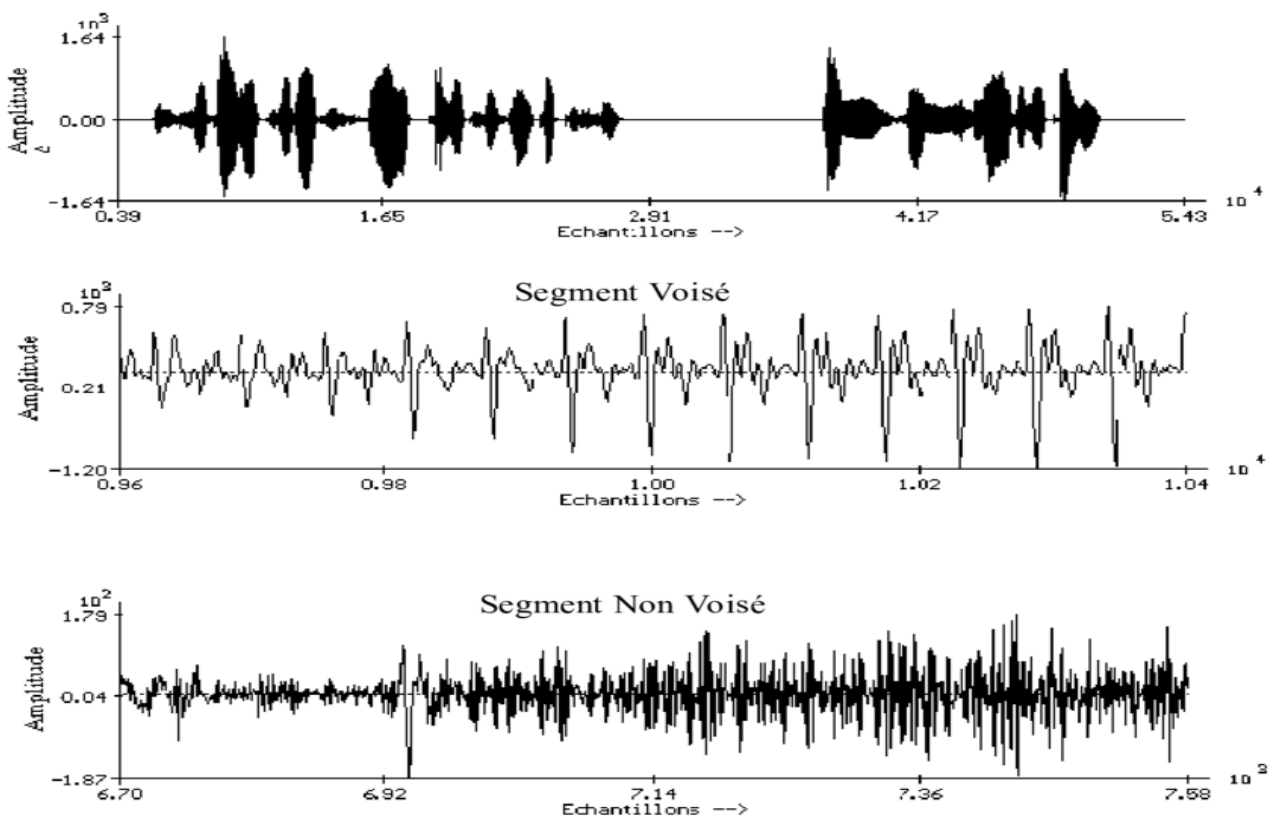


Figure 5: Exemple d'un signal de parole voisée et non voisée [33].

4 Les pathologies

4.1 Inflammation aiguë du larynx ou laryngite aiguë

L'inflammation aiguë du larynx est causée par l'effet des bactéries où le moindre refroidissement. Les bactéries introduites par le nez visent directement le larynx. La laryngite aiguë peut être l'effet de d'autres infections telles que la typhoïde et la tuberculose. Cette inflammation peut être la conséquence d'un forçage de la voix.

Par conséquent le laryngoscope indique un gonflement et une rougeur au niveau des cordes vocales qui sont naturellement blanches, chez les enfants leur gonflement rend la respiration difficile. [11]



Figure 2: Laryngite aiguë [30]

4.2 Laryngite chronique

La laryngite chronique est reconnue par un gonflement des cordes vocales, la couleur est gris-rouge, accompagnée de mucus et de pus. Cette maladie peut être le prolongement de laryngite aiguë mal soignée, ou même le forçage continu de la voix. Le passage de la poussière, le tabac et l'alcool sont aussi des causes et conduisent à cette maladie. [11][12]



Figure 3: Laryngite chronique [30]

4.3 Tumeurs du larynx

Le larynx risque de développer des tumeurs qui peuvent être malignes ou bénignes.

4.3.1 Tumeurs bénignes

4.3.1.1 Polypes

Les polypes des cordes vocales sont souvent secondaires à un malmenage vocal, le polype peut être favorisé par une intoxication due à l'alcool et au tabac ou à une exposition aux poussières. Il s'agit d'une tuméfaction bénigne de la corde vocale sur le 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. [13]



Figure 4: Polype de cordes vocales [30]

4.3.1.2 Nodules

Les nodules des cordes vocales se traduisent par un gonflement des cordes vocales localement, de façon le plus souvent bilatérale (des deux côtés) au niveau du bord libre de la corde vocale.

Ce gonflement des cordes vocales est localisé à la jointure du premier et du deuxième tiers des cordes vocales et bloque le contact entre ces cordes au cours de la phase de phonation.

La création de nodules est la manifestation de chocs répétitifs des cordes vocales entre elles, à cause d'une surutilisation ou utilisation défectueuse de la voix.

Certains éléments sont également susceptibles de favoriser une telle manifestation: infections récurrentes au niveau des conduits respiratoires supérieurs (rhinite, sinusite), bronchite, asthme, allergies, une forte exposition aux fumées de cigarettes, un reflux laryngo-pharyngé, une atmosphère chargée en produits toxiques (poussières, allergènes) [15][14].



Figure 5: Les nodules des cordes vocales [30]

4.3.1.3 Granulome

Le granulome est une pathologie de type inflammatoire, elle perturbe la fermeture des cordes vocales ce qui provoque une dysphonie lorsque les structures inflammatoires sont volumineuses, ils se révèlent un malaise pharyngé. Généralement, il est causé par un dépassement vocal, il est découvert par l'examen laryngé systématique.

Les granulomes sont unis ou bilatéraux, en général développés dans la partie postérieure de l'orifice glottique [16].



Figure 6: Granulomes [30]

4.4 Tumeurs malignes du larynx :

Ce cancer se manifeste par une petite tumeur sur une corde vocale, puis elle évolue progressivement, ce qui entraîne une modification de la voix jusqu'à la rendre étouffée.

Chapitre 1

Quant aux tumeurs malignes, le cas le plus grave est celui du cancer du larynx. Ce type est très grave nécessite une intervention chirurgicale sinon il peut finir par une paralysie des cordes vocales [17].

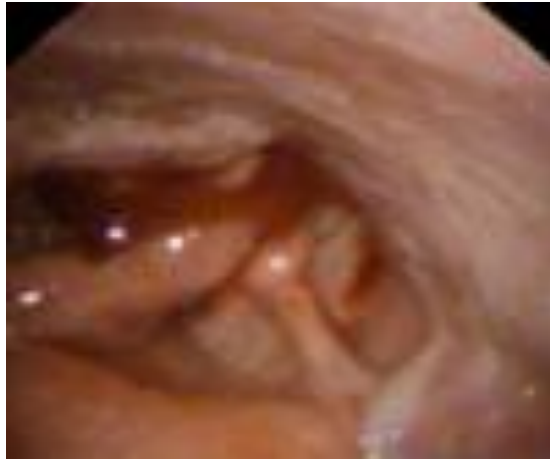


Figure 7: Cancer du larynx [30]

Conclusion

Ce chapitre a été consacré à l'état de l'art du signal vocal, la façon dont le son vocal est généré depuis l'appareil phonatoire. Nous avons abordé les différents organes ainsi que les différentes pathologies pouvant affecter ce système. Le chapitre suivant sera consacré à l'analyse objective de la voie.

Chapitre 2 : Analyse objective de la voix

Introduction

Afin d'extraire les informations portées par le signal vocal plusieurs méthodes d'analyses sont utilisées, telle que l'analyse cepstral.

Le signal vocal est une source pleine d'information relative à la personne et l'environnement.

Nous nous sommes intéressés à la portion qui porte les informations sur le conduit vocal du locuteur.

Le signal

Le signal est le vecteur de l'information diffusée par une source et adressée à un récepteur, Il sert à transporter les ordres dans les systèmes de contrôle, il assure le transit de l'information, particulièrement délicat et doit être traité de manière soigneuse. Pour cela il y a des techniques de traitement numérique très répandu à cette propriété et beaucoup utiliser, dont nous allons aborder quelques-unes.

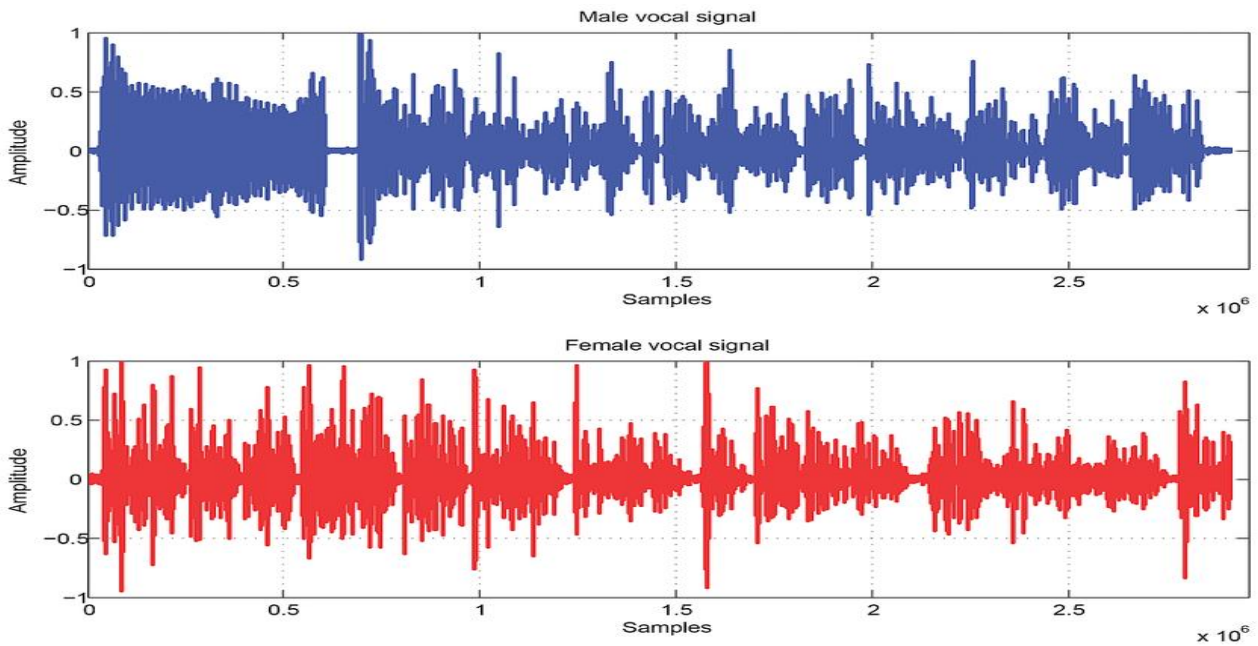


Figure 8: Signal audio [34].

Le signal peut être représenté dans le domaine du temps et même dans un domaine des fréquences, la figure (11) montre une représentation du signal dans le domaine temporel. Cette représentation n'est pas utile pour extraire toutes les informations portées par le signal, pour cela la représentation fréquentiel prend lieu pour aboutir au traitement des signaux numériques.

5 La transformée de Fourier

La transformée de Fourier nous permet de passer depuis le domaine temporel vers un domaine fréquentielle afin de faire une analyse sur les propriétés spectrales. Le signal est considéré formé d'une superposition des sinusoides qui diffère en fréquence et amplitude. [18]

La transformée de Fourier est l'outil qui nous permet de trouver les formant fondamentaux du signal par la transformation directe. Revenir au signal depuis ses formants est possible en passant par la transformation de Fourier inverse.

La Transformée de Fourier Discrète est équivalente à la transformée de Fourier calculée pour des temps discret N séparés par un temps d'échantillonnage T ,

La discrétisation d'une fonction continue $f(t)$ a N échantillons construit à une fonction discrète $f[n]$ telle que : $f[0], f[1], f[2], \dots, f[k], \dots, f[N]$.

La transformée de Fourier d'un signal continue f est définie par la formule suivante:

$$\hat{S}(j\omega) = \int_{-\infty}^{\infty} f(t)e^{-i\omega t} dt \quad 5.1$$

6 L'analyse cepstrale

L'analyse cepstrale est une technique d'analyse acoustique servant à connaître la fréquence fondamentale, et les harmoniques qui construisent le spectre de la voix. L'axe horizontal d'un spectre montre une échelle de fréquences. Par analogie, l'axe horizontal du cepstre est une dimension temporelle appelée "quéfrenc", anagramme de "fréquence", tout comme "cepstre" est une anagramme de "spectre".

Les fréquences uniformément espacées avec des pics de plus petite amplitude dans le cepstre sont appelées des harmoniques, qui sont similaires aux "harmoniques" du spectre.

Le pic cepstral du cepstre reconnu par le pic de haute amplitude, et la période fondamentale du spectre est le pic majoritaire, La période fondamentale est la quéfrenc de l'onde sinusoïdale dominante de l'onde complexe appelée spectre, tout comme la fréquence fondamentale qui est la fréquence de l'onde sinusoïdale dominante du signal vocal. Une forte crête présente dans un signal vocal fortement périodique représente la fréquence fondamentale et les petits pics indiquent les multiples de la fréquence fondamentale. Ces pics de haute amplitude se manifesteront à intervalles réguliers, et dans le cepstre ces intervalles correspondent à la période fondamentale.

L'analyse cepstrale est souvent utilisée pour le suivi de la hauteur des sons et elle reste toujours efficace que pour des voix saines. Les trames de signaux fenêtrée subissent une évolution du cepstre afin de détecter le pic dominant qui théoriquement la période est l'indicateur de la fréquence fondamentale. Cette fréquence fondamentale (tonalité laryngée, hauteur) est présente avec plusieurs harmoniques spectrales qui sont l'effet du passage de la voix à travers des cavités de résonance (telles que les cavités nasales et la gorge).

6.1 Préaccentuation :

Cette étape consiste à filtrer les fréquences élevées. L'émission de la voix depuis un endroit avec pression élevée, à travers une constriction, dans une zone de basse pression, une coupure spectrale se produit pour réduire l'amplitude des hautes fréquences.

Ainsi, la parole enregistrée à l'extérieur de la bouche sera différente de celle enregistrée à l'intérieur de la bouche. Puisqu'il est préférable d'analyser le signal produit par le conduit vocal sans l'influence des rayonnements des lèvres. Nous pouvons donc contrecarrer les effets du rayonnement labial en effectuant une préaccentuation avant l'analyse. Pour un signal vocal $s(n)$, la préaccentuation serait généralement réalisée à l'aide d'un filtre, la relation entrée/sortie du signal est définie par la formule:

$$y(n) = s[n] - \alpha * s[n - 1], \quad 0.9 \leq \alpha \leq 1 \quad 6.1. [32]$$

où $y(n)$ est le signal résultant pour lequel a été appliqué le filtre de préaccentuation, avec α le coefficient de la préaccentuation, et $s[n]$ signal avant filtrage.

La figure suivante représente la réponse des différentes valeurs de Alpha :

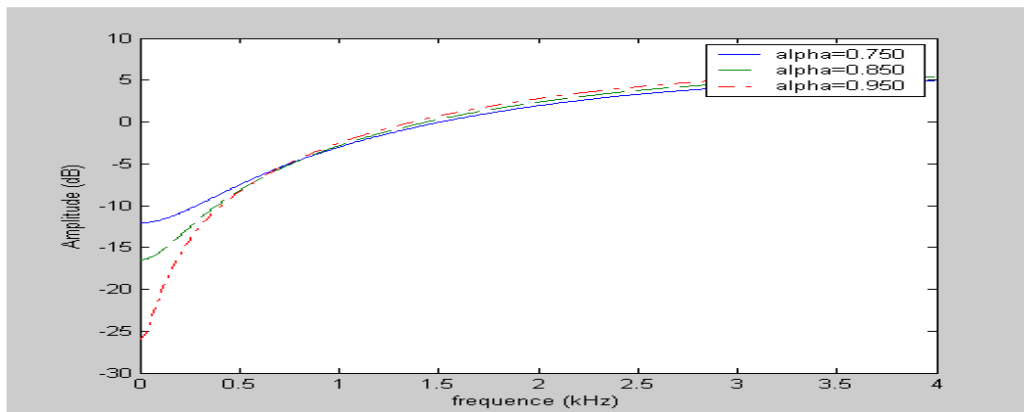


Figure 9: Réponse d'un filtre de préaccentuation pour des différentes valeurs d'alpha [29]

6.2 Segmentation :

Le processus de segmentation des échantillons de parole obtenus par conversion analogique numérique (ADC) en une petite trame dont la longueur est comprise entre 20 et 40 ms.

Le signal vocal est divisé en trames de N échantillons. Les trames adjacentes sont séparées par M ($M < N$).

6.3 Chevauchement :

Le chevauchement signifie qu'au lieu de la segmentation directe du vecteur audio en morceaux séquentiels, chaque nouvelle trame contient une partie de la trame précédente et une partie de la trame suivante. Le chevauchement permet d'éviter la perte de caractéristiques aux limites des trames.

La sélection de la longueur de la fenêtre est un facteur primordial pour effectuer une véritable analyse spectrale, en raison de la compensation existante entre les résolutions temporelle et fréquentielle.

La taille de fenêtre doit être suffisamment importante, afin d'offrir une résolution fréquentielle permanente, tout en restant relativement courte pour capturer les propriétés spectrales locales. Le plus souvent, nous employons un intervalle de 10 à 30 millisecondes [20].

La longueur des fenêtres se diffère entre le sujet masculin et féminin ou le pitch est souvent plus aigu.

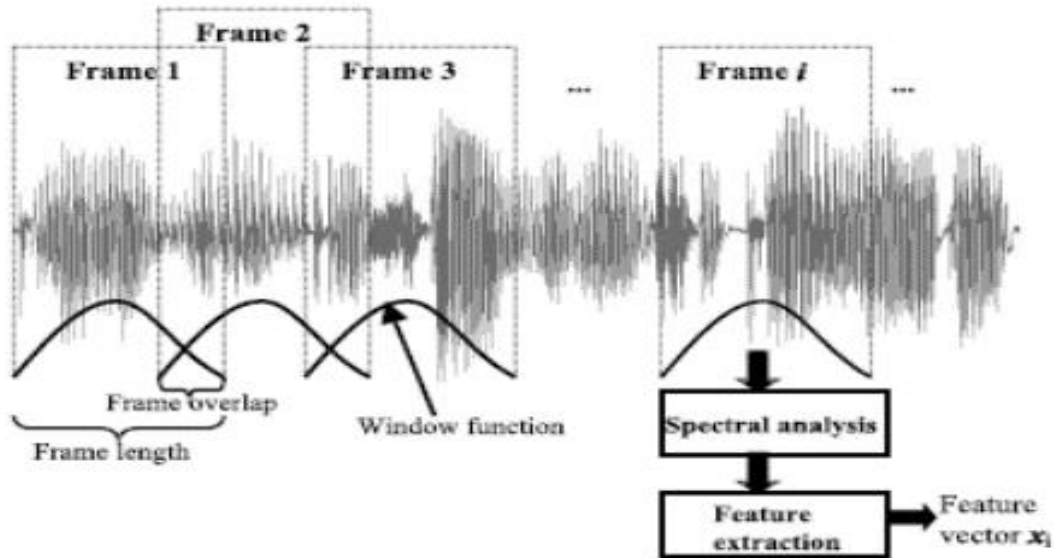


Figure 10: Analyse spectrale à court terme [29]

6.4 Fenêtrage :

Le fenêtrage réduit l'amplitude des discontinuités aux limites de chaque séquence finie acquise par le numériseur, Le fenêtrage consiste à multiplier l'enregistrement temporel par une fenêtre de longueur finie dont l'amplitude varie doucement et progressivement vers zéro sur les bords [20][21]. Cela permet de faire rencontrer les extrémités de la forme d'onde et, par conséquent, d'obtenir une forme d'onde continue sans transitions nettes. Cette technique est également appelée "application d'une fenêtre".

Il existe de nombreux types de fenêtres prédéfinis fréquemment utilisés, chacun d'entre eux ayant des caractéristiques différentes dans les domaines fréquentiel et temporel.

Pour convertir chaque trame de N échantillons du domaine temporel au domaine fréquentiel. La transformée de Fourier permet de convertir la convolution de l'impulsion glottale et de la réponse impulsionnelle du conduit vocal dans le domaine temporel.

On pose $w(n)$ fenêtre avec $0 \leq n \leq N - 1$ ou N est le nombre des échantillons

Les fenêtres les plus utilisées sont :

Fenêtre de Hamming :

Chapitre 2

$$\omega(n) = \begin{cases} 0.54 - 0.46\cos\left(\frac{2\pi n}{N-1}\right), & 0 \leq n \leq N-1 \\ 0, & \text{sinon} \end{cases} \quad 6.1$$

Fenêtre rectangulaire :

$$\omega(n) = \begin{cases} 1, & 0 \leq n \leq N-1 \\ 0, & \text{sinon} \end{cases} \quad 6.2$$

Fenêtre triangulaire :

$$\omega(n) = \begin{cases} 0.5 - 0.5\cos\left(\frac{2\pi n}{N-1}\right), & 0 \leq n \leq N-1 \\ 0, & \text{sinon} \end{cases} \quad 6.3$$

Fenêtre Hann :

$$\omega(n) = \begin{cases} 0.5 - 0.5\cos\left(\frac{2\pi n}{N-1}\right), & 0 \leq n \leq N-1 \\ 0, & \text{sinon} \end{cases} \quad 6.4$$

Fenêtre Blackman :

$$\omega(n) = \begin{cases} 0.54 - 0.46\cos\left(\frac{2\pi n}{N-1}\right) + 0.08\cos\left(\frac{4\pi n}{N-1}\right), & 0 \leq n \leq N-1 \\ 0, & \text{sinon} \end{cases} \quad 6.5$$

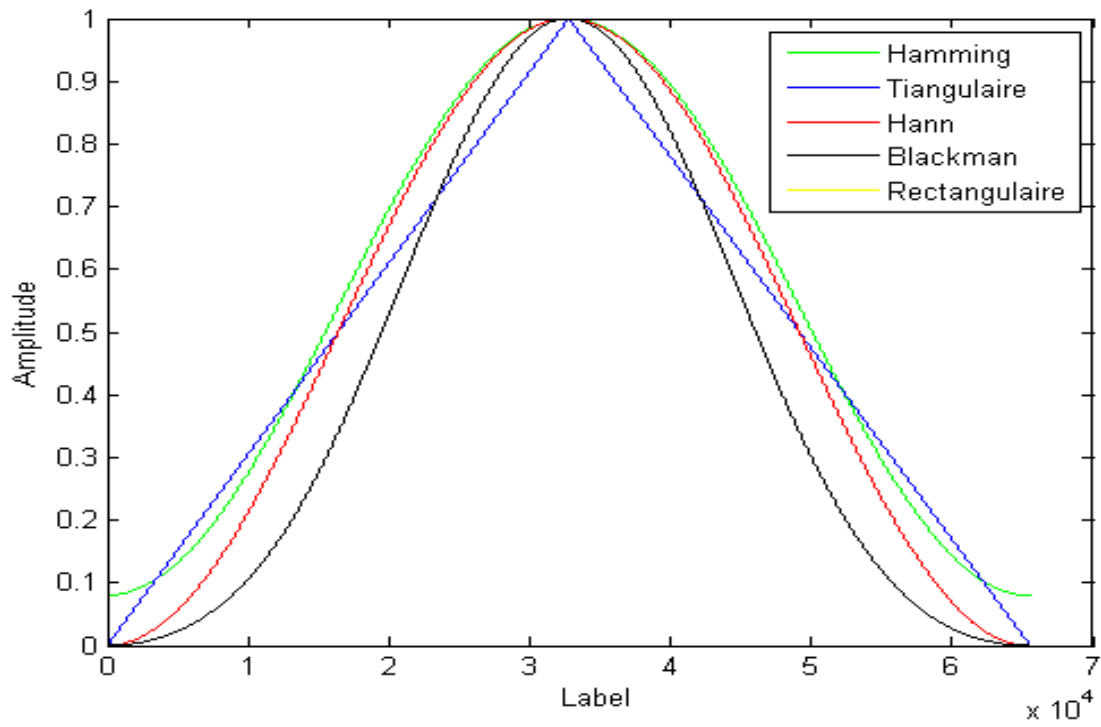


Figure 11: Fenêtres les plus utilisées [35].

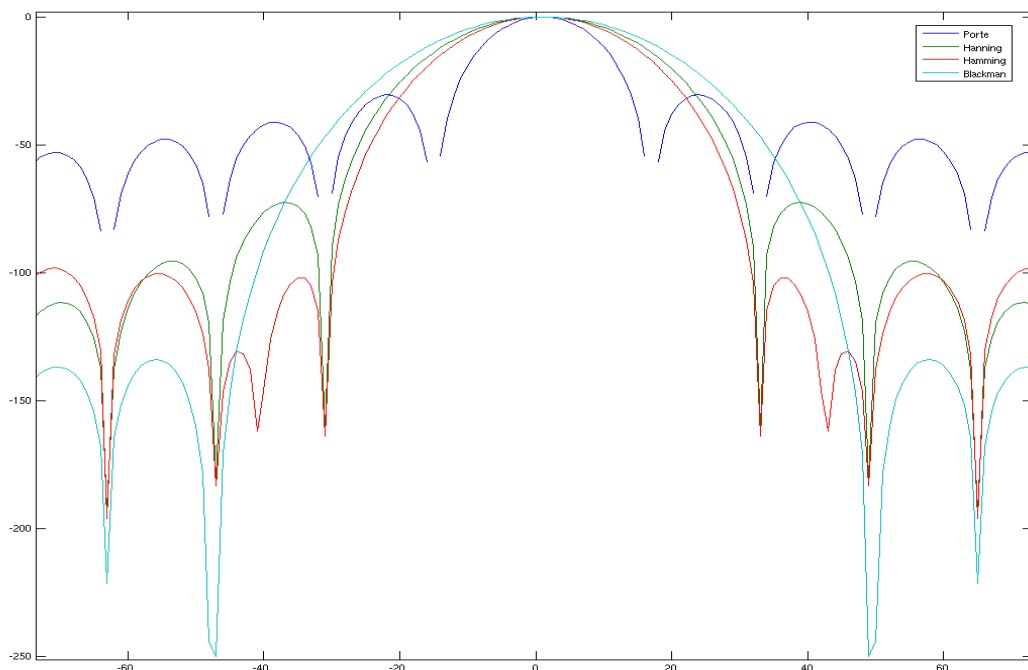


Figure 12: Schéma illustrant allures fréquentielles de quelques fenêtres [35].

6.5 Le Cepstre :

Cette technique est particulièrement efficace pour détecter la périodicité des signaux dans le spectre des fréquences, et de séparer les composantes de ces signaux complexes composés de plusieurs éléments simultanés mais différents combinés ensemble, comme la parole.

Le cepstre est généré comme la transformée de Fourier inverse du logarithme de la transformée de Fourier du signal original, bien qu'en pratique, la deuxième étape est souvent exécutée comme une transformée de Fourier inverse [18][22].

$$C(\tau) = \mathcal{F}^{-1} \{ \log (F_x(f)) \} \tag{6.7}$$

$$\text{où } F_x(f) = \mathcal{F} \{ f_x(t) \} . \tag{6.8}$$

L'analyse cepstrale est une méthode qui permet de détecter la périodicité dans un spectre de fréquence, elle est souvent utilisée pour l'identification des voix pathologiques [23].

6.6 La préminence du pic cepstral (CPP)

La préminence du pic cepstral est la distance mesurée entre le pic du cepstre du signal vocal et la ligne tracée depuis le cepstre (ligne de régression linéaire), cette ligne de régression linéaire indique l'énergie sonore moyenne.

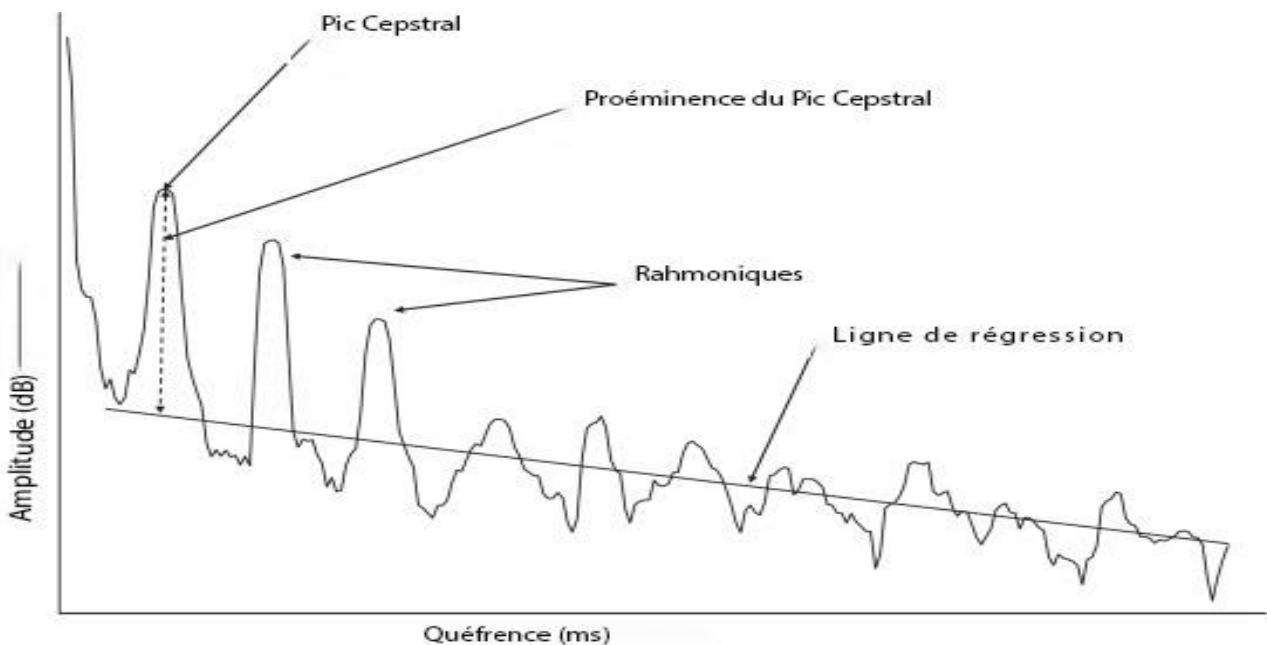


Figure 13: Schéma illustrant les composantes du cepstre et de sa ligne de régression pour déterminer la CPP dans un signal quasi-périodique.[31]

Chapitre 2

Les avantages de l'utilisation du cepstre au lieu de la fonction d'autocorrélation dans certaines circonstances peuvent être interprétés de deux manières différentes. En ce qui concerne les bandes latérales, cela signifie que grâce à la conversion logarithmique, un poids plus important est accordé aux composantes de bas niveau, ce qui est avantageux lorsqu'il s'agit de confirmer l'existence d'une périodicité, où l'amplitude du CPP diffère entre un signal périodique et un autre faiblement périodique, plus l'amplitude du CPP est élevée plus le signal est fortement périodique et vice versa [24].

Le signal vocal peut être quantifié en utilisant cette mesure acoustique (CPP) qui est une mesure fiable équipée d'un algorithme d'analyse vocale robuste.

Les travaux récents en analyse acoustique de la voix apportent de plus en plus la prééminence du pic cepstral (CPP) comme mesure objective de la respiration et de la dysphonie dans son ensemble, En 2018, les recommandations de l'American Speech-Language-Hearing Association (ASHA) ont recommandé la CPP comme outil de "mesure de la dysphonie".

La CPP est donc confirmée comme une mesure objective de la respiration et de la dysphonie dans son ensemble [25].

Beaucoup d'études démontre que la CPP est capable de différencier perceptivement les voix dysphoniques et non dysphoniques à travers les langues, les types de troubles et les tâches de parole. Ces études montrent que les plus faibles valeurs de CPP sont bien liées aux augmentations de la gravité de la dysphonie basée sur des évaluations auditives [25], [26].

Conclusion

Dans ce chapitre nous avons abordé quelques méthodes d'analyse objective du signal, en commençant par l'analyse spectrale qui se base sur la transformée de Fourier, nous avons accentué nos recherches sur l'analyse cepstrale. Nous avons abordé cette transformée, sa particularité et ses caractéristiques distinctives des méthodes temporels. Le chapitre suivant sera consacré à l'application de cette dernière méthode et à la discussion des différents résultats.

Chapitre 3 : Résultats obtenus et discussions

Introduction

Dans ce chapitre, nous allons présenter notre application destinée au traitement numérique du signal vocal pour une analyse objective: l'analyse cepstrale. Nous avons commencé par l'acquisition du signal vocal et appliqué l'analyse cepstrale afin d'extraire le pic prédominant et d'en déduire la fréquence fondamentale f_0 . Nous avons aussi utilisé des signaux enregistrés ainsi que la base de données Physionet [28] afin d'extraire les différentes fréquences fondamentales pour différents sujets. Cette analyse va nous permettre de différencier entre les sujets sains et pathologiques ainsi d'avoir une indication sur le degré de cette pathologie.

1 Présentation du Programme

Le programme a pour but de faire une analyse cepstrale du signal vocal afin de détecter le pic de la proéminence cepstral CPP qui est très important pour notre étude.

Nous avons commencé par l'acquisition des signaux (.wav), puis nous sommes passés à l'étape de prétraitement. La segmentation du signal vocal a été établie en précisant la taille des segments ainsi que le taux de chevauchement, dans notre cas, nous avons utilisé la fenêtre de Hamming.

2 Traitement du signal numérique par exécution de la fréquence fondamentale par la méthode cepstral

Pour faire ce traitement, nous allons commencer par charger le signal vocal de la lettre 'a' pour un sujet sain sur une durée de 3s.

2.1 L'affichage du signal d'entrée :

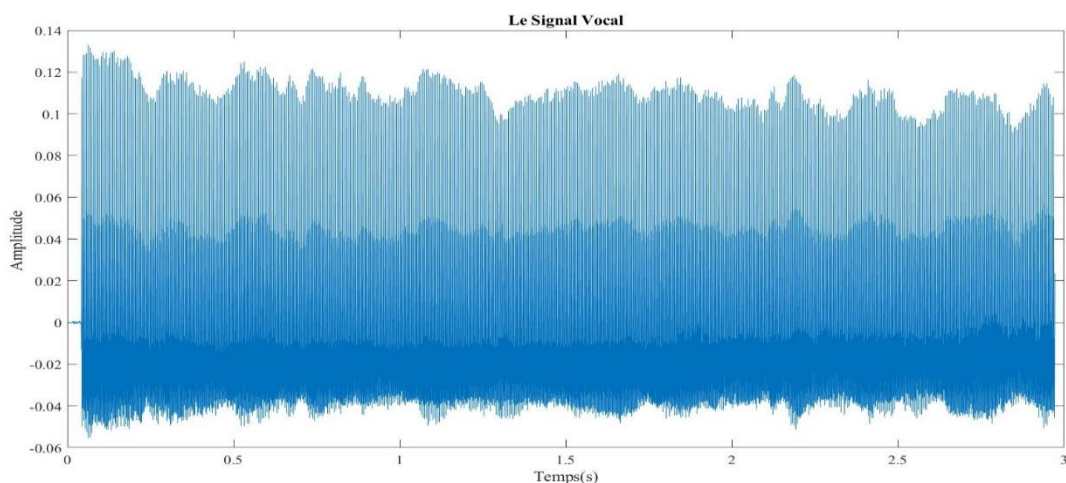


Figure 14: Signale vocal

2.2 Segmentation et filtrage :

L'étape suivante consiste au pré-traitement (segmentation et filtrage), où le filtre utilisé est le filtre de Savitsky Golay d'ordre 9, où le filtre est présenté par la formule suivante :

$$Y_j = \sum_{i=-\frac{m-1}{2}}^{\frac{m-1}{2}} C_i y_{j+i} , \quad \frac{m+1}{2} \leq j \leq n - \frac{m-1}{2} \quad 2.1$$

Les données consistent en un ensemble de points $\{X_j, Y_j\}$, $j = 1, \dots, n$, où Y_j est une variable indépendante et Y_j une valeur observée. Elles sont traitées avec un ensemble de m coefficients de convolution, C_i . [35]

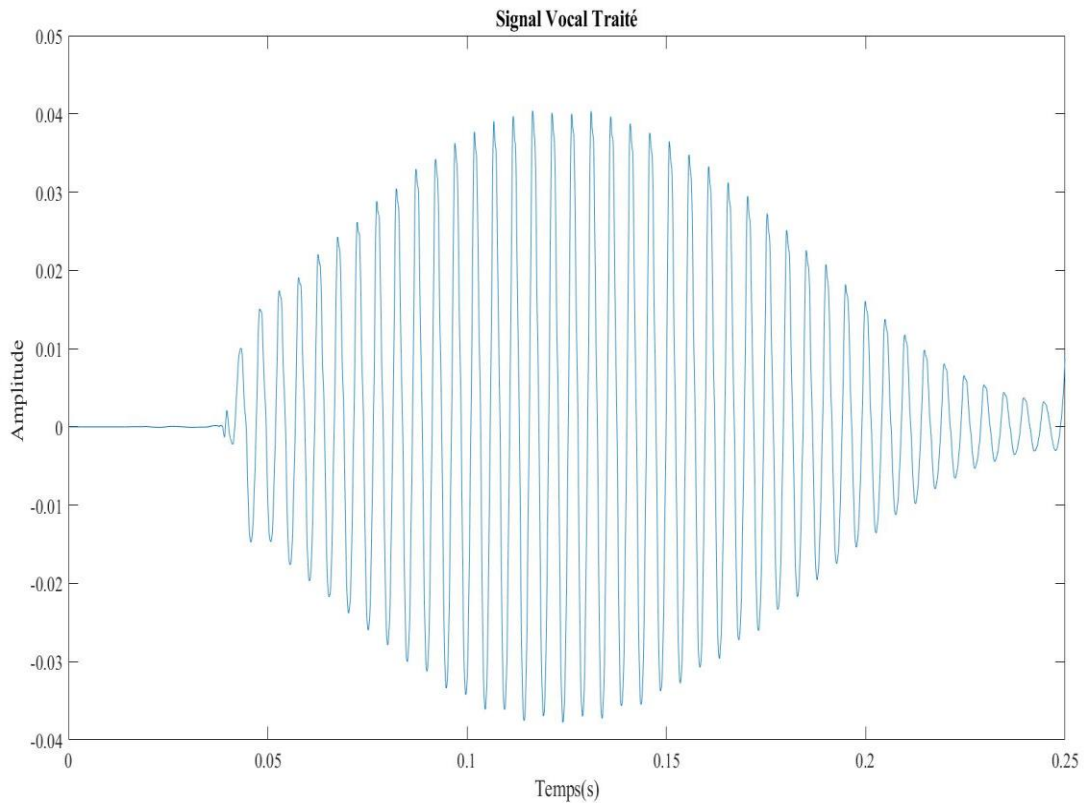


Figure 15:Signal vocal prés traité

2.3 Formation du cepstre :

Chapitre 3

- Le cepstre est calculé par la transformée de Fourier inverse du logarithme décimale de la transformée de Fourier du signal vocal à analyser. La méthode cepstrale devenue efficace pour traiter les signaux vocaux au détriment des méthodes temporelle.
- Le plus grand pic du cepstre est très important comme dans le spectre pour définir la fréquence fondamentale du signal F_0 , ce pic est appelé le pic cepstrale.

Le tracé suivant représente le cepstre du signal vocal de la lettre « a » :

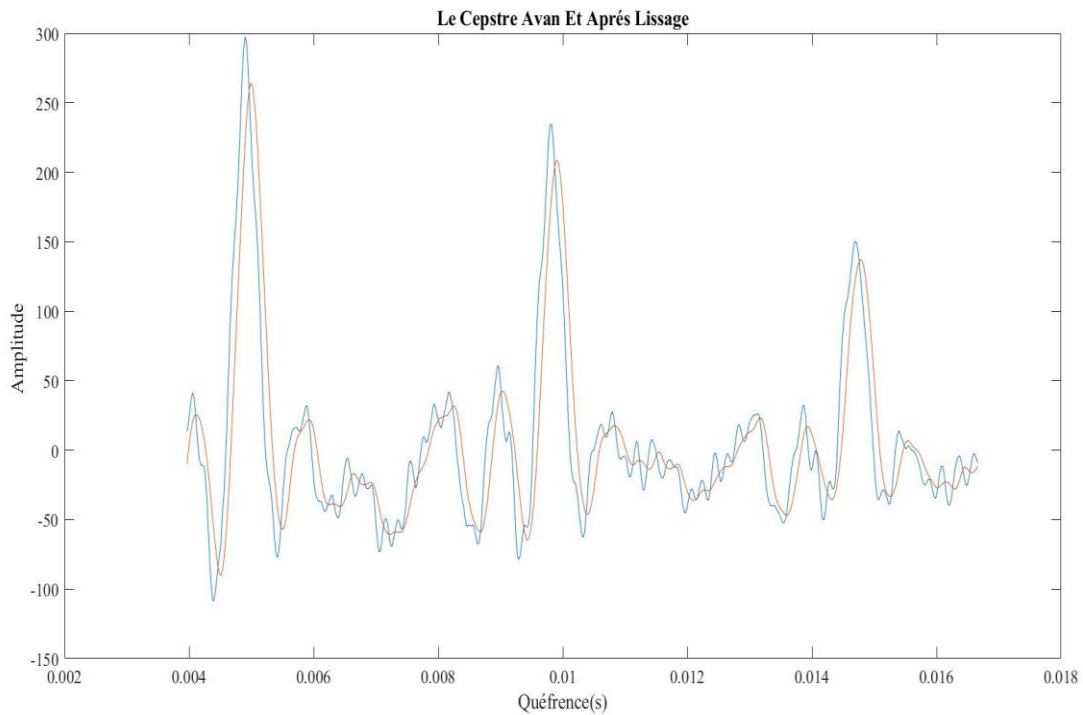


Figure 16: Cepstre du signal vocal 'a'

Une autre donnée indispensable se référant à la périodicité des signaux : la proéminence du pic cepstral. Elle est définie par la différence d'amplitude entre le pic cepstral et la ligne de régression.

2.4 Calcul du cepstre en dB et détection du pic cepstral

Tracé suivant représente le cepstre du signal vocal ‘a’, ainsi que la ligne de régression, avec détection du pic prédominant:

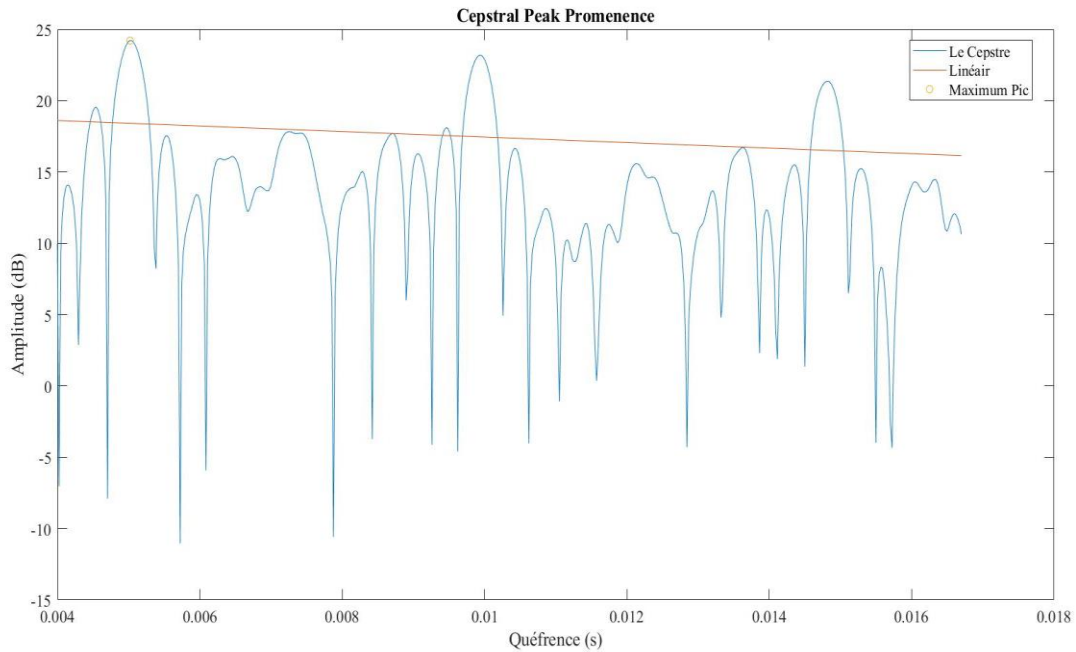


Figure 17:Tracée du cepstre avec la ligne de régression

A partir de ce tracé, nous pouvons clairement détecter le pic cepstral pour en déduire la CPP (Cepstral Peak Promenence). La fréquence fondamentale du signal (F_0) est ainsi calculée,

D'où: $F_0 = 1/Q$,

avec(Q) la quéfrenc du pic cepstrale.

La CPP prend des valeurs minimales quand la mesure est faite pour des signaux de parole continue contrairement à des sons des voyelles soutenue pour quelque seconde, ainsi les valeurs obtenues des voix féminines et des personnes âgées sont faibles.

Donc quatre paramètres influx sur la CPP, La qualité de la voix (saine ou pathologique), le genre, l'âge et le type du son continue ou pas.

Les valeurs du CPP et F_0 sont comprises entre 60 et 300 Hertz, cette mesure contrairement au CPP a des valeurs élevées pour les femmes par rapport aux hommes, où ces valeurs restent constantes pour un même sujet.

Nous allons examiner des signaux vocaux de la voyelle ‘a’ de la base donnée physionet, des enregistrements de trois secondes. Les sujets examinés sont des hommes et des femmes saines et malades de différents âges. Les résultats obtenus sont représentés sur le tableau suivant :

Tableau 1: Résultats obtenus par l'analyse cepstrale

Les Résultats finals		
Les différents sujets	CPP (dB)	F ₀ (CPP(Hz))
Sujet 1 : homme sain 55 ans 'a'	12.16	117.41
Sujet 2 : femme 34, nodule. 'a'	7.89	241.35
Sujet 3 : femme 47 ans, hypokinetic dysphonia	9.62	109.05
Sujet 4 : homme 43 ans, reflux laryngites	11.42	113.45
Sujet 5 : homme sain, 30 ans. Parole continue.	8.31	109.05
Sujet 6 : femme 21 ans, sain. Parole continue.	6.78	213.12

Conclusion :

Dans ce chapitre nous avons pu extraire la fréquence fondamentale à partir de l'analyse cepstrale, où nous avons exploité certains signaux de la base donnée Physionet. Nous estimons que pour pouvoir comparer nos résultats, il est intéressant de faire plusieurs enregistrements afin de constituer une base de donnée pour enfin dire que les résultats obtenus sont satisfaisants, et comparable aux résultats des méthodes existant citées dans la littérature.

Conclusion générale

L'outil de traitement numérique du signal a beaucoup aidé dans le domaine de la laryngologie, en particulier l'aspect diagnostique, le rendant plus précis et objectif.

Ce travail s'inscrit dans le cadre de la manipulation de la voix dans le but d'extraire des données liées à l'état du larynx du locuteur, où nous nous sommes concentrés sur le degré de périodicité du signal et ceci à l'aide du programme Matlab qui est doté d'algorithmes spécifiques pour traiter ce type de signal tel que la transformée de Fourier et la transformée de Fourier inverse, avec lesquels on a réussi à faire une analyse cepstrale.

Au terme de ce travail, nous avons obtenu des résultats, dont le nombre de sujets étudiés restent limité pour pouvoir dire que les résultats obtenus sont satisfaisants, proches de la théorie et en accord avec les travaux déjà établis.

Références

- [1] <http://www.orl-parc.com/Activite-ORL-Medicale-Pathologie-Voix-Cordes-Vocales>
- [2] <http://lapc.free.fr/tutorials/Troubles%20de%20la%20voix.pdf>
- [3] Giovanni A., Yu P., Révis J., Guarella MD., Teston B., Ouaknine M. (2006) « Analyse objective des dysphonies avec l'appareillage EVA. Etat des lieux, Revue Oto-rhinolaryngologie Française, 90, p3 183-192
- [4] GIOVANNI, A., HEIM, C., DEMOLIN, D., and TRIGLIA, J.M. (2000). « Estimate subglottic pressure in normal and dysphonic subjects » Otol. Rhinol. Laryngol, 109, 500-504.
- [5] <http://scgwww.epfl.ch/courses>
- [6] Castellengo(M) ; Ronbeau ; Valette(C) « Study of the acoustical phenomena characteristic of the transition between chest voice and falsetto proceedings of SMCA références », Stockholm,1983, P 113-117.
- [7] J. Cisonni « modélisation et inversion d'un système complexe de production de signaux acoustique application a la voix et aux pathologies » Doctorat de l'institut polytechnique de Grenoble. Soutenu le 24 Février 2009.
- [8] T. Dutoit « Introduction au traitement automatique de la parole », notes de cours Faculté polytechnique de Mons, 2000.
- [9] V.A. Dubsesset «la langue française parlée complétés (LPC) production et perception », Thèse de doctorat de l'institut national polytechniques de Grenoble, soutenu le 25 novembre 2005.
- [10] Lawrence R. Rabiner & Ronald W. Schafer. « Introduction to digital speech processing ». Now Publishers Inc., Hanover, MA, USA, 2007.
- [11] <http://www.orl-parc.com/Activite-ORL-Medicale-Pathologie-Voix-Cordes-Vocales>
- [12] <http://lapc.free.fr/tutorials/Troubles%20de%20la%20voix.pdf>
- [13] http://www.edimark.fr/phototheque/galerie_detail.php?id_galerie=998
- [14] S.D. Courbiere médecin ORL, F. Resuel phoniatre, « la voix : la corde vocale et sa pathologie ». Collège international de médecine et chirurgie de l'hôpital américain de paris, ISBN 2-905238-23-2.

- [15] <http://blogopedie.canalblog.com/archives/2007/04/05/4535001.html>
- [16] I. Atallah « la voix : la corde vocale et sa pathologie. Examen du larynx : de laryngoscopie a la stroboscopie », Université Joseph Fourier Science Technologie Médecine, CHU Grenoble.
- [17] A. Grovanni « la voix : la corde vocale et sa pathologie ». Orientation diagnostique devant : troubles aigus de la parole. Professeur à la faculté de médecine de Marseille, Juin 2005.
- [18] McLoughlin, Ian. "Speech and audio processing : a Matlab-based approach / Ian Vince McLoughlin, University of Kent." New York, NY : Cambridge University Press, 2016.
- [19] Lindasalwa Muda, Mumtaj Begam and I. Elamvazuthi. "Voice Recognition Algorithms using Mel Frequency Cepstral Coefficient (MFCC) and Dynamic Time Warping (DTW) Techniques". JOURNAL OF COMPUTING, VOLUME 2, ISSUE 3, MARCH 2010, ISSN 2151-9617.
- [20] R. Boite, H. Bourlad, T. Dutoit, J. Hancq, and H. Leich, Traitement de la parole, presses polytechniques et universitaires romandes. France, décembre 99.
- [21] The Scientist and Engineer's Guide to Digital Signal Processing.
- [22] R.B. Randall B. Tech., B.A. "CEPSTRUM ANALYSIS AND GEARBOX FAULT DIAGNOSIS". Bruel & kjaer Instruments, Inc. Naerum Danemark.
- [23] Ruben Fraile, Juan Ignacio Godino-Llorente. "Cepstral peak prominence : A comprehensive analysis. " Circuits & Systems Engineering Department, ETSIS Telecomunicacion, Universidad Politecnica de Madrid, Campus Sur, Carretera de Valencia Km.7, 28031 Madrid, Spain.
- [24] Sujitha P S. "Cepstral Analysis of Voice in Young Adults". MSc Speech Language Pathology. India,2020.
- [25] Olivia Murton, Robert Hillman, and Daryush Mehta, "Cepstral Peak Prominence Values for Clinical Voice Evaluation". American Journal of Speech-Language Pathology. Vol. 29. August 2020.
- [26] Tria Khelif, "Réalisation d'une Interface Dédiée au Traitement Numérique du Signal Vocal". Université Abou BakrBelkaïd de Tlemcen. Juillet 2021.
- [27] Neal B. Gallagher, "Savitzky-Golay Smoothing and Differentiation Filter".
- [28] <https://physionet.org/content/voiced/1.0.0/>.

[29] Hadri Cherif, "La recherche des paramètres de la trace acoustique et son application dans la reconnaissance de la parole". Université Badji Mokhtar-Annaba, 2008.

[30] Abdelouahed Sara, "Etude et réalisation d'une plateforme télé médicale dédiée a l'évaluation objective et au suivi des dysphonies chroniques d'origines laryngée par analyse SPECTRO-TEMPORELLE du signal vocal ". Université Abou BakrBelkaïd de Tlemcen.2015.

[31] Calvin P Baker, Johan Sundberg "CPPS and Voice-Source Parameters: Objective Analysis of the Singing Voice".

[32] Djadoudi Mohammed Sobir, Naim Marwa, "Réalisation d'un dispositif d'acquisition, d'analyse et classification de pleurs de bébés". Université Abou BakrBelkaïd de Tlemcen.2021.

[33] Loukhaoukha Khaled, "Quantification vectorielle des paramètre LSF (Line Spectral Frequencies) Application au codeur de parole (CELP)". Université SAAD DAHLAB Blida Institut D'Electronique.

[34] Marco Barbato, Gianni Orlandi, and Massimo Panella, "Real-time Identification and Tracking Using Kinect Multimodal Interaction and Performance Analysis". Département de Information Engineering, Electronics and Telecommunications, Université de Rome "La Sapienza", Rome, Italy.

[35] William H. Press, Saul A. Teukolsky, "Savitzky-Golay Smoothing Filters". American Institute of Physics.

[36] "Transformée de Fourier Rapide", Centre Interuniversitaire de préparation à l'Agrégation de Physique. Montrouge, 2015.