

Tables des matières

Introduction Générale	11
Chapitre I : Généralités sur les sons de la parole	14
1. Introduction :.....	14
2. Production de la parole :	14
2.1. Aperçu anatomique [1] [2]:	14
2.1.1. Partie sous glottique [3]:	15
2.1.2. Partie glottique [3]:	15
2.1.3. La partie supraglottique[3]:.....	16
2.2. Caractéristiques de la production de la parole [2'] :	16
2.3. Mécanisme de la phonation [1]:	17
2.4. Caractéristiques phonétiques [4'] :	20
2.4.1. Définition du phonème [1] [2] [4]:	20
2.4.2. Classification des phonèmes de la langue française [4']:.....	20
2.4.3. Les résonances [7'] [3] [5'] :	22
2.4.3.1. Les résonateurs :	23
2.4.3.2. Organes dont dépend la taille des résonateurs :	24
2.4.3.3. Les muscles du pharynx :.....	24
2.4.3.4. Le larynx :	25
2.4.3.5. Le voile du palais :	25
2.4.3.6. Les lèvres :	26
2.5. Description acoustique des sons de la parole :	27
2.5.1. Introduction :.....	27
2.5.1.1. Les voyelles [3']:	27
A). Définition :	27

B). Le mode d'articulation :	28
C). Lieu d'articulation :	28
2.5.1.2. Les consonnes [3'] :	28
A). Définition :	28
B). Le mode d'articulation :	29
C). Lieu d'articulation :	29
2.5.1.3. Les semi voyelles [3'] :	30
3. Conclusion :	30
Chapitre II : Modélisation du système vocal	32
1. Introduction :	32
2. La Modélisation [5] :	32
2.1. Le modèle physique de la production de la parole [21]. :	33
3. Nonstationarité [5] :	34
4. Représentation numérique du signal vocal [5] :	35
5. Analyse temporelle a court terme [3] [6] :	35
5.1. Energie et puissance à court terme [4] :	35
5.2. Amplitude moyenne à court terme [3] :	35
5.3. Taux de passage par zéro [4] :	36
5.4. Fonction d'auto-corrélation à court terme [4] :	36
6. Analyse spectrale à court terme [3] [6] :	36
6.1. Transformer de Fourier a court terme [3] [11] :	36
6.2. Spectrogramme [4] :	37
7. Conclusion :	37
Chapitre III : Les dispositifs d'acquisition du signal vocal.....	39
1. Introduction :	39
2. Définition de la Carte son [8'] :	39
3. Les connecteurs de la carte son [9'] [[8'] :	39
4. Le format WAVE [9'] [8'] :	42
4.1. Format [9'] [8'] :	42

4.2. Descriptif du son [9'] :	42
4.2.1. Datas [9'] :	44
5. Audacity [9] :	44
5.1. Interface d'AUDACITY [9] :	44
6. La base de donné [11'] :	47
6.1. Définition :	47
6.2. Utilité d'une base de données [10] :	48
7. Conclusion :	48
Chapitre IV : Caractéristiques Spectro-Temporelle du signal vocal multiparamétrique	50
1. Introduction :	50
2. L'analyse acoustique [7] :	50
2.1. La hauteur : son harmonique [8]	50
2.2. L'intensité sonore [8] :	51
2.3. La durée [8] :	52
2.4. Le timbre [8] :	53
2.5. La fréquence fondamentale F0 [7] :	56
2.6. Les Formants [6'] :	56
3. Les mesures de stabilité fréquentielle [7] :	58
3.1. La perturbation à court terme : le jitter	58
3.2. Les perturbations à moyen terme [7] :	59
3.2.1. La mesure de stabilité de l'intensité : le shimmer[7] :	60
3.2.2. Le rapport signal-bruit [7] :	61
4. L'évaluation de la fonction phonatoire [7] :	61
4.1. Le temps maximum de phonation [7] :	61
4.2. La pression sous-glottique phonatoire [7] :	62
5. Les altérations observées en cas de pathologie cordale bénigne [7] :	62
6. Conclusion :	63
Chapitre V : Protocole expérimentale	65
1. Introduction :	65

2. Conception de la base de données Télémédecine :	65
2.1. Conception de la base de données :	65
3. Sélection des Clés primaires :	66
4. L'implémentation de la base de données avec le SGBD :	66
4.1. Système de Gestion de Bases de Données[13] :	66
4.1.1. Définitions :	66
4.1.2. Fonctions des SGBD :	67
4.2. Qu'est ce que MySQL?	67
4.3. Administration avec l'outil web phpMyAdmin :	68
4.3.1. Création de la base de données :	69
4.3.2. La création des tables :	70
5. La réalisation d'un serveur de Télémédecine [10]:	73
5.1. Le serveur web:	73
5.2. L'implémentation du serveur Télémédecine :	73
5.2.1. Interpréteur PHP :	73
5.2.2. L'outil WAMP :	74
6. Aperçu du site web :	76
A. La page d'accueil :	76
B. La page de la connexion :	77
C. La page d'accueil administrative :	77
D. La page du formulaire du patient :	78
E. La page de visualisation :	79
F. La page coordonné d'un patient :	80
G. La page des administrateurs :	81
H. La page des patient rechercher avec (ID : 1) :	81
I. La page de tous les patients enregistrés dans la base de données :	82
7. Les algorithmes de l'interface :	83
7.1. L'algorithme de la TFD CT:	83
7.1.1. Explication :	84

7.1.2. Relation de définition de TFD CT :.....	85
7.2. l'algorithme de l'interface :.....	86
8. Conclusion :	86
Chapitre VI : Evaluation clinique des résultats	88
1. Introduction :.....	88
2. Présentation de l'interface graphique :.....	88
3. Validation clinique des résultats :.....	91
3.1. sujets sains :	91
3.1.1. le premier sujet (masculin) :.....	91
3.1.2. le deuxième sujet (femme) :	94
3.1.3. Le 3eme sujet (femme) :	96
3.2. Sujets atteint d'une pathologie des cordes vocales :.....	99
3.2.1. le premier sujet (masculin)atteint d'un cancer du larynx :.....	99
3.2.2. le 2eme sujet (masculin) atteint d'un polype inflammatoire :.....	101
3.2.3. Le 3eme sujet (masculin) atteint d'un cancer du larynx :.....	104
3.2.3.1. Premier enregistrement :.....	104
3.2.3.2. 2 ^{eme} enregistrement (2 mois après) :.....	106
4. Récapitulation des résultats :	109
5. Interprétation des résultats :.....	109
5.1. Sujets sains :	109
5.2. sujets malades :.....	109
6. Discussion diagnostique :	110
6.1. Les sujets sains :	110
6.2. Les sujets malades :.....	110
7. conclusion :	111
Conclusion générale et perspectives :	112
Bibliographie	114
Neto-graphie	115

Liste des Figures :.....	116
Liste des Tableaux :.....	118
Annexe	119
Résumé	121

Introduction Générale

Le traitement de la parole est aujourd'hui une composante fondamentale des sciences de l'ingénieur. Située au croisement du traitement du signal numérique et du traitement du langage (c'est-à-dire traitement de données symboliques), cette discipline scientifique a connu depuis les années 60 une expansion fulgurante, liée au développement des moyens et des techniques de l'information et de la communication. L'importance particulière du traitement de la parole dans ce cadre plus général s'explique par la position privilégiée de la parole comme vecteur d'information dans notre société humaine.

L'extraordinaire singularité de cette science, qui la différencie fondamentalement des autres composantes du traitement de l'information, tient sans aucun doute au rôle fascinant que joue le cerveau humain à la fois dans la production et dans la compréhension de la parole et à l'étendue des fonctions qu'il met, inconsciemment, en œuvre pour y parvenir de façon pratiquement instantanée.

La parole est en effet produite par le conduit vocal, contrôlé en permanence par le cortex moteur. L'étude des mécanismes de phonation permettra donc de déterminer, dans une certaine mesure, ce qui est parole et ce qui n'en est pas. De même, l'étude des mécanismes d'audition et des propriétés perceptuelles qui s'y rattachent permettra de dire ce qui, dans le signal de parole, est réellement perçu. Mais l'essence même du signal de parole ne peut être cernée de façon réaliste que dans la mesure où l'on imagine, bien au-delà de la simple mise en commun des propriétés de production et de perception de la parole, les propriétés du signal dues à la mise en boucle de ces deux fonctions. Parmi les applications du traitement de la parole nous distinguons :

- 1) l'analyse spatio-spectro-temporelle du signal vocal en vue de la caractérisation objective des dysphonies d'origine laryngée.
- 2) l'estimation quantitative des paramètres caractéristiques du signal vocal au moyen de sa représentation acoustique notamment la fréquence fondamentale des sons voisés et sa dispersion exprimée par l'indice STD, les timbres, les formants, le jitter, le shimmer.

Dans le cadre de ce mémoire nous réalisons une plate forme Télémédicale dédiée à l'acquisition, au traitement, à l'archivage et à la constitution d'une base de donnée interactive pour l'exploration, le traitement, et le suivie des dysphonies d'origine laryngée et plus particulièrement dans le cas des syndromes tumoraux malins.

Pour cela nous développons un algorithme pour qui prend en charge :

- a) la transformation du format WAV en format HEXADECIMAL codé sur 16 bits avec une fréquence d'échantillonnage de 8KHZ.
- b) le calcul de la transformée de fourrier discrète à court terme d'ordre 12.
- c) le calcule de la fréquence fondamentale moyennée sur 6 sélections (moving average).
- d) la mise en œuvre d'une base de données interactive de signaux vocaux acoustiques physiologiques et pathologiques en vue d'une étude clinique épidémiologique et pour une meilleure prise en charge thérapeutique.

Pour ce faire nous organisons notre mémoire en 6 chapitres :

Le premier chapitre intitulé « Généralités sur les sons de la parole » est consacré à l'aspect anatomique de l'appareil vocal et des différents organes responsables de la phonation.

Le deuxième chapitre intitulé « Modélisation du signal vocal » décrit de manière détaillée le modèle physique de la production de la parole et les caractéristiques stochastiques du signal vocal.

Dans le troisième chapitre intitulé « Dispositifs d'acquisition du signal vocal » nous mettons en œuvre la plate forme d'acquisition du signal vocal utilisons la carte son sur le plan hardware et l'environnement Audacity sur le plan software.

L'estimation objective multiparamétrique proprement dite des paramètres acoustiques fait l'objet du quatrième chapitre intitulé « Caractérisation spectro-temporelle multiparamétrique du signal vocal »

Le cinquième chapitre intitulé « protocole expérimental » prend en charge l'implémentation des différents algorithmes que nous avons développés ainsi que du site internet hébergeant la base de donnée interactive dédiée au suivie des dysphonies vocales d'origine laryngées.

Le sixième chapitre est réservé à la présentation et à l'interprétation des résultats obtenus. Nous achevons ce mémoire par une conclusion générale ouvrant les perspectives de sa poursuite.

Chapitre I :

Généralités sur les sons de la parole

Clicours.COM

Chapitre I : Généralités sur les sons de la parole

1. Introduction :

Dans toute forme de vie, il existe un besoin de communication. La forme la plus évoluée et la plus complexe est la fonction d'expression de la pensée et de la communication entre les humains, mise en œuvre au moyen d'un système de signes vocaux et éventuellement de signes graphiques qui constituent une langue.

Le premier et le plus facile des supports de langages est la parole. La parole est la faculté de communiquer la pensée par un système de sons articulés, elle se distingue des autres sons par des caractéristiques acoustiques ayant leurs origines dans les mécanismes de sa production. Les sons de parole sont produits soit par la vibration des cordes vocales (source de voisement), soit par l'écoulement turbulent de l'air dans le conduit vocal, soit lors du relâchement d'une occlusion de ce conduit (source de bruit).

Dans le processus de communication parlée, pour une longue donnée, les sons permettent de distinguer les différentes unités de signification du langage. L'unité sonore élémentaire permettant la distinction des différents mots est le phonème [1].

Dans ce travail toutes les notations des voyelles et des consonnes respectent le tableau des phonèmes du français de l'association phonétique internationale API [7]. (Voire annexe)

2. Production de la parole :

IL n'existe pas a proprement parler un organe spécifique destinée à la production de la parole. L'homme a cette fin fait un usage particulier de son système respiratoire et d'une partie de son système digestif.

2.1. Aperçu anatomique [1] [2]:

Les sons de la parole sont produits par la modification du courant d'air de l'expiration en provenance des deux poumons (partie sousglottique).ce souffle d'air se transforme, selon la configuration du conduit vocal, lors de son passage dans le larynx (partie glottique), le pharynx et les cavités buccale et nasale (partie supraglottique).

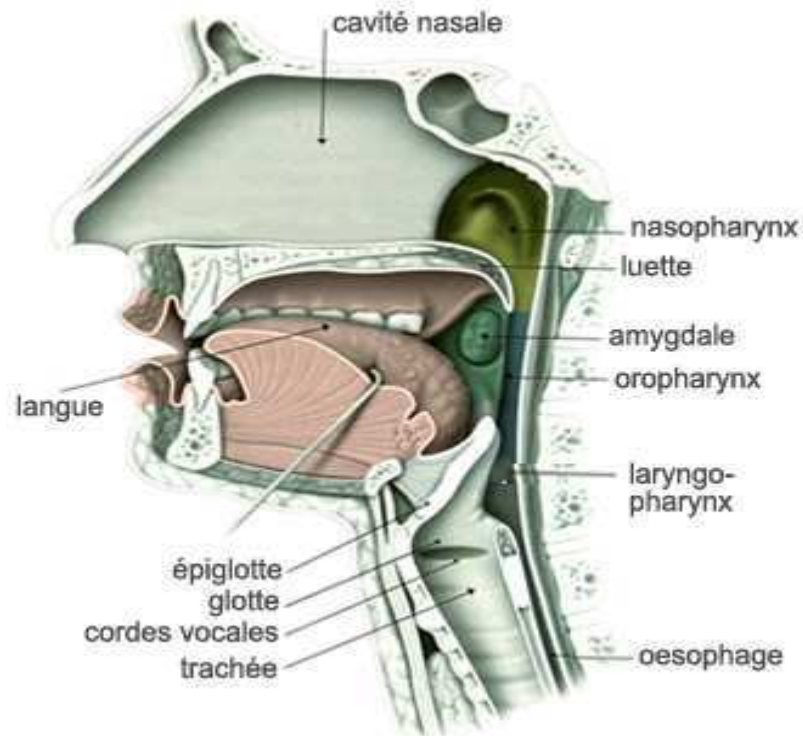


Figure 1 : Schéma de l'appareil phonatoire [3]

2.1.1. Partie sous glottique [3]:

La partie subglottique constituée des poumons et de la trachée qui ont pour rôle principal de fournir l'énergie nécessaire à la phonation en insufflant de l'air vers la partie glottique. Lors de l'expiration, les poumons provoquent une suppression d'air et chassent ainsi l'air préalablement inspiré.

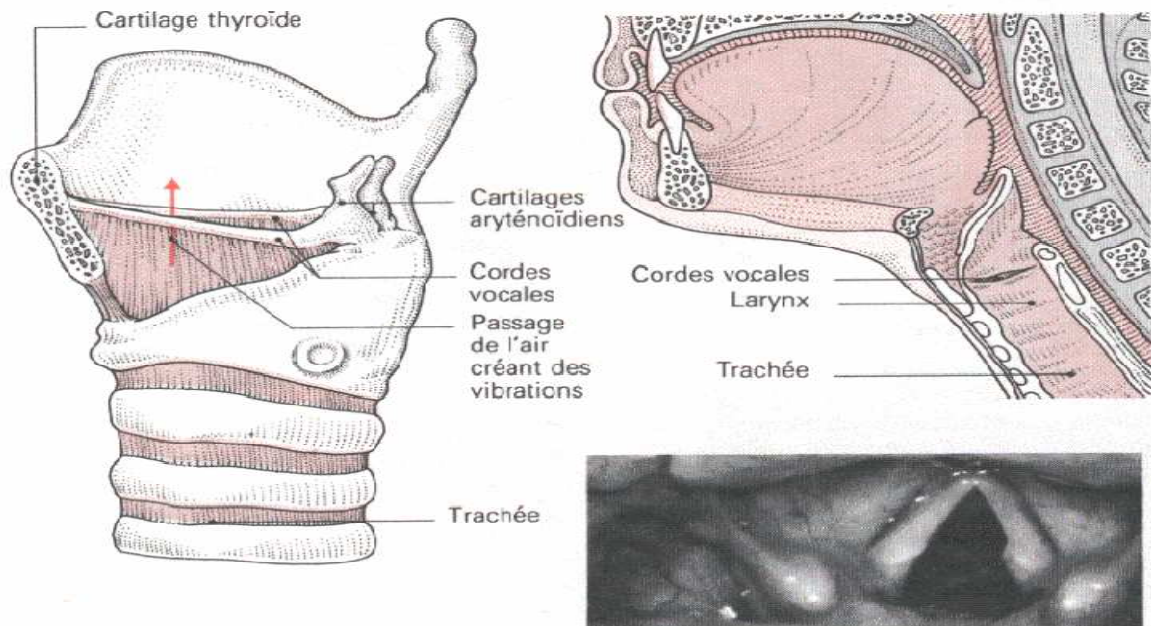
2.1.2. Partie glottique [3]:

Le larynx est l'organe qui joue le rôle le plus important dans tout le mécanisme de la parole. Il est constitué d'un ensemble très complexe de cartilages, de ligaments et de muscles, à la base duquel sont attachées les cordes vocales (Figure 2) qui constituent l'élément fondamentale de la phonation et la jonction entre l'étage supra glottique (bande ventriculaire, épiglotte, base de la langue) et l'étage sous glottique (sous glotte avec cartilage cricoïde, trachée cervical).

CORDES VOCALES

Les cordes vocales sont deux replis situés à la partie supérieure du larynx, tendus entre le cartilage thyroïde en avant et les aryténoïdes en arrière. Elles

sont unies et fixes en avant, mobiles en arrière, formant un V ouvert. Elles se joignent pour vibrer à l'émission des sons.

**Action**

Pour la production de la voix, les cordes vocales sont rapprochées l'une contre l'autre, par des muscles qui agissent sur les cartilages aryténoïdiens.

Vue des cordes vocales

Cette photographie montre les deux cordes vocales au repos, leur écartement dessinant un V.

Figure 2 : Cordes vocales [1']

Deux cordes, formées par des replis horizontaux, tendus symétrique de part et d'autre de la ligne médiane du larynx. Elles forment une ouverture en forme de fente (glotte) qui se comporte comme une vanne commandant le débit d'air lors de la phonation ou permettant la respiration.

2.1.3. La partie supraglottique[3]:

La partie supraglottique est constituée par la partie supérieure de larynx (bandes ventriculaires, ventricule, épiglote, base de langue, pharynx et cavité buccal) le tout relié à la cavité nasale par l'abaissement du voile du palais. Les deux paramètres : la forme de la cavité buccale et la position de la langue, sont utilisés par la phonétique articulatoire pour décrire les phonèmes.

2.2. Caractéristiques de la production de la parole [2'] :

Le processus de production de la parole présente certaines caractéristiques :

- Continuité : lorsqu'on écoute parler une personne, on perçoit une suite de mots que l'analyse de signal vocal sépare difficilement. Le même problème de

segmentation se retrouve à l'intérieur du mot, perçu comme une suite de sons élémentaires (les phonèmes).

- Variabilité : a contenu phonétique égal, le signal vocal est très variable, tant pour différents individus que pour un même locuteur, en raison des différences anatomiques.
- Le conduit vocal est un tuyau tridimensionnel qui est excité par une ou deux sources acoustiques. La source laryngienne peut être considérée comme quasi périodique, avec une fréquence pouvant évoluer très rapidement. La source génère du bruit de friction ou d'explosion (glotte, lèvres).
- Encodage : depuis l'idée jusqu'au signal sonore, interviennent plusieurs niveaux successifs de traitement : sémantique (concept), syntaxique (structure du langage), lexical (mots), morphologique, phonétique (phonèmes et leurs interactions).

Le larynx est un lieu important pour les mécanismes phonatoires. La langue joue un rôle dans la phonation, car sa mobilité lui permet d'agir avec précision et rapidité sur la taille du conduit vocal.

Les lèvres sont situées à l'extrémité du conduit vocal et c'est leur écartement (la variation de cet écartement) qui est important du point de vue acoustique.

2.3. Mécanisme de la phonation [1]:

En simplifiant, on peut dire que la parole est le résultat de l'excitation des cavités supra glottiques par une ou deux sources acoustique. La première, essentielle, génère une onde quasi périodique, c'est la source du larynx. L'autre peut s'ajouter ou se substituer à la première : il s'agit, cette fois, de bruit de friction ou d'explosion qui peuvent naître à l'intérieur du conduit vocal (de la glotte aux lèvres), s'il y a un rétrécissement nettement marqué ou si une occlusion se relâche brusquement. Ces sources vont exciter le conduit vocal, dont la forme dépend de l'articulation (figure 3).

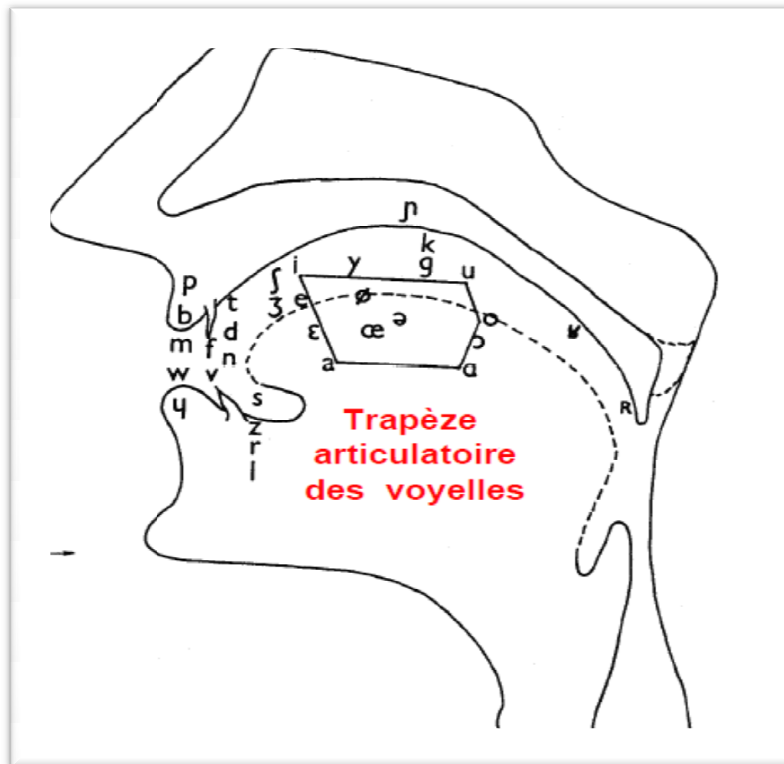


Figure 3 : la phonation des phonèmes français

Donc la structure du signal de la parole est complexe tantôt périodique tantôt aléatoire, ceci nous ramène à répartir les sons de la parole en deux catégories : les sons voisés et les sons non voisés.

- Les sons voisés : ils résultent d'une vibration périodique des cordes vocales (excitation voisée) et d'une configuration quasi-stable du conduit vocale, ce sont des signaux quasi-périodiques, exemple : les voyelles (a, i,...) et certaines consonnes (v, b,...) (figure 4)

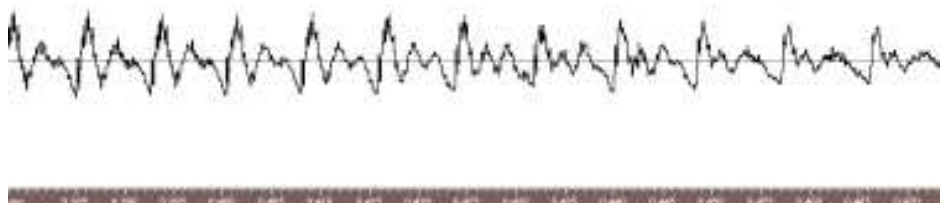


Figure 4 : exemple de sons voisé

- Les sons non voisés : ces signaux ne présentent pas une structure périodique en raison de la position écartée des cordes qui ne vibrent pas, exemple (f, p,...) (figure 5).



Figure 5 : Exemple de sons non voisé

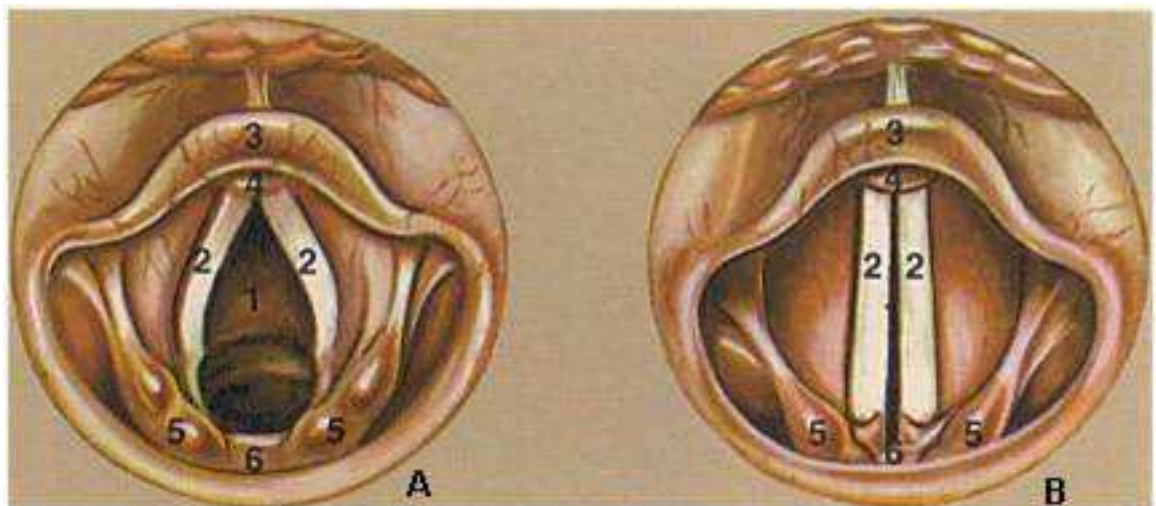


Figure 6 : A. Glotte pendant la respiration B. Glotte pour la phonation

Glotte 2. Cordes vocales 3. Epiglottite 5. Cartilages

Si on fait une classification plus détaillé on peut aussi distinguer :

- Les sons fricatifs : les sons fricatifs résultent de l'écoulement de l'air dans une constriction (resserrement) étroite située en un point du conduit vocal. En particulier au niveau des lèvres et des dents. les sons fricatifs sont voisés comme (v, z,...) ou non voisés comme (f, s,...).
- Les sons occlusifs : ces sons sont produits par une occlusion (fermeture) momentanée du conduit vocal en un point donné, suivie par une ouverture brusque : ils peuvent être voisés (b, d,...) ou non voisés (p, t,...).
- Les sons liquides : l'obstruction n'existe que sur une petite largeur au centre du conduit vocal, laissant un écoulement libre sur les cotés. Exemple (l, r,...).
- Les sons nasals : produit avec le voile du palais abaissé, ils sont souvent décrits comme des occlusives dans la mesure où l'on tient compte que de la partie buccale du conduit vocal, ou elles représentent que une occlusion. Exemple (m, n,...).

2.4. Caractéristiques phonétiques [4'] :

La principale fonction des sons dans une langue est d'établir des distinctions entre des unités de signification, ce sont des phonèmes qui permettent cette distinction.

2.4.1. Définition du phonème [1] [2] [4]:

C'est la plus petite unité présente dans la parole qui est susceptible par sa présence de changer la signification d'un mot par exemple pari/mari, pas/bas, et mie/mes. Cette définition ne tient compte que des caractéristiques qui sont pertinentes pour les distinctions de signification.

Les réalisations physiques d'un phonème peuvent donc varier considérablement en fonction du contexte du dialecte, du style du locuteur par exemple la langue française comprend 36 phonèmes.

2.4.2. Classification des phonèmes de la langue française [4']:

Les phonèmes d'une langue ne forment pas une liste amorphe, se regroupent en catégories naturelles dont les éléments partagent des « trait distinctifs » la notion de trait exprime une similarité au niveau articulatoire, acoustique ou perceptif. La figure 7 donne une classification des phonèmes du français en traits distinctifs.

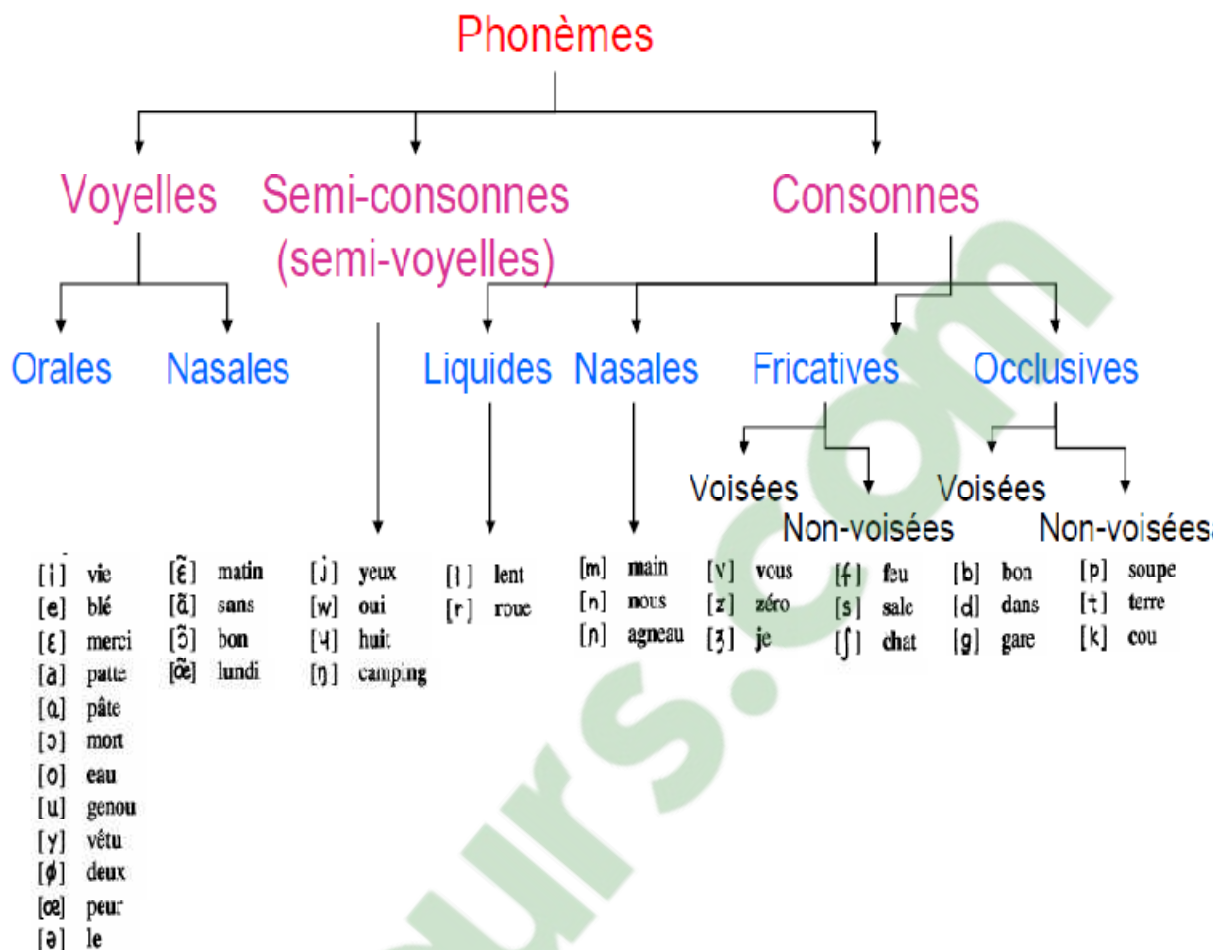


Figure 7 : Les phonèmes de la langue française

Les voyelles sont classées selon quatre traits :

- La nasalité.
- Le degré d'ouverture du conduit vocal.
- La position de la constriction principale du conduit vocal (antérieure/postérieure).
- L'arrondissement des lèvres.

Les consonnes sont classées selon :

- Le voisement.
- Le mode d'articulation (occlusif, nasal, fricatif, glissant ou liquide).
- La position de constriction principale ou « lieu d'articulation » (labial, dentale ou vélo-palatale).

2.4.3. Les résonances [7'] [3] [5'] :

Le conduit vocal constitue une cavité de résonance, les fréquences de résonance dépendent du volume d'air et de la forme du résonateur. Elles peuvent être modifiées par le jeu complexe des organes articulatoires.

Le conduit vocal modifie le son glottique et lui confère ses caractéristiques définitives : on dit qu'il possède une fonction de transfert. Le spectre de la parole émise à la sortie des lèvres est le produit du spectre de la source et de la fonction de transfert. La fonction de transfert du conduit vocal détermine les fréquences qui seront renforcées par le phonème de résonance : on leur donne le nom de fréquences de formants.

Les fréquences de formant dépendent de la forme du conduit vocal. Les mouvements des organes articulatoires modifient la forme des résonateurs. Pour chaque position de la langue et pour chaque nouvelle configuration du conduit vocal, la cavité bucco-pharyngale va résonner à des fréquences spécifiques. Les voyelles sont ainsi caractérisées par des fréquences de formants.

Les formants sont numérotés dans leurs ordres d'apparitions depuis les basses fréquences : formant 1, formant 2..., formant n. On utilise généralement la notation F1, F2,..., Fn.

Les voyelles sont souvent représentées, positionnées sur un plan, dont les axes sont les formants F1 et F2. Elles tracent alors un triangle dont les extrémités sont occupées par les voyelles «extrêmes», c'est-à-dire [a] [u] [i]. Ce triangle représente également les positions moyennes de la langue dans la bouche selon deux axes.

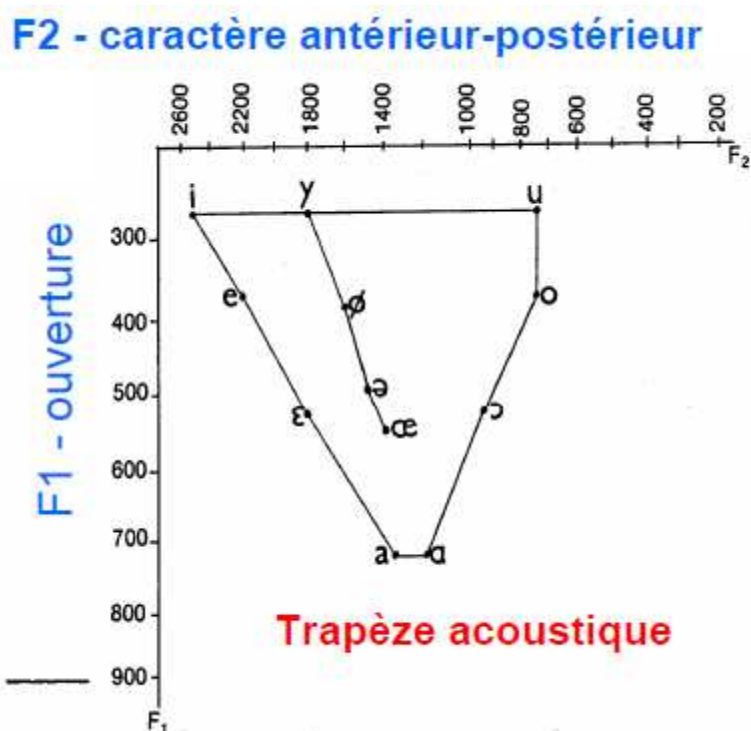


Figure 8 : Trapèze acoustique des voyelles orales du français

2.4.3.1. Les résonateurs :

On donne le nom de résonateurs aux cavités que le son laryngé traverse avant d'arriver à l'air libre, à savoir l'hypopharynx et l'oropharynx, la cavité buccale, et pour certains sons, le nasopharynx et les fosses nasales.

Les résonateurs, situés entre les cordes vocales et l'ouverture de bucco-nasale, correspondent au système de production de la parole.

Ils comprennent un ensemble d'organes, comme le larynx, ne sont pas seulement dévolus à la fonction de communication.

Ces organes constituent le tractus ou conduit vocal. Ce tractus est responsable des modifications de l'énergie sonore produite par le larynx pour en faire les sons du langage, par un ensemble d'actions telles que l'articulation (qui correspond à des mouvements relatifs de ses organes entre eux), la production de bruit de turbulences (tels que des claquements, des sifflements...) et les résonances (qui sont des modes de vibrations particulières à certaines fréquences préférentielles). La particularité du conduit vocal est de pouvoir faire varier sa forme, sa longueur, ses volumes, et donc ses caractéristiques acoustiques.

Clicours.COM

2.4.3.2. Organes dont dépend la taille des résonateurs :

La taille et la forme des résonateurs varient dans des proportions considérables, car leurs parois sont sous la dépendance d'organes mobiles : le maxillaire inférieur, la langue, les muscles du pharynx, le voile du palais, les lèvres.

- Le maxillaire inférieur ou mandibule :

Le maxillaire inférieur est relié au crâne par l'articulation temporo-mandibulaire. Ses mouvements sont sous la dépendance des muscles masticateurs et en particulier du muscle masséter (élévation de la mandibule et fermeture buccale) et des muscles sus-hyoïdiens (abaissement de la mandibule et ouverture buccale).

Sur le plan phonatoire, l'ouverture de la mandibule entraîne un agrandissement de la cavité buccale par abaissement du plancher buccal, ainsi que, très souvent, un abaissement laryngé (bâillement), d'où un agrandissement de la cavité pharyngée.

- La langue :

La langue présente une structure extrêmement complexe. Ses muscles sont au nombre de 17 et ses mouvements sont nombreux et variés. Sa base, postérieure, est relativement peu mobile, mais sa position plus ou moins avancée dans la cavité buccale influence la taille du résonateur pharyngé.

La pointe ou apex, antérieure, est très souple et peut prendre les formes les plus diverses. Son rôle est capital dans l'articulation des sons et des bruits de la parole.

2.4.3.3. Les muscles du pharynx :

Le pharynx est un conduit musculo-membraneux qui s'étend verticalement depuis le larynx en bas jusqu'à la cavité buccale et aux fosses nasales en haut.

Ce conduit à la forme d'un entonnoir irrégulier dont la longueur et le diamètre peuvent varier en fonction de l'activité des muscles qui le constituent. Ces derniers se divisent en :

Muscles élévateurs (muscles du voile du palais, muscles élévateurs du larynx)

Muscles constricteurs : en se contractant, ses muscles rétrécissent les diamètres antéropostérieur et transversal du pharynx.

2.4.3.4. *Le larynx :*

Le larynx constitue la limite inférieure de l'entonnoir pharyngé. En fonction de sa position dans le cou, il va donc modifier la longueur du résonateur pharyngé.

Toute élévation du larynx raccourcit la longueur du pharynx, alors que tout abaissement laryngé allonge la cavité pharyngée.

Par ailleurs, le ventricule de Morgani, cavité située entre les bandes ventriculaires et les cordes vocales, est un résonateur dont la taille varie elle aussi, en fonction de la position du larynx dans le cou. L'abaissement du larynx étire le larynx et agrandit le ventricule de Morgani. L'élévation du larynx réduit la taille du ventricule de Morgani.

L'augmentation de la taille du ventricule de Morgani renforce son rôle de résonateur, lui permettant notamment le renforcement des fréquences aiguës.

2.4.3.5. *Le voile du palais :*

Le voile du palais est une cloison musculo-membraneuse mobile qui prolonge en arrière et en bas la voûte palatine.

Sa face antero-inférieure est concave (buccale), sa face postéro-supérieure est en continuité avec le plancher des fosses nasales.

Son bord postérieur présente un prolongement conique médian : la luette. De chaque côté de la luette, le bord postérieur présente 2 replis curvilignes : les piliers antérieur et postérieur du voile du palais. Entre les piliers se loge l'amygdale palatine.

Le voile du palais est constitué par une aponévrose (lame fibreuse) sur laquelle s'insèrent des muscles (5 paires), le tout recouvert d'une muqueuse.

Le voile du palais est une «porte» qui autorise ou pas le passage de l'air vers les cavités nasales.

Elevé et plaqué contre la paroi pharyngée postérieure, le voile oblige le son laryngé à transiter par la cavité buccale. En fonction de sa position plus ou moins élevée, le voile influence la taille du résonateur pharyngé et du résonateur buccal abaissé, le voile autorise le son laryngé à pénétrer dans les fosses nasales, permettant ainsi la nasalisation des sons.

2.4.3.6. *Les lèvres :*

A l'état de repos, les lèvres se situent au contact l'une de l'autre. Dans la parole, elles produisent des mouvements complexes modifiant la longueur du résonateur buccal et son degré d'aperture. Pour certains, la cavité labiodentale serait responsable de la formation du troisième formant F3.

La source glottique, ou fourniture laryngée, est produite par la vibration de chaque corde vocale. Elle est composée d'une multitude de sons purs regroupés en un son complexe. Le son pur dont la fréquence est la plus grave correspond au son produit par la vibration de la corde vocale dans toute sa longueur et prend le nom de fondamental laryngé ou F0.

Les autres sons composants la fourniture laryngée sont produits par la vibration de chaque portion de corde vocale.

Cette fourniture laryngée va être modifiée par les ajustements articulatoires supra-laryngés qui définissent le mode, la zone, l'aperture et la force articulatoire des sons produits. Ces ajustements sont réalisés principalement par la langue, la mâchoire, les lèvres et le voile du palais.

La phonétique articulatoire nous enseigne qu'il existe des cibles articulatoires pour les différents segments phonétiques d'une langue. Elle nous enseigne aussi que, en raison des différences morphologiques entre individus, les mêmes gestes articulatoires ne produisent pas rigoureusement les mêmes sons chez 2 individus : c'est une des sources de la variabilité inter individuelle. Elle concerne aussi bien la longueur du tractus vocal, l'axe vertical, le diamètre du tractus, l'axe transversal, ou encore la tension générale des organes articulatoires et la nasalisation.

2.5. Description acoustique des sons de la parole :

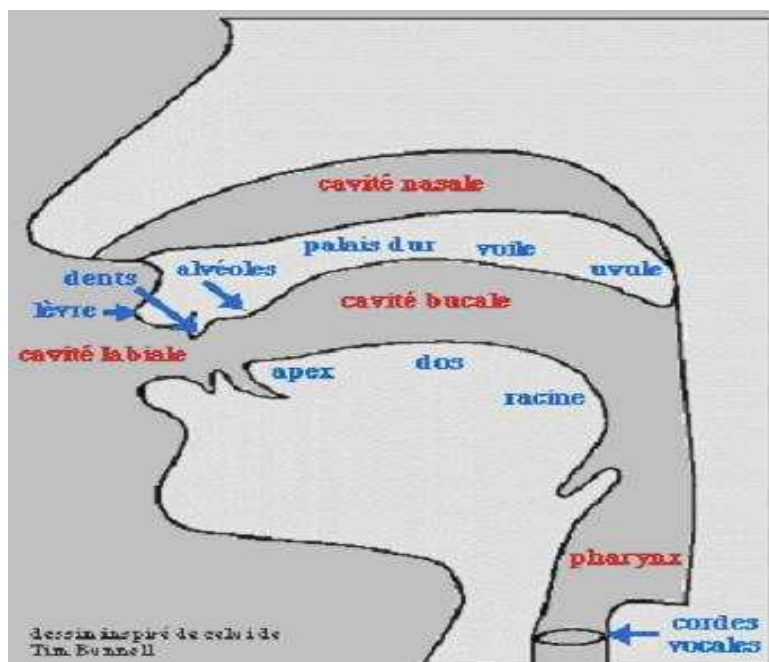


Figure 9 : Lieu d'articulation des phonèmes [4']

2.5.1. Introduction :

La parole se distingue des autres sons par des caractéristiques, qui ont leurs origines dans les mécanismes de production. Les sons de la parole sont produits soit par des vibrations des cordes vocales (sources de voisement), soit par une turbulence créée par l'air s'écoulant rapidement dans une constriction du conduit vocal ou lors de relâchement d'une occlusion de ce conduit (sources de bruit). Donc il est intéressant de grouper les sons de parole en classes phonétiques, en fonction de leur mode articulaire. On distingue généralement trois classes principales : les voyelles, les semi voyelles et les consonnes.

2.5.1.1. Les voyelles [3']:

A). Définition :

Les voyelles diffèrent de tous les autres sons par le degré d'ouverture du conduit vocal et par le degré d'activité des cordes vocales (excitation voisée), qui donnent au conduit vocal une configuration quasi-stable.

Dans la description des voyelles de la langue française, on peut distinguer deux dimensions.

D'un côté, le mode d'articulation qui décrit la configuration générale des organes articulaires dans la production d'une voyelle donnée. D'un autre côté, le lieu

d'articulation : est le lieu où se produit une modification du diamètre du conduit vocal pendant la production du son.

B). Le mode d'articulation :

En français, le mode d'articulation permet de distinguer quatre grandes classes de voyelles :

- Les voyelles orales : elles se prononcent avec le voile du palais, ce qui ferme le passage nasal, donc il y a absence de nasalité (exemple : i (lit), e(les)).
- Les voyelles nasales : Elles se prononcent avec le voile du palais abaissé, ce qui laisse passer de l'air par la bouche, et par le nez. En français on distingue quatre voyelles nasales (exemple (matin, **plante**, long, **lundi**)).
- Les voyelles arrondies : On dit que la voyelle est arrondie si les lèvres sont arrondies exemple (**lu**, **peu**, **loup**).
- Les voyelles non-arrondies :

Les voyelles non arrondies si les lèvres sont ou bien écartées ou bien dans une position neutre (exemple : **papa**).

C). Lieu d'articulation :

Pour bien comprendre le lieu d'articulation, il faut imaginer la bouche comme un espace à deux dimensions, allant du haut vers le bas et de l'avant vers l'arrière de la bouche.

C'est dans l'espace ainsi défini que se situe le point de rétrécissement maximal qui détermine le lieu d'articulation.

- Voyelles antérieures : dans ce cas le bout de la langue se déplace vers l'avant de la bouche (**lait**, **lu**).
- Voyelles postérieures : Dans ce cas le dos de la langue descend dans la bouche et se trouve près du palais (exemple : **leur**, **peu**).
- Voyelle ouverte : lors de la prononciation d'une voyelle ouverte la langue se trouve au fond de la bouche et le passage de l'air est ouvert (exemple : **bâton**).

2.5.1.2. Les consonnes [3']:

A). Définition :

On a une consonne lorsque le passage de l'air venant des poumons est partiellement ou totalement obstrué. Elles peuvent être voisées ou non voisées suivant que les cordes vocales vibrent au passage de l'air ou non.

Comme c'était le cas pour les voyelles, on distingue un mode d'articulation et un lieu d'articulation.

B). Le mode d'articulation :

- Le voisement : Pour les consonnes on peut noter les consonnes voisées et les consonnes non voisées. Ce sont les cordes vocales qui font cette distinction, lorsqu'elles vibrent le son est voisé et la consonne correspondante est une consonne voisée ou sonore. Les consonnes sans vibration sont des consonnes non voisées ou sourdes.
- L'oralité et la nasalité : Dans les consonnes nasales, il y a la composante nasale qui s'ajoute à la composante buccale (l'air sort par le nez et la bouche), tandis que pour les consonnes orales ; l'air sort par la bouche seulement
- Les consonnes occlusives et fricatives : On appelle consonnes fricatives les consonnes qui n'ont pas de fermeture totale du passage de l'air. Par contre, les consonnes occlusives ferment totalement le passage de l'air.
- Les consonnes latérales et vibrantes : On dit consonne latérale, lorsque la langue se met contre les dents supérieures et laisse passer de l'air des deux côtés.

Pour les consonnes vibrantes la langue prend la même position précédente, mais produit un battement qui laisse passer de l'air [ʒ].

Occlusives	Non-voisées	[p]pais	[t]taie	[k]quai
	Voisées	[b]baie	[d]dais	[g]gai
Nasales		[m]mais	[n]nez	[ŋ]gagner
Fricatives	Non-voisées	[f]fait	[s]sait	[ʃ]chez
	Voisées	[v]vais	[z]zéro	[ʒ]geai
Liquides			[l]lait	[R]raie

Tableau 1 : Les semi voyelle du français.

C). Lieu d'articulation :

- Consonnes labiales : on appelle une consonne labiale lorsque les deux lèvres ferment le passage de l'air [p], [b] et [m].
- Consonnes apico-dentales : Dans le cas où le bout de la langue (apex) s'appuie contre les dents Supérieures [t], [d] et [n].

- Consonnes dorso-vélaire : C'est lorsque le dos de la langue s'appuie contre le voile de palais [k], [g].
- Consonnes labio-dentales : Lorsque les dents supérieures, entrent en contact avec la lèvre inférieure, pour fermer le passage partiellement on dit que c'est une consonne labio-dentale [f] et [v].
- Consonnes pré-dorso-alvéolaires : Lorsque la partie antérieure du dos de la langue s'approche du palais dur on peut dire que la consonne est pré-dorso-alvéolaire [z].

2.5.1.3. Les semi voyelles [ʃ] :

Ce sont des sons qui ont la sonorité des voyelles mais l'absence de syllabicité des consonnes on peut les appeler aussi semi-consonnes. Pour résumer, les semi-voyelles ont le même lieu d'articulation que les voyelles, mais ne donnent pas de syllabe. Dans la langue française il y'a quatre semi-voyelles :

Symbole	Exemple
[j]	Yeux
[w]	Oui
[ɥ]	Huit
[ŋ]	Camping

Tableau 2 : Les semi voyelle du français.

3. Conclusion :

Dans ce chapitre ont été présentées quelques notions fondamentales sur la production des sons du langage et les divisions principales dans les quelles le système de l'Alphabet Phonétique International (A.P.I) classe ceux-ci selon deux types de critères : critère acoustique, et articulaire.

Chapitre II : Modélisation du système vocal

Chapitre II : Modélisation du système vocal

1. Introduction :

La parole est un signal réel, continu, d'énergie finie, non stationnaire. sa structure est complexe et variable dans le temps : tantôt pseudo-periodique pour les sons voisés, tantôt aléatoire pour les sons non voisés.

Le traitement du signal vocal a pour but de fournir une représentation moins redondante de la parole que celle obtenue par l'onde temporelle tout en permettant une extraction précise des paramètres significatifs. les principaux problèmes posés en traitement du signal proviennent de la dualité source /bruit de l'appareille phonatoire, de la grande dynamique et de la variété des voix ainsi que des variations rapides de la parole.

2. La Modélisation [5]:

L'analyse de la parole est une étape indispensable à toute application de synthèse, de codage, ou de reconnaissance. Elle repose en général sur un modèle. Celui-ci possède un ensemble de paramètres numériques, dont les plages de variation définissent l'ensemble des signaux couverts par le modèle.

Pour un signal et un modèle donné, l'analyse consiste en l'estimation des paramètres du modèle dans le but de lui faire correspondre le signal analysé.

Pour ce faire, on met en œuvre un algorithme d'analyse, qui cherche généralement à minimiser la différence, appelée erreur de modélisation, entre le signal original et celui qui serait produit par le modèle s'il était utilisé en tant que synthétiseur.

Il existe de nombreux modèles de parole. On distingue les modèles articulatoires, les modèles de production, et les modèles phénoménologiques :

- **modèles articulatoires :**

Les premiers réalisent une simulation numérique du mécanisme de phonation. Leurs paramètres sont essentiellement de nature articulatoire (position de la langue, ouverture des lèvres,...). La parole est décrite comme le résultat du passage d'un flux d'air à travers un ensemble de tubes de section variable.

L'analyse par modèle articulatoire, également appelée inversion acoustique articulatoire, est un problème complexe. Les algorithmes font intervenir les équations de la mécanique des fluides.

- **modèles de production :**

A la différence des précédents, les modèles de production ne cherchent à reproduire que le schéma de principe du mécanisme phonatoire, par le biais de son équivalent électrique. On y décrit la parole comme le signal produit par un assemblage de générateurs et de filtres numériques. Les paramètres de ces modèles sont ceux des générateurs et filtres qui les constituent.

- **modèles phénoménologiques :**

Enfin, les modèles phénoménologiques cherchent à modéliser le signal de parole sans se soucier de la façon dont il a été produit. Les algorithmes d'analyse qui y sont associés se rapportent par conséquent plus au traitement du signal en général qu'au traitement de la parole. Les modèles basés sur l'analyse de Fourier en est un exemple.

2.1. Le modèle physique de la production de la parole [21]. :

Le processus de production de la parole est un mécanisme très complexe qui repose sur une interaction entre les systèmes neurologique et physiologique. La parole commence par une activité neurologique. Après que soient survenues l'idée et la volonté de parler, le cerveau dirige les opérations relatives à la mise en action des organes phonatoires. Le fonctionnement de ces organes est bien, quant à lui, de nature physiologique.

Une grande quantité d'organes et de muscles entrent en jeu dans la production des sons des langues naturelles. Le fonctionnement de l'appareil phonatoire humain repose sur l'interaction entre trois entités : les poumons, le larynx, et le conduit vocal.

Le larynx est une structure cartilagineuse qui a notamment comme fonction de réguler le débit d'air via le mouvement des cordes vocales. Le conduit vocal s'étend des cordes vocales jusqu'aux lèvres dans sa partie buccale et jusqu'aux narines dans sa partie nasale.

La parole apparaît physiquement comme une variation de la pression de l'air causée et émise par le système articulatoire. L'air des poumons est comprimé par l'action du diaphragme. Cet air sous pression arrive ensuite au niveau des cordes vocales. Si les

cordes sont écartées, l'air passe librement et permet la production de bruit. Si elles sont fermées, la pression peut les mettre en vibration et l'on obtient un son quasi périodique. L'air mis ou non en vibration poursuit son chemin à travers le conduit vocal et se propage ensuite dans l'atmosphère. La forme de ce conduit est déterminée par la position des articulateurs tels que la langue, la mâchoire, les lèvres ou le voile du palais, détermine le timbre des différents sons de la parole. Le conduit vocal est ainsi considéré comme un filtre pour les différentes sources de production de parole telles que les vibrations des cordes vocales ou les turbulences engendrées par le passage de l'air à travers les constriction du conduit vocal. Le son résultant peut être classé comme voisé ou non voisé selon que l'air émis a fait vibrer les cordes vocales ou non.

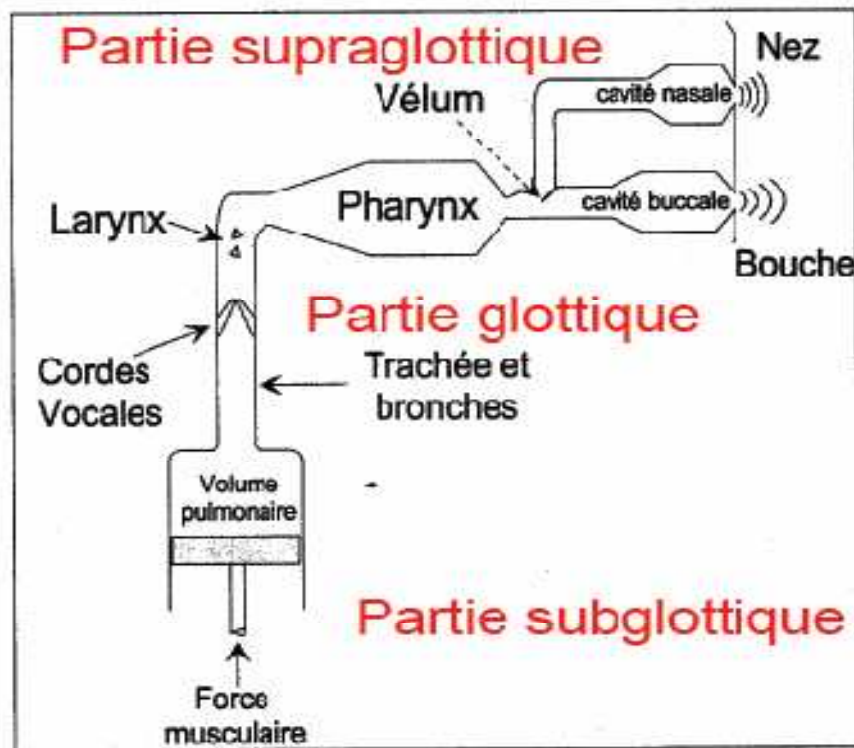


Figure 10 : Le modèle physique de la production de parole

3. Nonstationarité [5]:

Il est essentiel de rappeler que le signal vocal n'est pas un signal stationnaire : le conduit vocal se déforme d'une façon continue et les paramètres du modèle sont donc variables dans le temps.

4. Représentation numérique du signal vocal [5]:

La représentation numérique du signal de la parole implique :

un échantillonnage effectué à une fréquence F_e compatible avec les exigences du théorème de Shannon : la perte d'information entre un le signal discret correspondant est nulle si et seulement si la fréquence d'échantillonnage est supérieure ou égale au double de la fréquence la plus haute contenue dans ce signal. une quantification de chaque échantillon avec un pas de quantification en rapport avec la précision souhaitée.

Un codage approprié selon la nature du problème technique particulier.

5. Analyse temporelle a court terme [3] [6]:

Des mesures simples (énergie, taux de passage par zéro et fonction d'autocorrélation) à court terme sur le signal de parole donnent déjà beaucoup d'information. L'évolution de l'énergie à court terme indique la succession des voyelles (très énergétiques) et des consonnes (d'énergie moindre). le comptage et le tracé d'histogrammes des passages par zéro du signal traduisent, bien que grossièrement, le contenu spectral. L'estimation de période fondamentale peut être réalisée à l'aide de la fonction d'autocorrélation.

5.1. Energie et puissance à court terme [4]:

L'énergie à court terme d'un signal $x(t)$:

$$E_n = \sum_{m=-\infty}^{\infty} x^2(m) \cdot w^2(n-m) \quad (\text{II.1})$$

Avec $w(n)$ la fonction fenêtre (rectangulaire, hamming,...).

La puissance a court terme est définie par :

$$P_n = \frac{1}{N} \sum_{m=-\infty}^{\infty} x^2(m) \cdot w^2(n-m) \quad (\text{II.2})$$

Où N la longueur de la fenêtre.

5.2. Amplitude moyenne à court terme [3]:

La méthode de l'amplitude moyenne à court terme utilise une expression plus simple que la méthode de la puissance à court terme :

$$M_n = \frac{1}{N} \sum_{m=-\infty}^{\infty} |x(m)| \cdot w(n-m) \quad (\text{II.3})$$

5.3. Taux de passage par zéro [4] :

Pour un signal échantillonné, il ya le taux de passage par zéro lorsque deux échantillons successifs sont de signes opposés. Le taux de passage par zero a court terme peut être estimé par la formule:

$$Z_n = \frac{1}{2N} \sum_{m=-\infty}^{\infty} |sgn[x(m)] - sgn[x(m-1)]|.w(n-m) \quad (\text{II.4})$$

5.4. Fonction d'auto-corrélation à court terme [4] :

Elle s'exprime par la relation :

$$R_{xx}(k) = \frac{1}{N} \sum_{m=-\infty}^{\infty} x(m).w(n-m).x(m+k).w(n-k-m) \quad (\text{II.5})$$

$$k=0, 1 \dots K$$

6. Analyse spectrale à court terme [3] [6]:

La méthode spectrale est fondée sur une décomposition fréquentielle du signal sans connaissance a priori de sa structure fine.

Dans certaines mesures, ce choix peut être considéré comme dépendant de la structure de la parole.

Pratiquement, on utilise des versions numériques de ces méthodes, la TFD (Transformée de Fourier Discrète) avec un algorithme de calcul rapide(FFT).en principe, le concept de la TFD ne s'applique qu'à un signal stationnaire à durée limitée par une fenêtre temporelle.

6.1. Transformer de Fourier a court terme [3] [11]:

Les caractéristiques spectrales du signal de la parole a analyser évoluent avec le temps. On cherche une représentation temps fréquence adaptée : c'est pourquoi on introduit la notion de la transformer de Fourier a court terme :

$$x(k, n) = \sum_{m=-\infty}^{\infty} w(n-m).x(m).e^{-j(2\pi/N)mk} \quad (\text{II.6})$$

Il s'agit d'une opération de modulation du signal $x(n)$ par l'exponentielle complexe à la fréquence $(-k/N)$ suivie d'une opération de convolution.la transformées de Fourier

discrète a court terme s'interprète comme une translation du spectre du signal $x(n)$ suivie par un filtrage caractérisé par la réponse en fréquence $H(\omega)$ ($H(\omega)$ est la fonction de transfert d'une fenêtre temporelle).

Lorsqu'on utilise une fenêtre rectangulaire, il s'agit d'un filtre passe-bas. Il en est de même pour tout autre type de fenêtre (Hamming, Kaiser,...).

Ce filtre est d'autant plus sélectif que la fenêtre d'analyse est longue. la résolution fréquentielle est donc directement fonction de la longueur de la fenêtre. Le choix de N est le résultat d'un compromis car on désire généralement avoir aussi des fenêtres courtes pour pouvoir analysé des phonèmes courts.

Si l'on met en parallèle N filtres de même réponse $H(\omega)$ translatés en fréquences, on obtient un banc de filtre d'analyse ou de N est la longueur de la fenêtre.

6.2. Spectrogramme [4]:

Dans la plus part des cas, des analyses successives sur une durée limitée de signal s'obtient en prélevant N échantillons de parole à l'aide d'une fenêtre temporelle glissante.

Les propriétés de la transformée de Fourier a court terme dépendent beaucoup du choix de la fonction fenêtre. La longueur de la fenêtre doit d'une part être suffisante pour assurer une bonne résolution fréquentielle ; d'autre part elle doit être limitée si l'on veut suivre fidèlement l'évolution dans le temps du spectre vocal.

Ces deux exigences sont contradictoires.

Pour suivre mieux les transitoires de la parole, il est nécessaire de prendre des fenêtres temporelles avec recouvrement. Mais il demeure malgré tout un effet de lissage temporel du a la longueur de cette fenêtre ainsi qu'une distorsion du spectre dépendant de fenêtre utilisé.

En générale, on prend $N=256$ ou 512 , et le recouvrement est de moitié soit $N/2=128$ ou 256 .

En pratique, cette méthode est très souvent utilisée pour confectionner un spectrogramme numérique.

7. Conclusion :

Dans ce qui précède nous avons étudié les caractéristiques temporelle et spectrale à court terme de la voix et le modèle physique de la production de la parole, ce contexte est indispensable pour la compréhension de mécanisme de la production de parole.

Chapitre III : Les dispositifs d'acquisition du signal vocal

Chapitre III : Les dispositifs d'acquisition du signal vocal

1. Introduction :

Pour pouvoir manipuler et traiter le signal vocal il faut d'abord penser à acquérir le signal vocal. Dans ce chapitre nous avons parlé de la carte son et le format WAVE du signal vocal et l'acquisition de se dernier qui est faite par des logiciels d'enregistrement, conversion et de l'archivage de celui-ci.

2. Définition de la Carte son [8']:

Une carte audio contient une mémoire et un processeur encasté spéciaux permettant de traiter des fichiers audio et de les écouter par le biais des haut-parleurs de l'ordinateur. Un fichier audio représente l'enregistrement d'un son capturé qui peut être repassé. Les fichiers audio sont généralement compressés afin d'être stockés ou afin de permettre une transmission plus rapide. Le son doit être continu afin que les utilisateurs reçoivent un son en temps réel (effet multimédia), écoutent de la musique ou prennent part à une vidéoconférence ou à une audioconférence. Le format de fichier audio le plus populaire actuellement est le MP3.

3. Les connecteurs de la carte son [9'] [[8']:

Les principaux éléments d'une carte son sont :

- Le processeur spécialisé, appelé **DSP** (*digital signal processor*) chargé de tous les traitements numériques du son (écho, réverbération, vibrato chorus, tremolo, effets 3D, etc.) ;
- Le **convertisseur digital-analogique** appelé **DAC** (digital to analog converter) permettant de convertir les données audio de l'ordinateur en signal analogique vers un système de restitution sonore (enceintes, amplificateur, etc.) ;
- le convertisseur analogique/numérique appelé **ADC**.
- Les connecteurs d'entrées-sorties externes :

- Une ou deux sorties ligne au format jack standard 3.5 mm (notée Line Out ou bien Speaker output ou SPK, signifiant « hauts parleurs » en anglais), habituellement de couleur vert clair.
- Une entrée ligne (Line in).
- Une entrée microphone (notée parfois *Mic*), généralement au format jack 3.5 mm et de couleur rose.
- Une sortie numérique SPDIF (Sony Philips Digital Interface, noté également S/PDIF ou S-PDIF
- Une interface **MIDI**, généralement de couleur or (ocre) permettant de connecter des instruments de musique et pouvant faire office de port de jeu

Les connecteurs d'entrées-sorties internes :

- Connecteur CD-ROM / DVD-ROM, possédant un connecteur noir, permettant de connecter la carte son à la sortie audio analogique du CD-ROM à l'aide d'un câble CD Audio.
- Entrée auxiliaire (AUX-In) possédant un connecteur blanc, permettant de connecter des sources audio internes telles qu'une carte tuner TV.
- Connecteur pour répondeur téléphonique (**TAD**, Telephone Answering Devices) possédant un connecteur vert.

➤ **Enregistrement du son [9'] [8']:**

Un signal sonore est un signal analogique visible sous une forme sinusoïdale. or, en numérique, on utilise des bit (0 ou 1) pour stocker l'information. Lorsque l'on enregistre du son, il faut trouver le moyen de le convertir d'une source analogique en numérique. Voici à quoi ressemble un signal analogique :

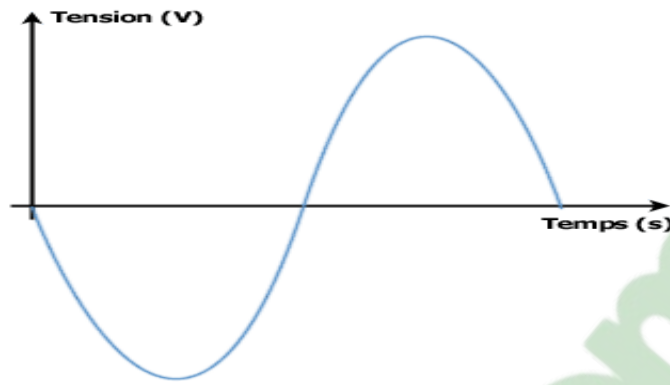


Figure 11 : Représentation temporelle d'un signal analogique

- Pour enregistrer du son il faut d'abord savoir combien de fois par seconde on va "récupérer une valeur sur la courbe". En effet, il faut bien se limiter en un nombre de points fini pour pouvoir acquérir numériquement un signal analogique. On parle de fréquence d'échantillonnage. Plus la fréquence d'échantillonnage est élevée, mieux ça sera au niveau qualité. Le signal devient donc découpé en périodes de temps. On obtient un signal échantillonné :

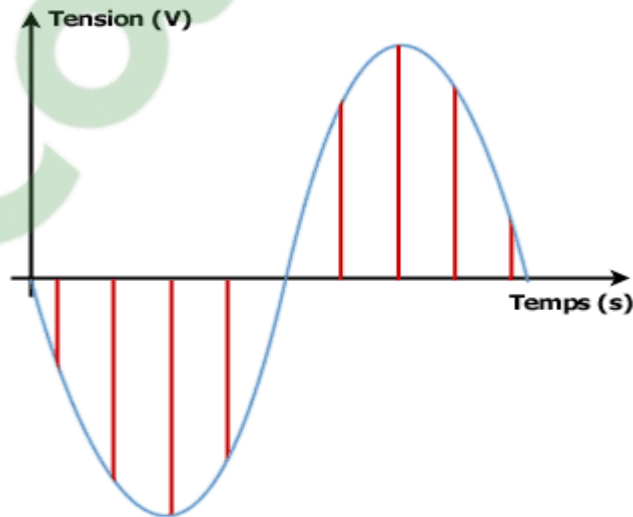


Figure 12 : Représentation temporelle d'un signal échantillonné

- Plus la fréquence d'échantillonnage est élevée, plus on diminue l'intervalle de temps entre deux "prises de valeur", et donc plus la précision est grande. A chaque période il va donc falloir récupérer la "valeur de la courbe", c'est là que l'on va parler de son "24 bit" par exemple. Quand on parle de son "24 bit" il s'agit en fait du niveau de précision avec laquelle on va lire la courbe. On regarde donc la

"hauteur" de la courbe et on la convertit en nombre stocké sur un nombre de bit défini. En effet, si nous stockons par exemple ce nombre sur 1 bit, il n'y a que 0 ou 1 comme possibilités (en gros si la courbe est en dessous de l'axe des abscisses on utilisera 0 et sinon 1). le son ne ressemblera donc à rien car la précision utilisée est bien trop faible. Il faut donc augmenter la précision verticale (la fréquence d'échantillonnage augmentant la précision horizontal pourra donc par exemple coder la hauteur en 16 bit ce qui permet 65536 valeurs possibles (2^{16}). En 24 bits on peut coder 16777216 valeurs différentes, ce qui permet d'avoir une précision plus grande.

4. Le format WAVE [9'] [8']:

Le format PCM : C'est le format de fichier "standard" pour les samples (les enregistrements de sons), car les données sont brutes, c'est-à-dire qu'elles ne sont ni modifiées, ni compressées. Le fichier possède une en-tête de 44 octets (en tout cas en général), permettant de connaître le type du sample: Son format, sa fréquence, le nombre de voies, etc...

Cet en-tête (« header » en anglais) peut, dans certaines variantes du format PCM, avoir une taille supérieure à 44 octets. Mais voyons d'abord le format le plus standard :

4.1. Format [9'] [8']:

- On trouve d'abord la mention « RIFF » dans les 4 premiers octets du fichier (\$52, \$49, \$42, \$42 en hexadécimal. Ce sont les codes ASCII des lettres R, I, F, F).
- On trouve ensuite la taille TOTALE du fichier codée sous la forme d'un entier long (4 octets, donc)
- On trouve ensuite la mention « WAVE » soit 4 caractères ce qui nous donne en hexadécimal : \$57,\$41,\$56,\$45 .

4.2. Descriptif du son [9']:

Fréquence d'échantillonnage, nombre d'octets par échantillons, etc. Les données qu'il contient ainsi que leurs types et leur ordre correspondent exactement à la structure

WAVEFORMAT (cette structure, définie par Microsoft, se retrouve dans pas mal de docs. Elle est aujourd'hui officiellement remplacée par la structure WAVEFORMATEX qui est exactement la même pour les 6 premiers champs, c'est-à-dire tout les champs que l'on retrouve dans l'en-tête du fichier WAVE/PCM).

- Indicateur de zone : « fmt » soit 4 caractères (attention, le dernier caractère est un espace) ce qui nous donne en hexadécimal : \$66,\$6D,\$74,\$20. « fmt » est l'abréviation de « format ». Cet indicateur nous prévient que les informations qui vont suivre concernent le format du son.

- Taille de la structure WAVEFORMAT sous la forme d'un entier long (soit 4 octets). La structure WAVEFORMAT comporte 16 octets. On trouve donc, logiquement, la valeur 16 (ou \$10 en hexadécimal) stockée à cet endroit.

wFormatTag.w (2 octets) : ce champ contient un code correspondant au format exact de codage des données. Pour les fichiers de type PCM, ce champ contiendra la valeur 1.

- nChannels.w (2 octets) : nombre de canaux. On aura la valeur 1 pour les sons mono, 2 pour les sons stéréo, et éventuellement plus pour les sons moins standards.

- nSamplesPerSec.l (4 octets) : nombre d'échantillons par seconde .

- nAvgBytesPerSec.l (4 octets) : nombre d'octets par secondes dépend de :

- du nombre d'échantillon par second

- du nombre d'octets par échantillon

- du nombre de canaux

- nBlockAlign.w (2 octets) : contient la taille totale (en octets) d'un échantillon. Elle dépend :

- du nombre d'octets par échantillon

- du nombre de canaux

Clicours.COM

Pour finir, le fichier contient les données proprement dites.

4.2.1. Datas [9'] :

- Indicateur de zone : «data» soit 4 caractères ce qui nous donne en hexadécimal : \$64,\$61,\$74,\$61. Cet indicateur nous prévient que les informations qui vont suivre sont les données proprement dites.
- Taille des datas sous la forme d'un entier long (soit 4 octets). Cette taille est le nombre total d'octets des datas.

5. Audacity [9]:

Est un logiciel libre pour la manipulation de données audio numériques. Audacity permet d'enregistrer du son numérique par le biais des entrées ligne/micro/cd des cartes sons. Il permet d'éditer (copier, coller, sectionner,...) les sons sur plusieurs pistes, et il est accompagné de divers filtres et effets : pitch, tempo, réduction de bruit, égaliseur, filtres de Fourier, augmentation de fréquences précises, compression, amplification, normalisation, écho, phaser, wahwah, inversion...

5.1. Interface d'AUDACITY [9]:

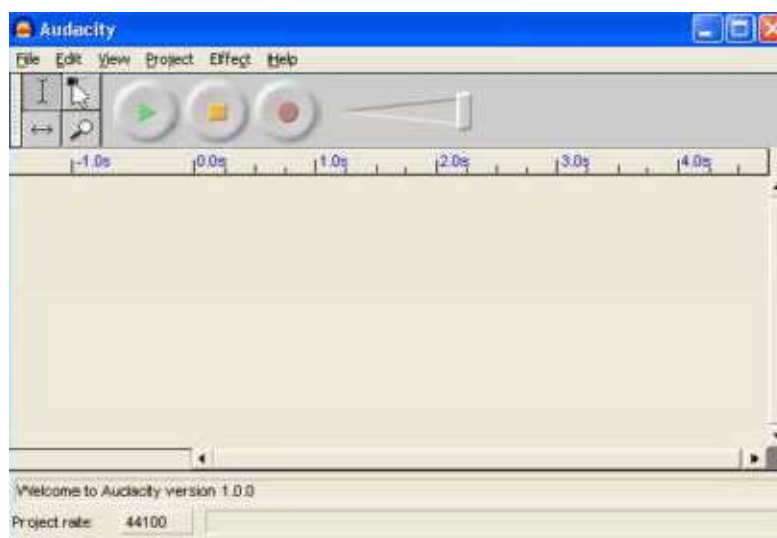


Figure 13 : L-interface de logiciel Audacity

L'interface est plutôt simple : les icônes ressemblent à une console sonore (playstop-record).

La barre d'outils est limitée, mais permet une bonne gestion de son est très facilement :



Figure 14 : La barre d'outils d'Audacity

L'outil sélection pour sélectionner une zone à traiter dans un son.

L'outil panoramisation afin de monter ou de diminuer le son sur des zones du son. Très facile de créer un fade-in ou un fade-out.

L'outil déplacement afin de réaliser un mini-montage.

L'outil loupe pour zoomer la sélection sonore et créer un traitement plus précis.

En peut enregistrer un son uniquement en branchant un microphone dans

L'ordinateur et en cliquant sur le bouton « record ».on Clique sur le bouton« stop » pour terminer l'enregistrement. Dès lors, en peut commencer à traiter notre son.

En peut également importer un son à partir d'un fichier (wav ou mp3) file > open.

Malheureusement, le logiciel ne permet pas de prendre directement un fichier sonore d'un disque compact. Vous devez auparavant extraire le fichier audio de votre disque compact et le convertir en wav ou en mp3.

- Audiograbber : ou tout autre logiciel de conversion (**cda to wav** ou **cda to mp3**) est indiqué.
- Traitement sonore :

Une fois le son mis dans notre logiciel, les ondulations du son est comme suit :

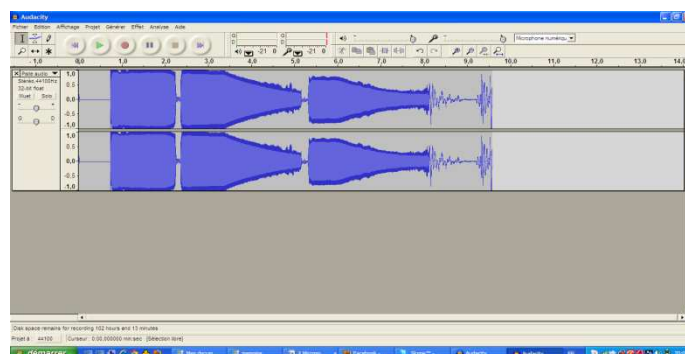


Figure 15 : Enregistrement du son sur Audacity

Nous pouvons utiliser les outils pour sélectionner des zones, pour créer un fade-in, fade-out ou pour déplacer le son dans le « timeline ».

On utilise le menu « effect » pour ajouter un traitement à une zone présélectionnée de notre son. Pour appliquer un effet à tout notre son, on sélectionne le son au complet.

Le menu « effect » :



Figure 16 : Le menu des effets d'Audacity

Voici un survol des effets disponibles :

- **AMPLIFY** : permet d'augmenter ou de diminuer le volume de notre son (parfois le son enregistré est trop faible). Le logiciel calcule le maximum permis à l'augmentation de notre son sans causer de distorsion ou de « clics » indésirables.
- **BASSBOOST** : permet d'amplifier les sons de fréquences graves afin d'avoir un effet de « basses » plus prononcés. Le logiciel nous suggère d'y aller doucement avec ce filtre en ajustant à 12dB.
- **ECHO** : Filtre qui ajoute un délai avant d'entendre l'écho de notre son. Peut augmenter la durée de votre son.
- **FADE IN – FADE OUT** : Ouverture et fermeture graduelle.
- **FFT FILTER** : Ce filtre nous permet d'isoler certaines fréquences de notre son.

Filtre en construction...

- **INVERT** : Fais jouer le son à l'envers (gauche dans droite et droite dans gauche). Cela ne sert à rien sauf pour dans la création d'effets panoramiques.

- NOISE REMOVAL : Enlève les « tics » et les bruits non désirables (climatisation de lieu, etc.) Mais s'utilise en deux étapes. On doit d'abord sélectionner le « bruit » à retirer sans autre son audible. Par la suite, le logiciel calcule la fréquence du son à retirer dans le reste de l'échantillon.
- PHASER : Crée un effet particulier à partir du signal de base. on peut ajouter du feedback et des variations de fréquence.
- REVERSE : Joue le son... à l'envers ! Très utilisé pour trouver des phrases démoniaques dans des « tounes » des beatles ou pour créer des sons nouveaux à partir de son très simples.
- WAHWAH : Effet de « guitare » électrique qui varie les ondes de gauche à droite (panoramique) et de hautes à basses fréquences.
- FREEVERB2 : Permet de créer un effet de « reverb » autrement dit, un écho contrôlé. Les options sont un peu difficiles d'accès, mais on peut créer un effet d'église (un gros effet de réverbération) ou un effet extérieur.

Pour pouvoir implémenter les algorithmes de traitement du signal vocal sous environnement Visual Basic il faut au préalable faire la conversion WAVE-HEXADECIMAL .Pour cela nous avons mis à contribution le logiciel WATEX.

6. La base de donné [11']:

6.1. Définition :

En informatique, une base de données (BD) est un lot d'informations stockées dans un dispositif informatique. Les technologies existantes permettent d'organiser et de structurer la base de données de manière à pouvoir facilement manipuler le contenu et stocker efficacement de très grandes quantités d'informations.

L'organisation logique des données se fait selon un modèle de données. La structure physique des fichiers comporte des index destinés à accélérer les opérations de recherche et de tri.

La motivation fondamentale de l'organisation des données est de permettre d'effectuer des requêtes sur celles-ci. Le modèle de données relationnel est aujourd'hui le plus utilisé parce qu'il est formellement démontré que ce type de représentation permet de résoudre toutes requêtes.

Le logiciel qui manipule les bases de données est appelé système de gestion de base de données (SGBD). Il permet d'organiser, de contrôler, de consulter et de modifier la base de données. Les opérations sont parfois formulées dans un langage de requête tel que SQL .

6.2. Utilité d'une base de données [10] :

Une base de données permet de mettre des données à la disposition d'utilisateurs pour une consultation, une saisie ou bien une mise à jour, tout en s'assurant des droits accordés à ces derniers. Cela est d'autant plus utile que les données informatiques sont de plus en plus nombreuses. Une base de données peut être locale, c'est-à-dire utilisable sur une machine par un utilisateur, ou bien répartie, c'est-à-dire que les informations sont stockées sur des machines distantes et accessibles par réseau.

L'avantage majeur de l'utilisation de bases de données est la possibilité de pouvoir être accédées par plusieurs utilisateurs simultanément.

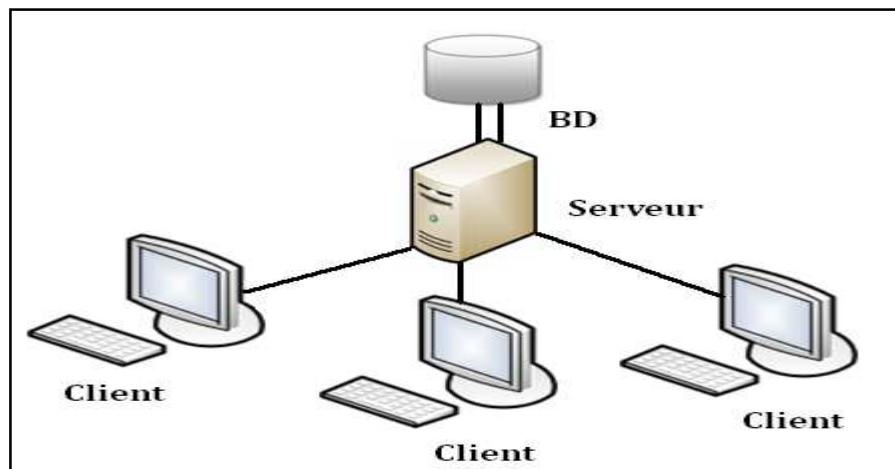


Figure 17 : La base de données dans un réseau

7. Conclusion :

Dans chapitre nous avons étudiés les différentes étapes qu'on doit effectuer pour avoir une information utile de notre signal c.à.d. l'enregistrement de la voix par le logiciel Audacity et la conversion de ceci par watex_1 pour avoir le format hexadécimale quelle doit être utiliser directement dans le traitement de notre signal et de même nous avons parlés de quelque notion sur la base de donné et sa utilité.

Chapitre IV :

**Caractéristiques Spectro-Temporelle
du signal vocal multiparamétrique**

Chapitre IV : Caractéristiques Spectro- Temporelle du signal vocal multiparamétrique

1. Introduction :

Pour mieux comprendre l'analyse de la production de la parole on détaille les différentes méthodes d'évaluations objectives de la production vocal pour diagnostiquer la maladie cordale, la gravité de celle-ci, pour évaluer le degré de dysphonies pour détecter le pronostic et pour analyser les modifications avant et après traitement.

Cette étude, basé sur plusieurs centaines de cas, décrire statiquement les méthodes d'évaluations objectives généralement au moyen d'histogrammes en fonction de différents paramètres tels que le débit d'aire buccal et le volume d'air associés au temps maximal de phonation, la fréquence fondamentale usuelle(F0)...

2. L'analyse acoustique [7] :

Bien que l'analyse acoustique nous paraît indispensable à l'évaluation du changement vocal .De très nombreux paramètres peuvent être mesurés ; nous n'avons pris en compte que certains d'entre eux :

2.1. La hauteur : son harmonique [8]

L'oreille détermine la hauteur d'un son par la fréquence la plus facilement décelable par le cerveau. Le cerveau va analyser un signal harmonique comme un signal dont la hauteur est la fréquence de la fréquence fondamentale, ou plus précisément, la fréquence associée à la période du signal. Petit rappel sur les signaux sonores périodiques ou signal harmonique :

$$X(t) = 5.\sin (2\pi 450t) + 3.\sin (2\pi 100t) + 0.5.\sin (2\pi 350t) \quad (\text{VI.1})$$

On voit que le signal obtenu est un signal périodique, de période 1/50 s (correspondant à la plus petite fréquence). On peut voir que toutes les fréquences sont des multiples de la plus petite des fréquences. Ce résultat est généralisé par la loi de Fourier : tout signal périodique peut s'exprimer comme une somme de sinusoides de fréquences

Chapitre IV : Caractéristiques Spectro-Temporelle du signal vocal multiparamétrique

multiples d'une première fréquence, cette dernière s'appelle la fréquence fondamentale f_0 . Les partiels sont alors appelés les harmoniques. L'ordre des harmoniques est appelé le rang.

L'harmonique de rang n aura une fréquence $f = nf_0$. L'auditeur entendra alors un son dont la hauteur correspond à $f = f_0$. Un signal harmonique est donc un signal périodique.

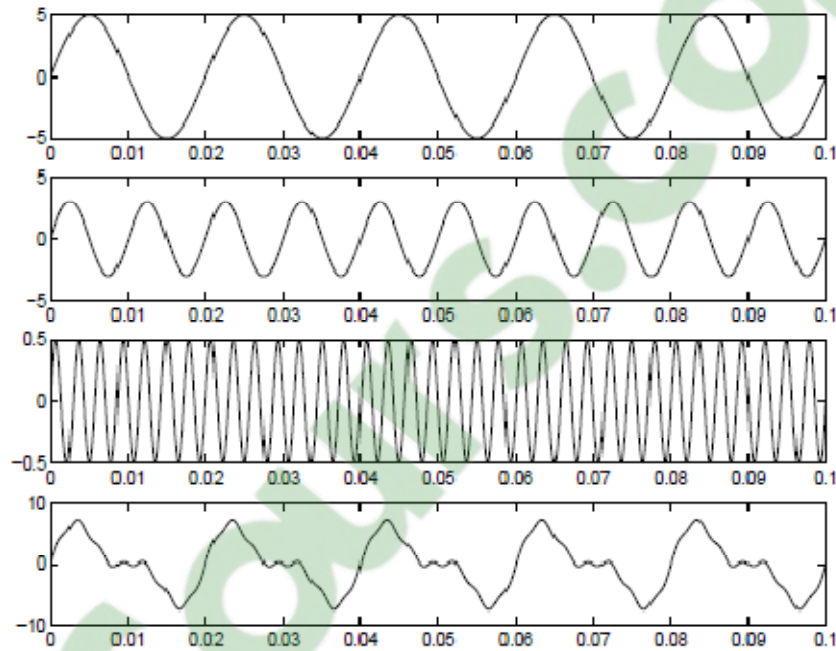


Figure 18 : Signal sonore périodique, somme de sons purs

La période de ce signal correspond à $\frac{1}{f_0}$ avec f_0 la fréquence fondamentale.

L'oreille est donc sensible à cette fréquence et la hauteur perçue dépend de cette fréquence. Plus la fréquence est grande (hautes fréquences), plus le son paraîtra aigu et plus la fréquence est petite (basses fréquences), plus le son paraîtra grave.

2.2. L'intensité sonore [8]:

L'intensité sonore est la sensation auditive qu'un son est plus fort qu'un autre. L'oreille humaine permet d'écouter des sons extrêmement faibles comme le bruit d'une goutte d'eau s'écrasant sur le sol d'une grotte silencieuse presque aussi facilement que le son extrêmement fort provoqué par les chutes du Niagara. Ou encore, on peut entendre le chuchotement d'une personne dans une salle vide et le bruit fait par 200 personnes discutant dans un hall.

Chapitre IV : Caractéristiques Spectro-Temporelle du signal vocal multiparamétrique

L'étendue de l'intensité des sons audibles est donc énorme. On peut se demander Combien de gouttes d'eau composent une chute d'eau, surement des milliards. Pourtant on ne peut pas dire que le son d'une chute d'eau plus un milliard de fois plus fort que celui de la simple goutte. L'intensité sonore perçue n'est pas directement proportionnelle à l'amplitude des signaux. Il ne suffit donc pas de doubler l'amplitude d'un son pour que le son paraisse deux fois plus fort. L'oreille fonctionne en fait comme une échelle logarithmique.

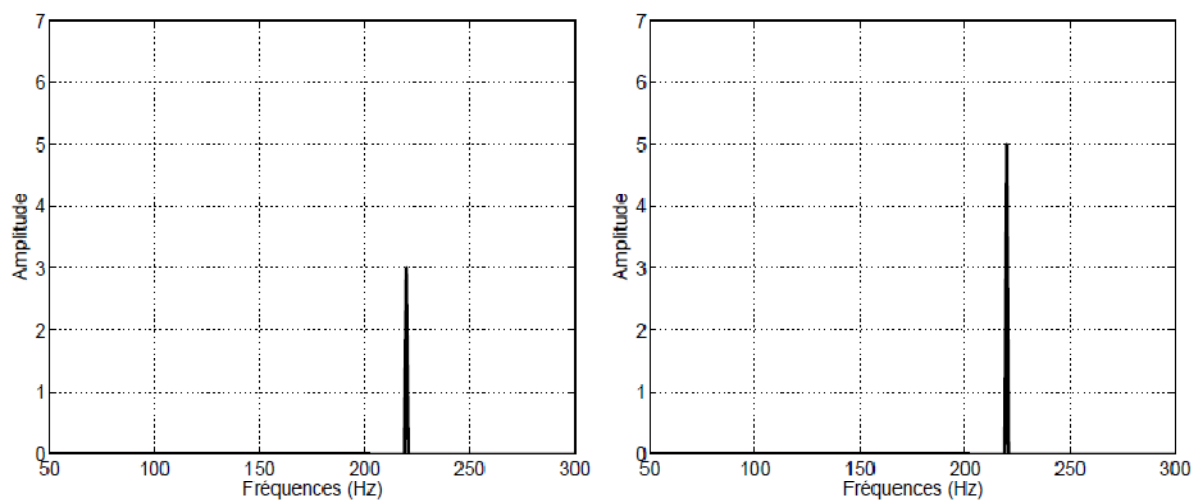


Figure 19 : Deux sons purs d'amplitude différente

2.3. La durée [8] :

La durée d'un son est le temps que dure le son. Elle peut se mesurer sur la représentation temporelle d'un son.

Chapitre IV : Caractéristiques Spectro-Temporelle du signal vocal multiparamétrique

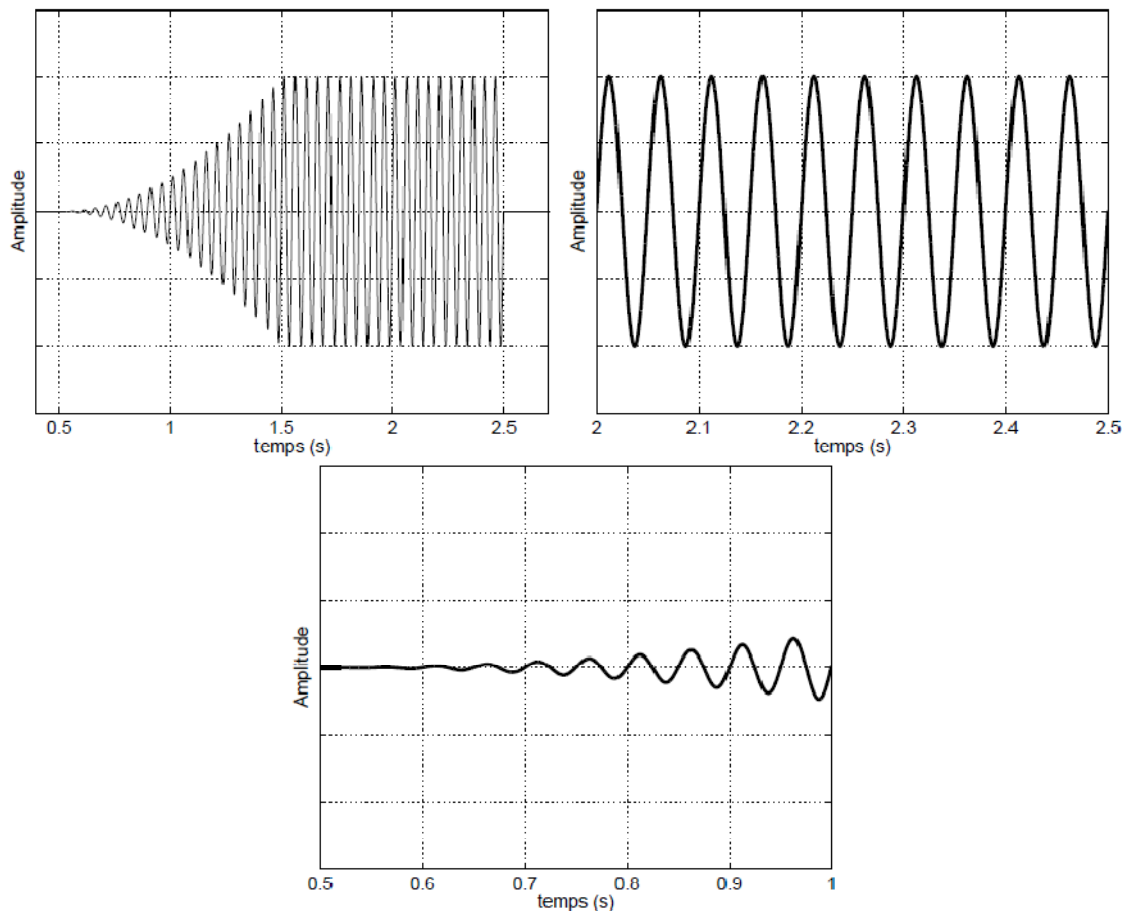


Figure 20 : Représentation temporelle d'un son : signal global et zoom

2.4. Le timbre [8] :

Imaginons deux instruments, comme une flûte et une guitare jouant la même note, pendant le même temps et avec un volume équivalent. On a donc deux sons de même hauteur, même durée et de même intensité. Pourtant, l'oreille différencie ces deux instruments et les trouve même très différents. Pourquoi ? Ce qui les différencie est ce que l'on appelle le "timbre", qui est propre à chaque instrument. On peut se faire une idée du timbre en regardant le sonogramme de mêmes notes jouées par deux instruments différents. La fréquence fondamentale sera la même mais les harmoniques auront des amplitudes différentes ou ne dureront pas le même temps.

Le timbre a des caractéristiques temporelles et fréquentielles :

- si on regarde la représentation fréquentielle d'un son pur issu de la vibration d'un diapason (La (440 Hz)) et la même note jouée au piano. Le timbre est très

Chapitre IV : Caractéristiques Spectro-Temporelle du signal vocal multiparamétrique

différent. Ceci vient de la composition spectrale des signaux sonores. En effet, le spectre du diapason ne contient qu'un pic à 440 Hz et le spectre de la note du piano contient plusieurs pics (les harmoniques) : un pic à 440 Hz et les harmoniques.

De même, si on regarde le spectre de deux instruments différents (Figure 21).

On peut voir que les deux sons vont être de la même hauteur car ils ont la même fréquence fondamentale, mais l'amplitude des harmoniques est différente. Ainsi, à partir de deux spectres de deux notes de même hauteur jouées par deux instruments mélodiques différents, les deux spectres n'auront pas la même composition.

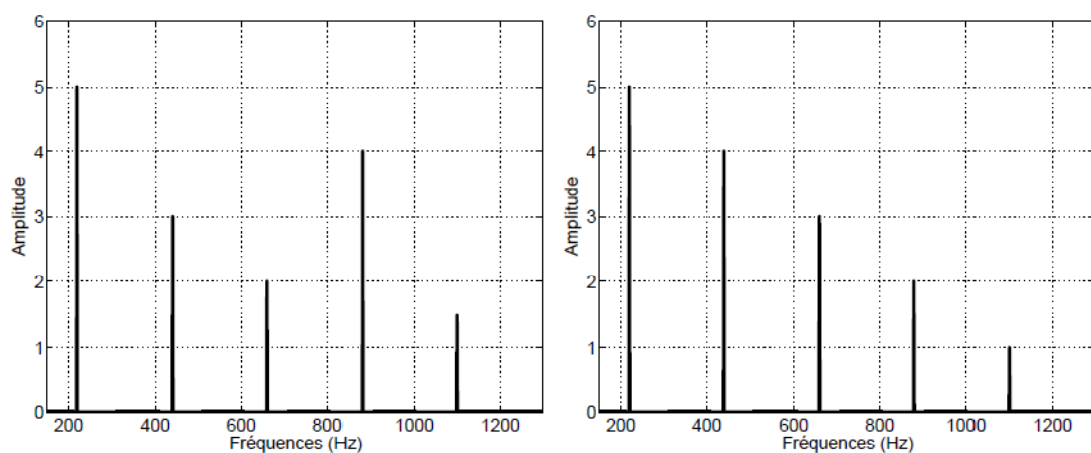


Figure 21 : Deux spectres de signaux harmoniques de même fréquence fondamentale (220Hz)

La fréquence fondamentale sera la même mais les harmoniques n'auront pas tous la même amplitude ou certains harmoniques seront présents dans un spectre et pas dans l'autre.

- si on regarde la représentation temporelle de deux même notes jouées par deux instruments différents (Figure 22), la période est la même, mais la forme est différente. Ceci veut dire que la note jouée est la même mais le timbre va être différent.

Chapitre IV : Caractéristiques Spectro-Temporelle du signal vocal multiparamétrique

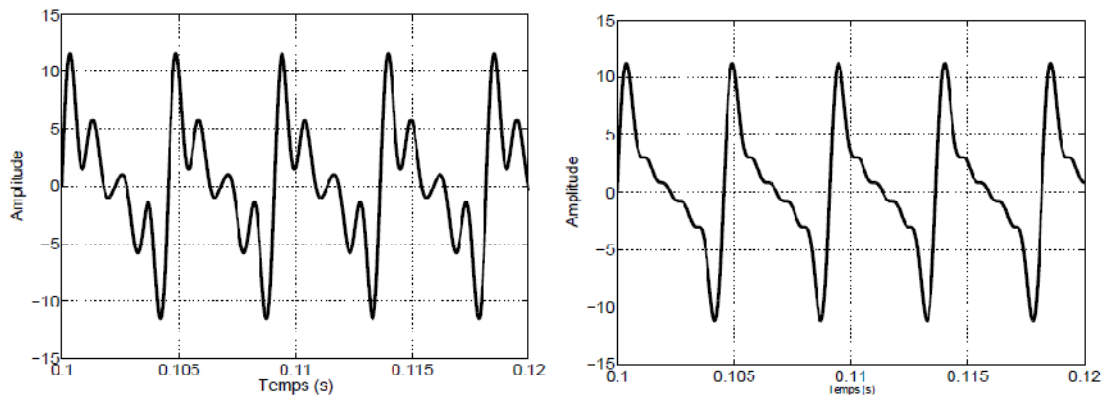


Figure 22 : Deux représentations temporelles de signaux harmoniques de même fréquence fondamentale (220 Hz)

Si on regarde l'enveloppe (l'enveloppe est la courbe reliant les maxima du son au cours du temps), on peut voir qu'elle est toujours du même type (Figure 23).

Chacune des différentes parties (attaque, décroissance, maintien et relâche) Peut varier en fonction de l'instrument (et du musicien) et contribuent à une différence de timbre.

- si on regarde la représentation temps/fréquence (sonogramme) de différents instruments jouant la même note, on peut voir que les harmoniques évoluent différemment ceci joue aussi sur le timbre.

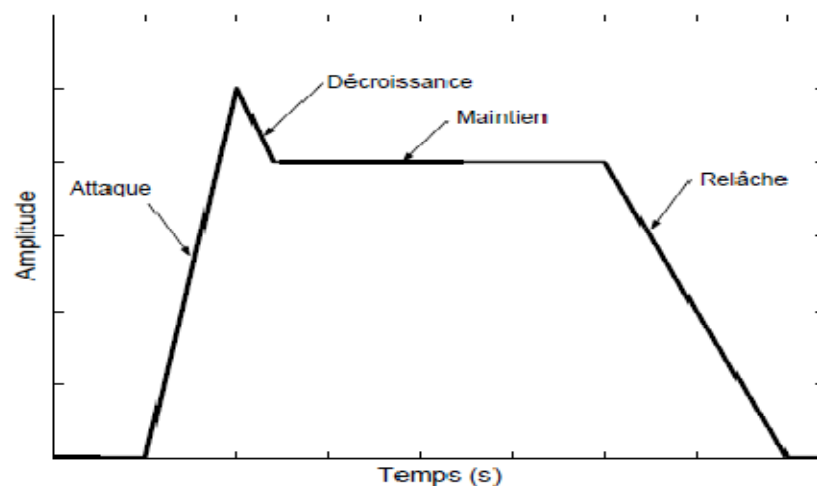


Figure 23 : Enveloppe (temporelle) d'une note jouée par un instrument

Chapitre IV : Caractéristiques Spectro-Temporelle du signal vocal multiparamétrique

2.5. La fréquence fondamentale F_0 [7]:

La fréquence fondamentale moyenne est calculée à partir de la valeur de la période moyenne du cycle.

$$F_0 = \frac{1}{N} \sum_{i=1}^N F_0^{(i)} \quad (\text{VI.2})$$

$$F_0^{(i)} = \frac{1}{T_0^{(i)}} \quad (\text{VI.3})$$

La période fondamentale :

$$T_0^{(i)}, i=1,2,\dots,N$$

La fréquence la plus basse ainsi que la fréquence la plus aiguë enregistrées lors de l'émission du « a » tenu, sont également mesurées.

Pour être valides, ces mesures nécessitent un signal périodique. Elles sont nécessaires aux mesures de stabilité fréquentielle décrites plus loin.

2.6. Les Formants [6'] :

On désigne par formant (acoustique) d'un son de parole l'un des maxima d'énergie du spectre sonore de ce son de parole. Le spectre sonore représente la distribution en fréquences de l'énergie du signal de parole. Le spectre sonore se mesure par analyse fréquentielle d'un segment très court (aussi appelé fenêtre d'analyse dans le vocabulaire du traitement du signal), de 10 à 40 millisecondes, du signal de parole. C'est une information variable au cours du temps, dépendant du type son de parole, c'est-à-dire en première approximation, du phonème prononcé, mais plus complètement, de la manière particulière dont il a été prononcé.

Comme le spectre sonore, dont ils représentent l'un des aspects, les formants sont mesurés sur des segments de parole très courts, et ils sont donc variables au cours du temps, dépendant du phonème prononcé et de la manière dont il est prononcé. On

Chapitre IV : Caractéristiques Spectro-Temporelle du signal vocal multiparamétrique

numérote les formants (qui sont donc les maxima du spectre sonore) en allant des basses fréquences vers les hautes fréquences. On les désigne par la notation physique F_i (mesurée en hertz) en partant du premier formant F_1 dans les fréquences basses. La notation F_0 est réservée à la fréquence fondamentale, dont les variations au cours du temps constituent l'intonation de la parole.

Les deux premiers formants F_1 et F_2 sont des caractéristiques particulièrement adaptées à la description des voyelles, car les voyelles sont des sons relativement stationnaires de la parole. Chaque voyelle se caractérise ainsi par son timbre spécifique, déterminé en première approximation par F_1 et F_2 . La discipline phonétique a permis d'établir une corrélation entre les valeurs mesurées en hertz de F_1 et F_2 et les articulations de l'appareil phonétique nécessaires à la réalisation des voyelles. Ainsi F_1 est corrélé avec l'aperture (ouverture de la bouche) et le F_2 avec la position antérieure (valeur élevée de F_2), postérieure (valeur basse de F_2) de la langue. Le troisième formant F_3 est également une caractéristique intéressante, corrélée avec la configuration des lèvres pour les voyelles antérieures.

On présente dans la table ci-dessous les valeurs typiques des formants F_1 et F_2 pour les voyelles du français. Si on reporte ces valeurs dans un graphique représentant F_1 et F_2 en abscisse et en ordonnée, on obtient le fameux triangle vocalique de la discipline phonétique.

Voyelle (API)	F_1 (Hz)	F_2 (Hz)
[u]	320	800
[o]	500	1000
[ɑ]	700	1150
[a]	1000	1400
[ø]	500	1500
[y]	320	1650
[ɛ]	700	1800

Chapitre IV : Caractéristiques Spectro-Temporelle du signal vocal multiparamétrique

[e]	500	2300
[i]	320	3200

Tableau 3 : Fréquences typiques des formants.

- **Le triangle vocalique [7'] :**

Le triangle vocalique est une représentation graphique de l'appareil vocal humain. Il permet de classer les voyelles selon deux axes: horizontalement la profondeur du point d'articulation (position de la langue dans la bouche (antérieure - centrale - postérieure)) et verticalement le degré d'aperture. La première représentation du triangle vocalique est due au médecin allemand Christoph Friedrich Hellwag (1754-1835) .Ce triangle enregistre des variations selon les langues et dans 10% des langues, dont l'anglais, les voyelles ne se conforment pas à ce schéma triangulaire; on utilise donc également le terme de trapèze vocalique

3. Les mesures de stabilité fréquentielle [7]:

3.1. La perturbation à court terme : le jitter

Le jitter est la variation de la période du cycle vibratoire, mesurée d'un cycle à l'autre (autrement dit, cycle-à-cycle).

Chez le sujet sain, cette variation cyclique est physiologique et inévitable en raison de petites irrégularités mécaniques, tissulaires et fonctionnelles (force et pression légèrement différentes d'un cycle à l'autre, par exemple).

Le jitter est qualifié de « mesure de perturbation à court terme », de la période du cycle. Le terme « perturbation » fait référence à une irrégularité minime du phénomène étudié.« A court terme » fait référence à la comparaison réalisée d'un cycle à l'autre.

La période vibratoire normale est d'environ 10 millisecondes chez un homme et de 5 millisecondes chez la femme. Le jitter absolu peut être mesuré en millisecondes et au départ été défini comme étant le pourcentage des perturbations égales ou supérieures à 0.5 millisecondes.

Chapitre IV : Caractéristiques Spectro-Temporelle du signal vocal multiparamétrique

Le jitter est plus grand si la période est plus longue (et donc si la fréquence de la voix est plus basse). Il est donc plus intéressant d'utiliser un jitter « relatif » en divisant la valeur moyenne de la perturbation par la période moyenne.

$$jitt = \frac{\frac{1}{N-1} \sum_{i=1}^{N-1} |T_0^{(i)} - T_0^{(i+1)}|}{\frac{1}{N} \sum_{i=1}^N T_0^{(i)}} \quad (\text{VI.4})$$
$$T_0^{(i)}, i=1,2,\dots,N$$

Le jitter varie chez un même sujet, d'une émission vocale à l'autre, au cours d'une même séance d'acquisition et d'enregistrement. Pour toutes ces raisons, et malgré les nombreuses recherches réalisées sur le jitter, le jitter est plus élevé à l'attaque sonore et en fin de tenue vocalique il augmente aussi avec l'intensité.

Une variation avec le type de voyelle émise est possible mais n'est pas confirmée: la voyelle « a » aurait un jitter plus élevé Les femmes possèderaient un indice de jitter plus élevé que les hommes.

Chez le sujet atteint de pathologie cordale, le jitter augmente en raison d'une instabilité vibratoire cordale plus grande, et est corrélé avec la raucité subjectivement audible.

3.2. Les perturbations à moyen terme [7] :

La fluctuation de F0 est susceptible de se manifester non pas de cycle à cycle, mais à plus long terme, soit sur la durée d'un phonème. Nous avons pris en considération l'étendue de fluctuation de la fréquence et l'écart-type de F0.

- L'étendue de la fluctuation de la fréquence est l'écart entre la fréquence la plus basse et la fréquence la plus élevée enregistrée.
- L'écart-type de F0 (STD). La détermination de l'écart-type de F0 est également liée à la raucité subjectivement audible.

$$STD = \sigma = \sqrt{\frac{1}{N} \sum_{i=1}^N ((F_0 - F_0^{(i)})^2} \quad (\text{VI.5})$$

$$F_0 = \frac{1}{N} \sum_{i=1}^N F_0^{(i)} \quad , \quad F_0 = \frac{1}{T_0^{(i)}}$$

$$T_0^{(i)}, i=1,2,\dots,N$$

3.2.1. La mesure de stabilité de l'intensité : le shimmer[7] :

Le shimmer est la mesure de stabilité de l'intensité et correspond donc à la mesure du jitter réalisée pour l'évaluation de la stabilité fréquentielle.

Le shimmer est une mesure de perturbation à court terme de l'amplitude du cycle vibratoire : c'est la différence d'amplitude de cycle à cycle.

$$shim = \frac{\frac{1}{N-1} \sum_{i=1}^{N-1} |A^{(i)} - A^{(i+1)}|}{\frac{1}{N} \sum_{i=1}^N A^{(i)}} \quad (\text{VI.6})$$

$$A^{(i)}, i=1,2,\dots,N \text{ données d'amplitude crête à crête}$$

Comme pour le jitter, nous avons choisi une mesure relative :

Le shimmer %. Il est défini comme la fluctuation moyenne de l'amplitude du cycle rapporté l'amplitude moyenne du cycle vibratoire. Tout Comme le jitter, sa mesure et son mode de calcul sont différents d'un logiciel à l'autre.

Chez le sujet sain, le shimmer (dB) est plus élevé chez les hommes sur les voyelles « i » et « u », mais pas sur la voyelle « a » .

Chez le sujet atteint d'une lésion bénigne cordale, le shimmer augmente pour la même raison que celle évoquée pour l'augmentation du jitter l'instabilité vibratoire est augmentée par la pathologie cordale.

3.2.2. Le rapport signal-bruit [7] :

Chez un locuteur sain, du bruit (phénomène non périodique) est associé à la fourniture harmonique. Ce niveau de bruit « normal » est, selon les auteurs à mettre en relation avec l'importance de la perturbation fréquentielle. Lorsque la voix est rauque, le bruit s'intensifie et remplace la fourniture harmonique. Deux mesures permettent de quantifier le rapport entre l'énergie du bruit et l'énergie harmonique du signal : le rapport signal/bruit (ou bruit/signal : noise-to-harmonics ratio) et l'énergie normalisée du bruit. Chez le sujet sain, le rapport signal/bruit est identique dans les deux sexes.

Chez le sujet pathologique, le rapport bruit/signal augmente, et ce dernier est corrélé avec la raucité.

4. L'évaluation de la fonction phonatoire [7] :

L'amélioration fonctionnelle vocale qui est recherchée grâce à la réalisation d'un geste microchirurgical n'est pas seulement acoustique mais aussi aérodynamique. Il convient en effet d'améliorer la voix, mais aussi de réduire l'effort nécessaire à son émission. L'amélioration du rendement aérodynamique laryngé est la clé d'une diminution de la fatigue vocale et de l'augmentation du confort vocal quotidien du patient.

4.1. Le temps maximum de phonation [7] :

Le temps maximum de phonation est une mesure simple à réaliser : c'est le temps maximum de tenue d'un « a » à intensité et fréquence spontanées confortables, après une inspiration maximum.

Il est enregistré et mesuré en même temps que le débit aérien moyen phonatoire.

Chez le sujet sain, le temps maximum de phonation est plus grand chez les hommes que chez les femmes en raison de leur plus grande capacité vitale. Il est compris entre 25 et 35 secondes chez l'homme, et entre 15 et 25 secondes chez la femme. Un temps maximum inférieur à 10 secondes chez la femme, et à 15 secondes chez l'homme, est considéré comme anormal. En cas de lésion cordale bénigne, le temps maximum de phonation se réduit.

Chapitre IV : Caractéristiques Spectro-Temporelle du signal vocal multiparamétrique

4.2. La pression sous-glottique phonatoire [7] :

La pression sous-glottique en phonation est conditionnée par l'importance de la fermeture glottique, par la force expiratoire et par l'importance de la force nécessaire à la mise en oscillation des cordes vocales. La pression sous-glottique est corrélée à l'intensité du son émis: chaque fois que la pression minimale nécessaire à l'entrée en oscillation des cordes vocales (« phonation treshold pressure ») est doublée, l'intensité augmente de 6dB.

Une élévation de la pression sous-glottique s'observe chez les patients atteints de pathologie cordale. Cette élévation de pression sous-glottique est multi conditionnée : D'une part, les modifications tissulaires cordales, en augmentant la masse et la rigidité cordales, rendent inévitable l'existence d'une pression sous glottique plus élevée d'autre part, le forçage compensatoire entraîne également une élévation de la pression sous-glottique.

5. Les altérations observées en cas de pathologie cordale bénigne [7] :

On va décrire, le sens de l'altération de la variable mesurée en présence d'une pathologie cordale bénigne altérant les caractéristiques biomécaniques du vibreur laryngé. on les résumons ci-dessous.

Amplitude du mouvement vibratoire : diminution

Fermeture glottique : diminution

Jitter % : augmentation

Shimmer % : augmentation

Fluctuation F0 : augmentation

Ecart-type F0 : augmentation

Rapport bruit/harmonique : augmentation

Temps maximum de phonation : diminution

Pression sous-glottique : augmentation

Débit aérien moyen : augmentation

Chapitre IV : Caractéristiques Spectro-Temporelle du signal vocal multiparamétrique

Une relation directe entre le grade de la pathologie et la valeur de la variable, existe pour les variables suivantes : le jitter %, la fluctuation de F_0 , l'écart-type de F_0 , le shimmer %, le rapport bruit/signal, la pression sous-glottique et le débit aérien moyen.

6. Conclusion :

Dans cette partie nous avons étudié les caractéristiques spectro-temporelle du signal vocal qui ont une relation directe avec le grade de la pathologie qui sont basé sur le calcul de la fréquence fondamentale calculé à partir du spectre de notre signal vocal.

Chapitre V :

Protocole expérimentale

Chapitre V : Protocole expérimentale

1. Introduction :

Après avoir étudié, dans les chapitres précédents, tous les aspects liés à l'appareil phonatoire et à la production du son ainsi qu'aux différents paramètres dédiés à sa caractérisation multiparamétrique, nous présentons dans ce chapitre, la conception et l'implémentation de la base de données interactive, du site web permettant d'administrer le système Télémédecine, ainsi que l'implémentation des algorithmes de traitement et de calcul des paramètres caractérisant le signal vocal.

2. Conception de la base de données Télémédecine :

2.1. Conception de la base de données :

La base de données comprend deux tables :

- 1) une table abritant le questionnaire clinique du patient qui constituera son dossier médical personnalisé.
- 2) une table consacré aux différents acteurs de cette base de données (administrateurs et praticiens de la médecine ayant accès à cette base de données).

elle comprend également des requêtes d'appel selon l'ID, le nom, la pathologie.

- le tableau clinique (les renseignements du patient) est structuré de la manière suivante :

ID : le numéro d'immatriculation du patient.

Nom :

Prénom :

Sexe :

Âge :

Adresse :

Téléphone :

Email :

Photo :

Profession :

Tabac : est ce que le patient est tabagique ou non.

Alcool : est ce que le patient est alcoolique ou non.

Pathologie : la pathologie du patient.

Rapport clinique : le rapport du patient.

Antécédent familiaux :

examen-CDS :

Date d'admission :

- la deuxième table « user » concernent les administrateurs et les personnes (praticiens) qui peuvent accéder à cette base de données :

id-user : le numéro d'immatriculation de l'administrateur.

Nom :

Prénom :

Login_user :

pass-user :

3. Sélection des Clés primaires :

Dans une base de données relationnelle, une **clé primaire** est une contrainte d'unicité qui permet d'identifier de manière unique un enregistrement dans une table. Une clé primaire peut être composée d'un ou de plusieurs champs de la table. Deux lignes distinctes de la table ne peuvent pas avoir les mêmes valeurs pour les champs définis au niveau de la clé primaire.

La clé primaire d'une table doit se placer sur des champs qui permettent d'identifier chaque ligne de la table. Il peut donc sembler intéressant de placer une clé primaire sur un numéro de sécurité sociale par exemple. Dans notre cas on a choisi les clés primaires suivant :

Pour la première table patient nous avons choisie la clé primaire pour ID du patient.

Pour la deuxième table la clé primaire est choisie pour id-user.

4. L'implémentation de la base de données avec le SGBD :

4.1. Système de Gestion de Bases de Données[13] :

4.1.1. Définitions :

Un SGBD est un logiciel qui prend en charge la structuration, le stockage, la mise à jour et la maintenance d'une base de données. Il est l'unique interface entre les

informaticiens et les données, ainsi qu'entre les utilisateurs et les données (consultation et mise à jour).

4.1.2. Fonctions des SGBD :

- ✓ Description des données
- ✓ Recherche de données
- ✓ Mise à jour des données

On désigne par mise à jour : insertion, modification, suppression ; cela grâce à un langage de manipulation de données (LMD).

- ✓ Transformation des données

Ex. Changement de format : date sur 2 chiffres date sur 4 chiffres.

- ✓ Contrôle de l'intégrité des données

Exactitude des données stockées dans la base par le maintien de la cohérence entre les copies multiples d'une même donnée.

- ✓ Gestion de transactions et sécurité

- Gestion des transactions : s'assurer qu'un groupe de mises à jour est totalement exécuté ou pas du tout.
- Sécurité : personnalisation des accès à la base par mots de passe [11].

4.2. Qu'est ce que MySQL?

MySQL est un système de gestion de base de données (SGBD). Il est un véritable serveur de base de données SQL multi-utilisateur et multi-threadé.

MySQL est une configuration client/serveur. SQL est un langage standardisé qui rend facile le stockage, la mise à jour et l'accès à l'information. Par exemple, on peut utiliser le SQL pour récupérer des informations sur un étudiant ou stocker des informations sur un site web. Les principaux objectifs de MySQL sont la rapidité, la robustesse et la facilité d'utilisation. MySQL a été originellement développé par ce que nous avons besoin d'un serveur SQL qui puisse gérer des grandes bases de données de manière plus rapide comme ceux que pouvaient offrir les distributeurs de bases de données.

La base sur laquelle MySQL est construite est un ensemble de routines qui ont été largement éprouvées pendant des années dans un environnement de production exigeant.

Selon le type d'application, sa licence est libre ou propriétaire. Il fait partie des logiciels de gestion de base de données les plus utilisés au monde, autant par le grand public (applications web principalement) que par des professionnels, en concurrence avec Oracle et Microsoft SQL Server.

Pour faire notre implémentation de base de données nous avons travaillé avec l'outil web phpMyAdmin qui va nous faciliter le travail avec le serveur MySQL.

4.3. Administration avec l'outil web phpMyAdmin :

L'outil phpMyAdmin est développé en PHP et offre une interface intuitive pour l'administration des bases de données du serveur.

Il est téléchargeable sur le site: <http://phpmyadmin.sourceforge.net>

Il s'agit de l'une des plus célèbres interfaces pour gérer une base de données MySQL sur un serveur PHP. De nombreux hébergeurs, qu'ils soient gratuits ou payants, le proposent ce qui permet à l'utilisateur de ne pas avoir à l'installer.

Cette interface pratique permet d'exécuter, très facilement et sans grandes connaissances dans le domaine des bases de données, de nombreuses requêtes comme les créations de table de données, les insertions, les mises à jour, les suppressions, les modifications de structure de la base de données. Ce système est très pratique pour sauvegarder une base de données sous forme de fichier **.sql** et ainsi transférer facilement ses données. De plus celui-ci accepte la formulation de requêtes SQL directement en langage SQL, cela permet de tester ses requêtes par exemple lors de la création d'un site et ainsi de gagner un temps précieux.

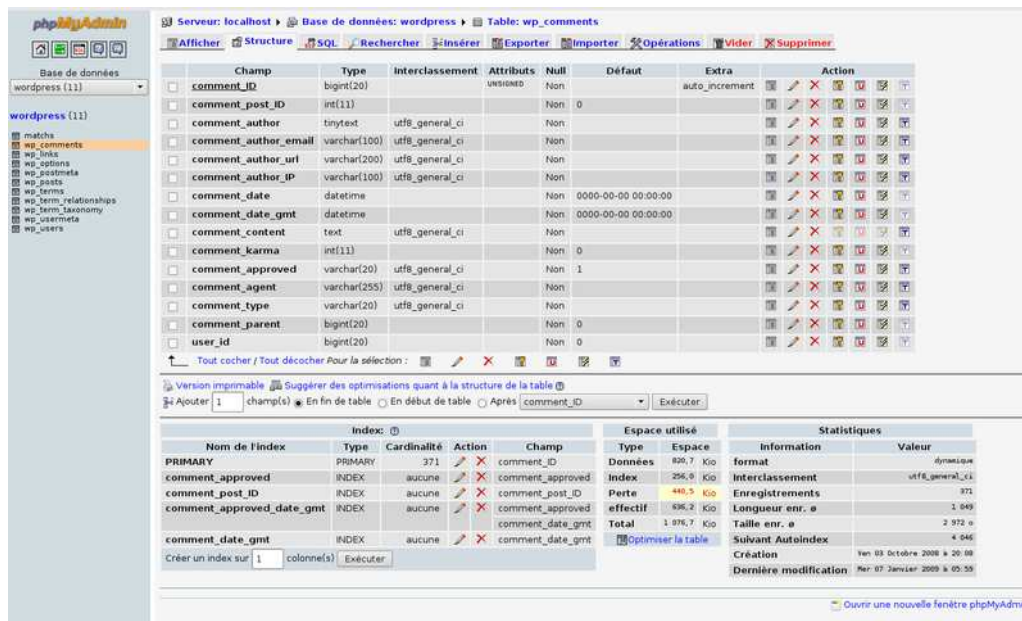


Figure 24 : Capture d'écran de phpMyAdmin

4.3.1. Création de la base de données :

Le SGBD MySQL que nous allons utiliser travaille avec le langage des requêtes Structurées SQL (Structured Query Language). Pour la création d'une base de données, on utilise l'instruction : « CREATE DATABASE [nom de la base] ; », mais comme l'outil PHPMyAdmin nous offre une interface intuitive qui va nous permettre de gérer notre base de données (créer des bases, des tables, etc...) sans nous préoccuper du langage SQL qu'il faut écrire, on va donc travailler avec.

Pour accéder à PHPMyAdmin, il faut taper dans le navigateur l'adresse suivante :

http://localhost/phpmyadmin

Si tout s'est bien passé, nous allons arriver devant un écran de ce style :



Figure 25 : Création d'une BD

Chapitre V : Protocole expérimentale

Nous allons maintenant pouvoir créer notre base de données. Notre base de données intitulée **Télémédecine**. On met donc **Télémédecine** dans la case **Créer une base de données** et on clique sur **Créer**.

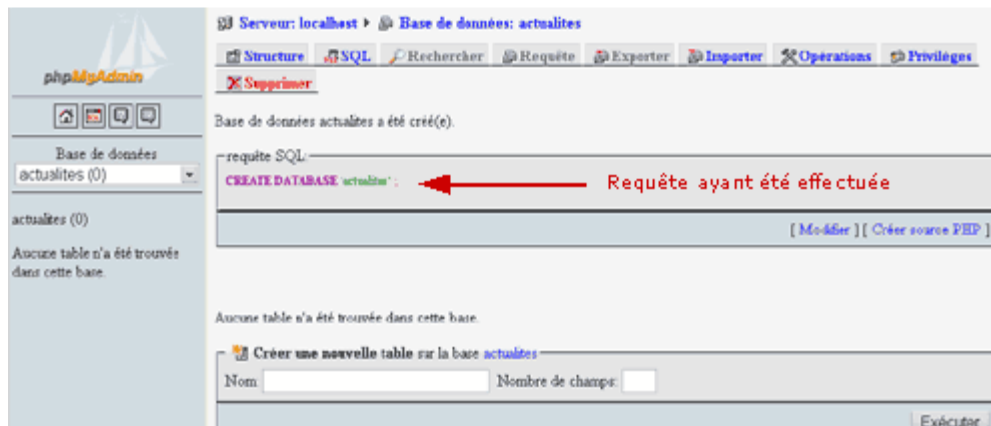


Figure 26 : Après la création de la BD

La base de données a été créée. Comme nous pouvons le constater, le langage SQL que PHPMYAdmin a utilisé pour créer cette base est **CREATE DATABASE Télémédecine**.

4.3.2. La création des tables :

La création d'une table utilise la commande CREATE TABLE selon la syntaxe suivante :

CREATE [TEMPORARY] TABLE nom_table [IF NOT EXISTS] (
nom_attribut TYPE_ATTRIBUT [OPTIONS]

...

)

TEMPORARY donne pour durée de vie à la table: le temps de la connexion de l'utilisateur au serveur, après, elle sera détruite. En l'absence de cette option, la table sera permanente à moins d'être détruite par la commande DROP TABLE. L'option IF NOT EXIST permet de ne créer cette table que si une table de même nom n'existe pas encore.

A l'intérieur des parenthèses, il sera listé tous les attributs, clés et indexes de la table, ainsi que les types des attributs et leurs options.

Les tables peuvent contenir différents types de données:

- Nombre entier signé ou non (température, quantité commandée, âge)
- Nombre à virgule (prix, taille)
- Chaîne de caractères (nom, adresse, article de presse)
- Date et heure (date de naissance, heure de parution)

Chapitre V : Protocole expérimentale

- Énumération (une couleur parmi une liste prédéfinie)
- Ensemble (une ou des monnaies parmi une liste prédéfinie).

Afin d'optimiser la place prise par les données et la rapidité de traitement, il convient de choisir avec efficacité le type de données que contiendront les différents champs de la table.

Ainsi les options vont aussi permettre d'économiser de la place, par exemple pour les champs de type numérique, l'option UNSIGNED va nous permettre en économisant de la place (on ne stocke plus le bit de signe) de stocker des nombres plus grands (positifs). Nous ne pourrions par contre plus stocker de nombres négatifs. L'attribut [M] définit le nombre de caractères maximum que le champ pourra contenir.

Maintenant on va créer une table avec PHPMyAdmin, pour ce faire dans le champ 'NOM' de l'interface PHPMyAdmin (Figure 26) on tape le nom de la table et dans le champ « 'Nombre de champs', en fait entrer le nombre d'attributs (appelées aussi champs).



Figure 27 : Aperçu des champs de la table avec PHPMyAdmin

Nous voici maintenant devant une liste d'attributs (champs) qu'il va falloir déterminer. Par exemple, on tape L'attribut Nom de la table patient dans la case champs, et on choisit le type et les différentes options. On répète ces étapes pour tous les attributs, puis on confirme et on aura une table complète.

a) Structure finale de la base de données :

- **Structure de la table patient :**

```
CREATE TABLE IF NOT EXISTS `patient` (  
  `ID` int(10) unsigned NOT NULL AUTO_INCREMENT,  
  `Nom` varchar(30) NOT NULL,  
  `Prenom` varchar(30) NOT NULL,
```

Chapitre V : Protocole expérimentale

`Sexe` enum('HOMME','FEMME') NOT NULL,

`Age` int(3) NOT NULL,

`Adresse` varchar(250) NOT NULL,

`Telephone` varchar(15) NOT NULL,

`Email` varchar(30) NOT NULL,

`Photo` varchar(30) NOT NULL,

`Profession` tinytext NOT NULL,

`Tabac` enum('OUI','NON') NOT NULL,

`Alcool` enum('OUI','NON') NOT NULL,

`Pathologie` tinytext NOT NULL,

`Rapport_clinique` tinytext NOT NULL,

`Antecedent_familiale` tinytext NOT NULL,

`Examen_CDS` tinytext NOT NULL,

`Date_Admission` tinytext NOT NULL,

PRIMARY KEY (`ID`)

) ENGINE=InnoDB DEFAULT CHARSET=latin1 AUTO_INCREMENT=2527259;

- **Structure de la table user :**

CREATE TABLE IF NOT EXISTS `user` (

`id_user` int(3) NOT NULL AUTO_INCREMENT,

`Nom` varchar(15) NOT NULL,

`Prenom` varchar(15) NOT NULL,

`login_user` varchar(10) NOT NULL DEFAULT "",

`pass_user` varchar(32) NOT NULL DEFAULT "",

PRIMARY KEY (`id_user`)

) ENGINE=MyISAM DEFAULT CHARSET=latin1 AUTO_INCREMENT=26 ;

5. La réalisation d'un serveur de Télémédecine [10]:

5.1. Le serveur web:

Un serveur Web est un ordinateur sur lequel fonctionne un serveur HTTP (*Hypertext Transfer Protocol*), ce dernier est un logiciel capable d'interpréter des requêtes respectant le protocole de communication client-serveur HTTP arrivant sur le port associé au protocole HTTP (par défaut le port 80), et de fournir une réponse avec ce même protocole. Ces requêtes utilisant pratiquement ce même protocole HTTP, demandent un service ou un contenu comme une page Web ou un fichier multimédia stocké dans ce même serveur Web. Pour cela, il pourra être couplé à des bases de données à travers un serveur MySQL pour générer du contenu et à Un interpréteur PHP ou des serveurs applicatifs.

5.2. L'implémentation du serveur Télémédecine :

Pour faire l'implémentation nous avons besoin de plusieurs outils :

Un serveur http ou appelé aussi serveur web, et puisque Le serveur HTTP le plus utilisé est Apache HTTP Server, et que ce logiciel est gratuit, on a choisi de travailler avec.

On a besoin aussi d'un interpréteur PHP, et de l'outil PHPMyAdmin qu'on va utiliser pour la gestion de la base de données et pour remplir le rôle du serveur MySQL.

5.2.1. Interpréteur PHP :

PHP (sigle de *PHP: Hypertext Preprocessor*), est un langage de scripts libre principalement utilisé pour produire des pages Web dynamiques via un serveur HTTP, mais pouvant également fonctionner comme n'importe quel langage interprété de façon locale, en exécutant les programmes en ligne de commande. PHP est un langage impératif disposant depuis la version 5 de fonctionnalités de modèle objet complètes. En raison de la richesse de sa bibliothèque, on désigne parfois PHP comme une plate-forme plus qu'un simple langage.

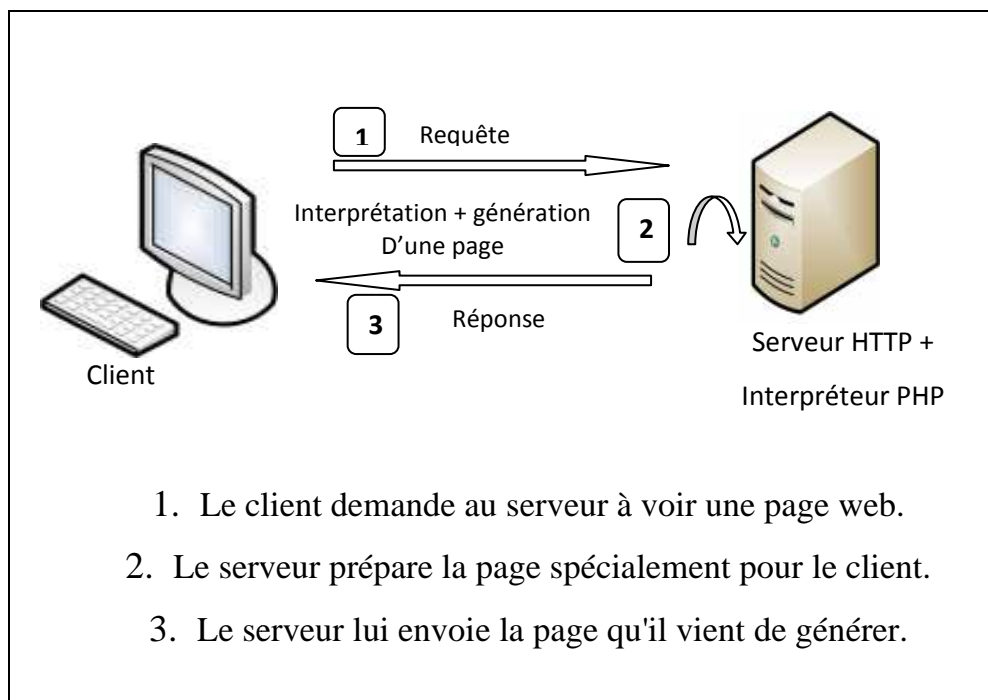


Figure 28: Fonctionnement du PHP

5.2.2. L'outil WAMP :

WAMP signifie Windows + Apache + MySQL + PHP (ou Perl, ou Python). WAMP est un utilitaire qui installe et configure automatiquement un environnement de travail complet sous Windows permettant de mettre en œuvre toute la puissance et la souplesse qu'offrent le langage dynamique PHP et son support efficace des bases de données. WAMP regroupe un serveur Apache, une base de données MySQL, le langage PHP ainsi que des outils facilitant le développement des sites comme PHPMyAdmin .

WAMP est en fait dérivé de LAMP, qui décrivait un ensemble d'outils permettant de bâtir un site Web à l'aide d'outils Open Source : Linux + Apache + MySQL + PHP (ou Perl, ou Python). La terminologie WAMP est donc utilisée pour décrire une des architectures destinée à faire fonctionner des sites web sur une machine Windows.

Donc il est très utile d'installer WAMP qui regroupe tous les outils nécessaire pour l'implémentation de notre serveur, au lieu de les installer un par un.

➤ *Récupération de WAMP :*

Le pack WAMP est disponible sur le site officiel : <http://www.wampserver.com/> ou sur n'importe quel site de téléchargement vu qu'il est gratuit, il suffit donc de télécharger la version la plus récente, actuellement, c'est la version 2i, et de l'installer.

➤ *Installation de WAMP :*

Pour installer WAMP, il suffit de double-cliquer sur le fichier d'installation téléchargé précédemment, une fenêtre s'affichera qui nous demandant de choisir la langue qui sera utilisée par l'assistant de l'installation. Après avoir cliquer sur OK l'assistant d'installation va nous indiquer les étapes à suivre (qui sont très simple) pour installer le logiciel.

➤ *Démarrage d' WAMP :*

Dés qu'on démarre le programme WAMP, une fenêtre s'affichera, nous indiquant que les serveurs Apache et MySQL sont démarrés. Une icône se placera alors dans la barre des tâches à coté de l'horloge, un clic droit sur cette dernière ou gauche sur le petit carré en haut à gauche de la fenêtre affichée précédemment permet d'accéder aux différents menus, les plus importants sont:

- ✓ . Fichier Log : renvoie les erreurs générées par Apache et MySQL.
- ✓ . Configuration : donne accès aux différentes configurations d' WAMP
- ✓ . Web local : ouvre la page "http://localhost/"
- ✓ . Démarrer/Arrêter : démarre/arrête Apache et MySQL
- ✓ . Quitter : ferme WAMP

Pour vérifier le fonctionnement d'WAMP, il suffit de taper dans le navigateur :

- ✓ . <http://localhost>
- ✓ . ou <http://127.0.0.1>

➤ *Utilisation du répertoire « www » ou des alias :*

Pour que les pages PHP soient interprétées, il faut placer le dossier les contenant dans le répertoire www. Le serveur Apache est configuré pour ouvrir automatiquement un fichier index lors de la saisie de l'adresse 'http://localhost/' (à condition évidemment que le serveur Apache soit en exécution). On peut aussi utiliser le menu « Répertoire Alias » pour créer des alias.

6. Aperçu du site web :

Dans les figures qui suivent nous allons présenter quelques captures d'écran pour notre site web.

A. La page d'accueil :



Figure 29 : Aperçu de la page d'accueil

B. La page de la connexion :

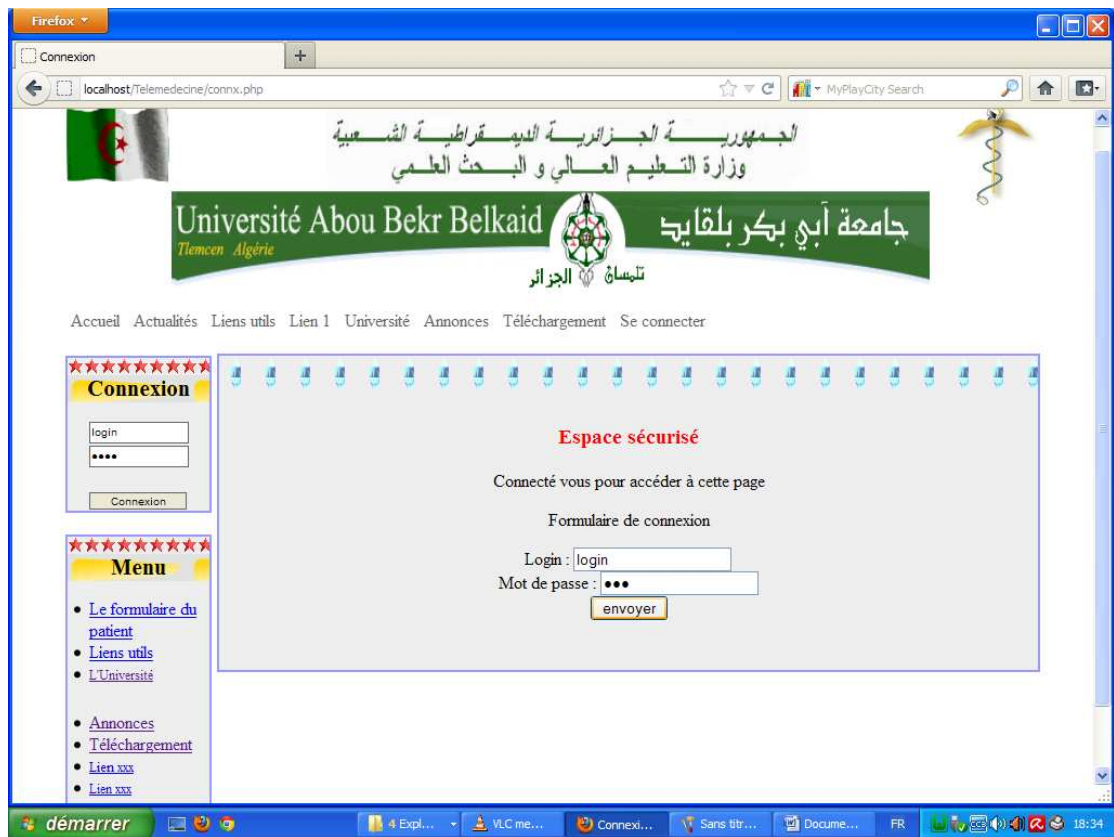


Figure 30 : Aperçu de la page de la connexion (concernant l'administration)

C. La page d'accueil administrative :

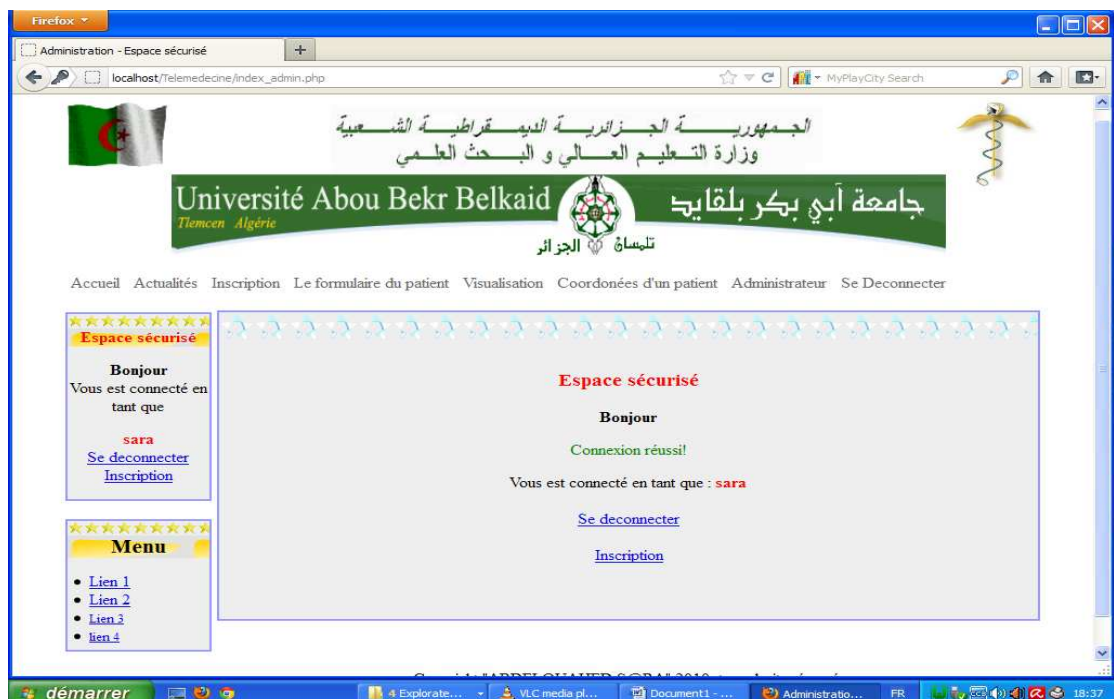


Figure 31 : Aperçu de la page d'accueil administrative (Page sécurisé)

D. La page du formulaire du patient :

The screenshot shows a web browser window with the address bar displaying 'localhost:8080/Telemedecine/formulaire_patient.php'. The page is titled 'Le Formulaire du Patient'.

Left Sidebar:

- Espace sécurisé**
 - Bonjour
 - Vous êtes connecté en tant que **sara**
 - [Se déconnecter](#)
 - [Inscription](#)
- Menu**
 - [Lien 1](#)
 - [Lien 2](#)
 - [Lien 3](#)
 - [Lien 4](#)

Main Form Fields:

- N° de patient (ID):
- NOM:
- PRENOM:
- SEXE:
- AGE:
- ADRESSE:
- TELEPHONE:
- COURRIER ELECTRONIQUE:
- Charger L'images du patient
 - Choisir une image:

Bottom Section: Charger L'images du patient

- Choisir une image:
- PROFESSION:
- Est ce que le patient est un fumeur ?
 - A. ☐ OUI
 - B. ☐ NON
- ALCOOL ?
 - A. ☐ OUI
 - B. ☐ NON
- PATHOLOGIE:
- RAPPORT CLINIQUE:
- ANTÉCÉDENT FAMILIAL:
- EXAMEN COMPLÉMENTAIRE DES SIGNAUX:
 -
- DATE D'ADMISSION:
(sous forme jour/mois/Année)
-

Figure 32 : Aperçu de la page formulaire du patient

E. La page de visualisation :

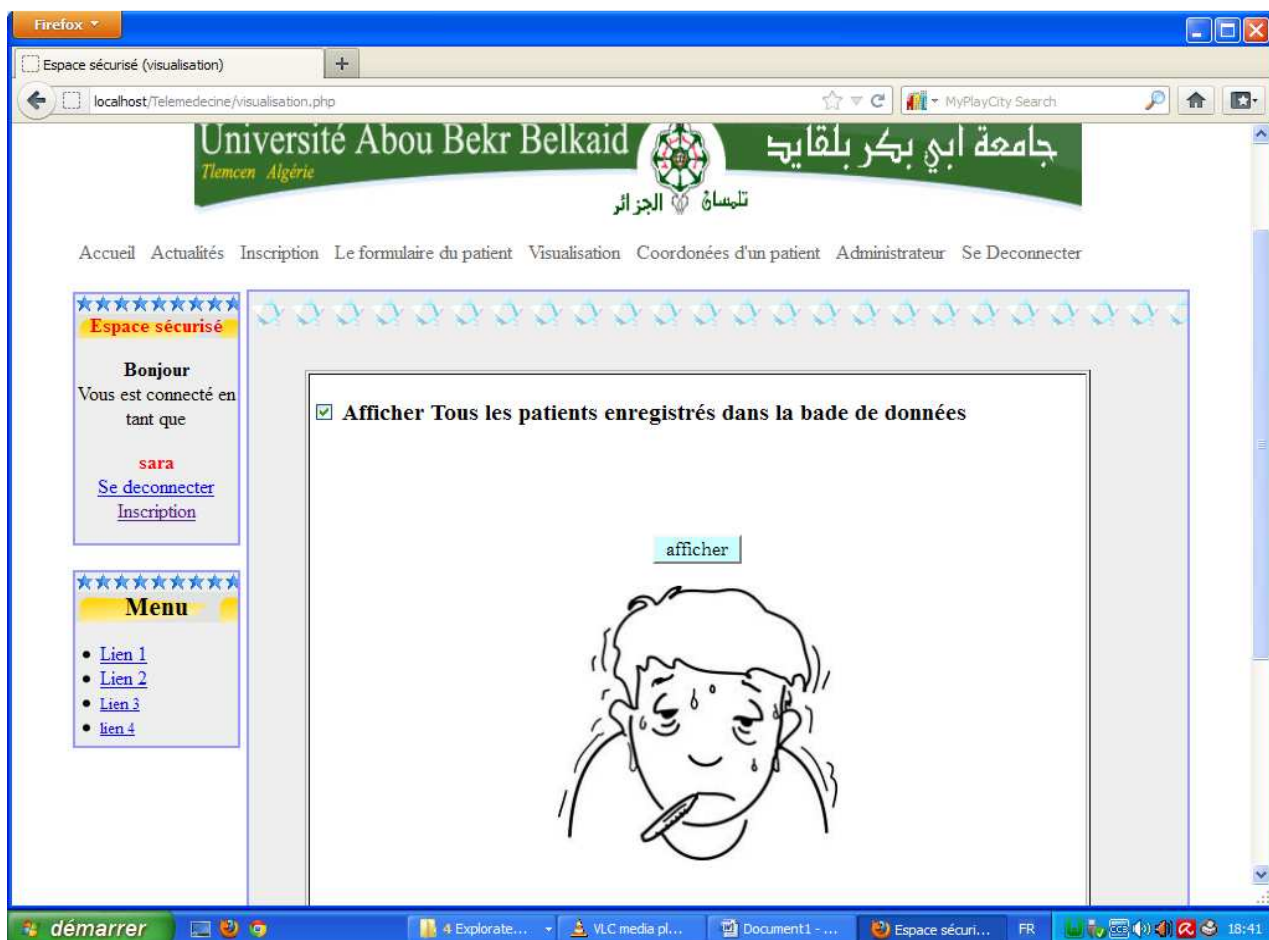


Figure 33 : Aperçu de la page visualisation

F. La page coordonné d'un patient :

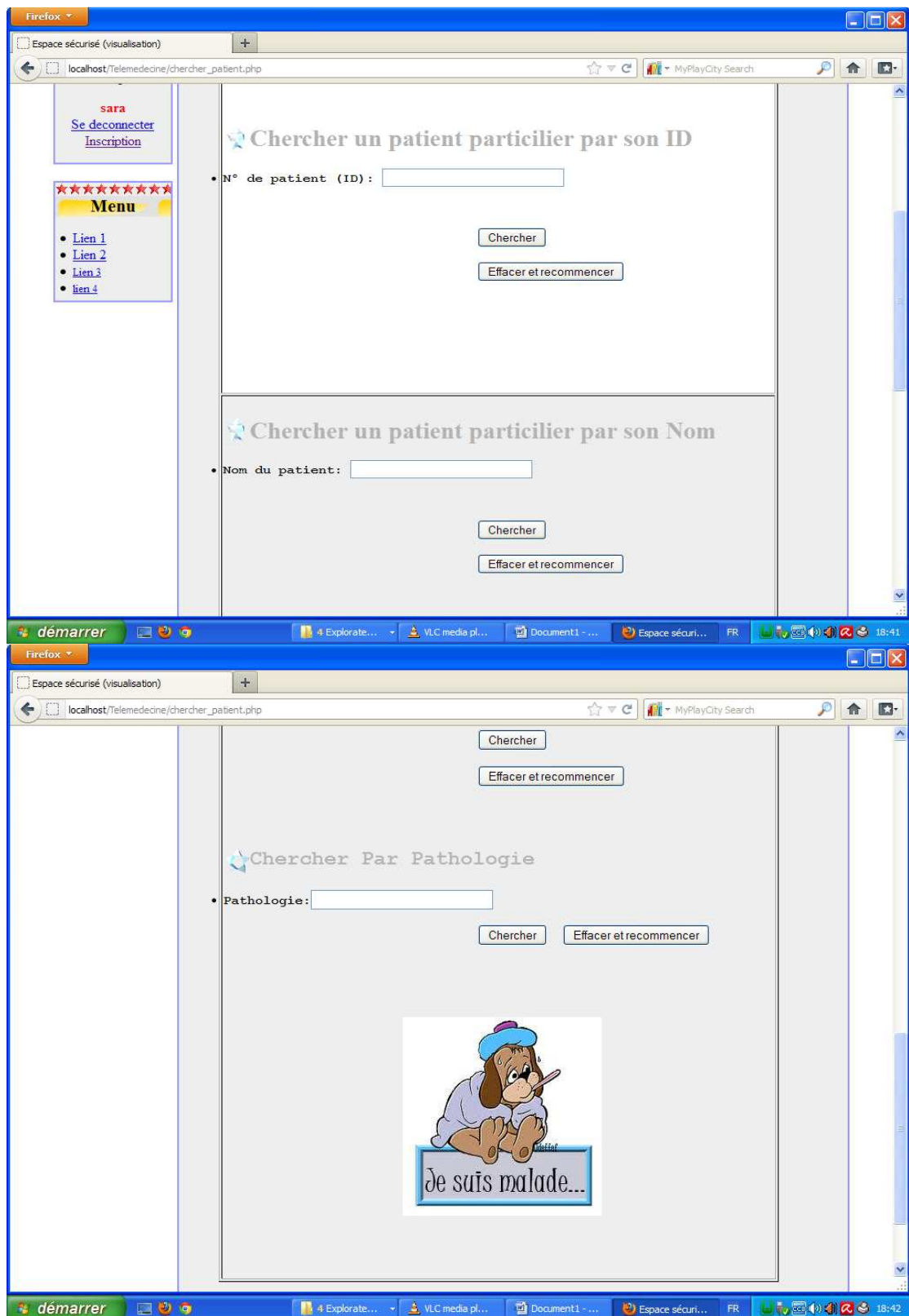


Figure 34 : aperçu de la page coordonné d'un patient

G. La page des administrateurs :

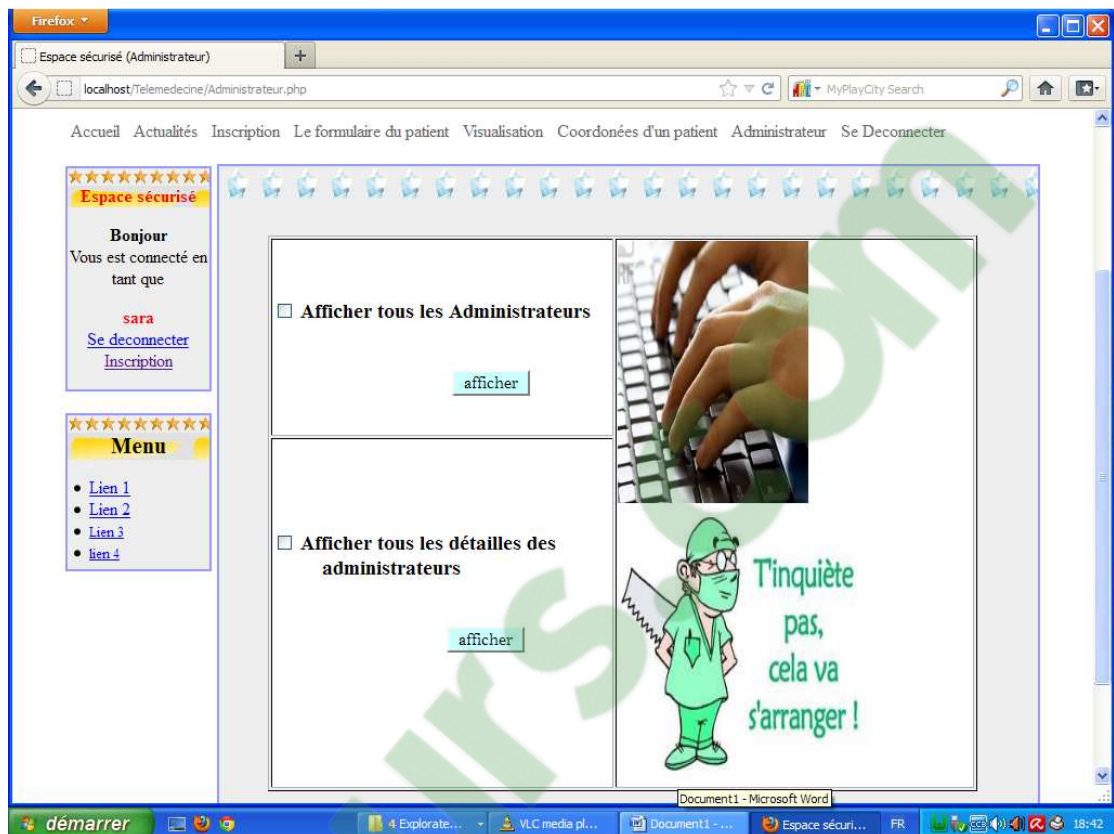


Figure 35 : Aperçu de la page administrateur

H. La page des patient rechercher avec (ID : 1) :



Figure 36 : Aperçu des patients recherchés

Chapitre V : Protocole expérimentale

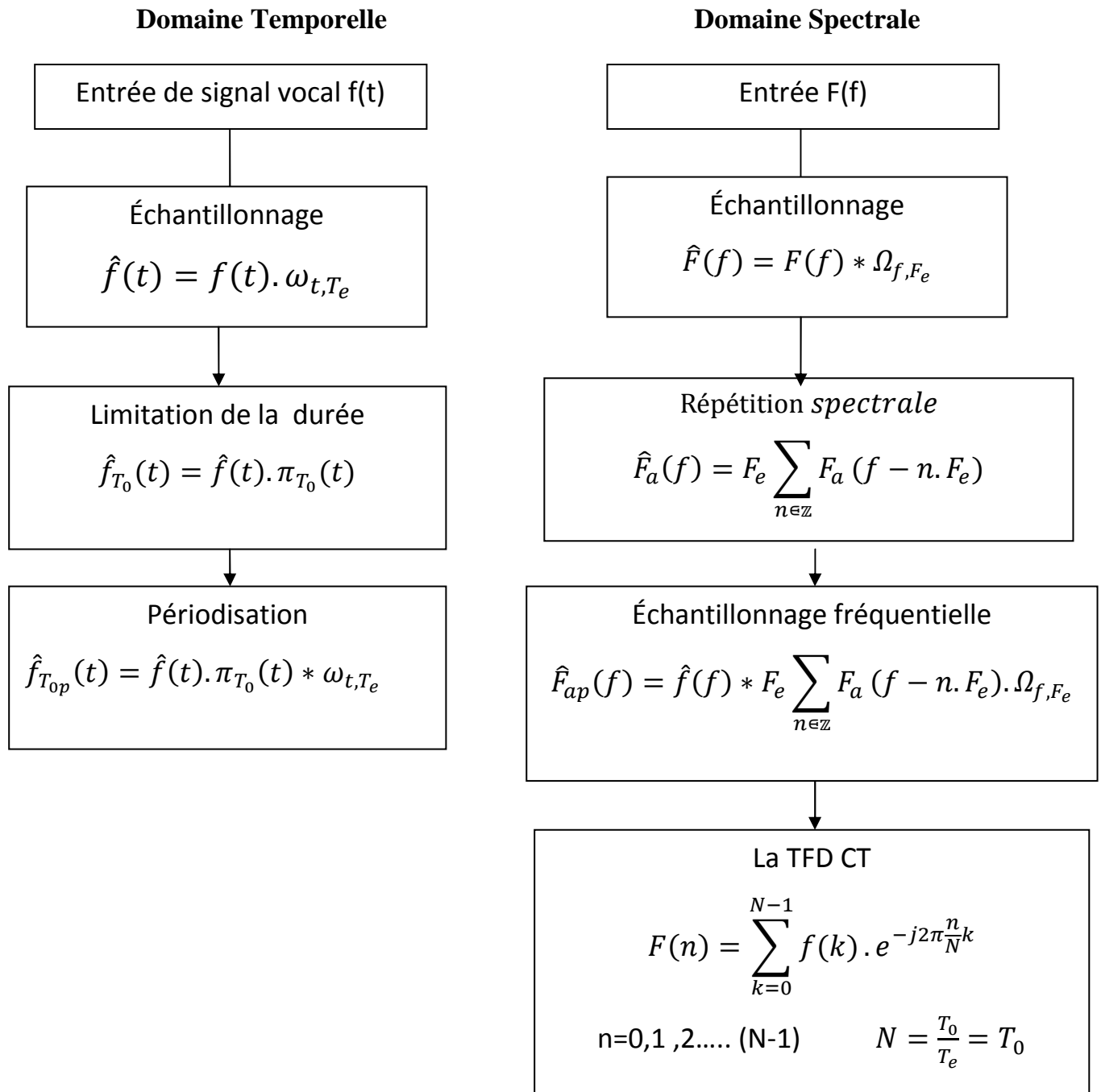
I. La page de tous les patients enregistrés dans la base de données :

Photo	ID	Nom	Prénom	Sexe	Age	Adresse	Telephone	E-mail	Profession	Tabac	Alcool	Pathologie	Rapport clinique
	1	Bengrine	Ahmed	HOMME	45	remchi	45455336	ahmoim	sans profession	OUI	OUI	tabagique chronique(30 paquets mois-ethyisme chronique-pneumothorax en 2008-bronchopneumopathie chronique obstructive	dysphonie chronique évoluant depuis 06 mois sans signes associés Examen clinique dont la laryngoscopie indirecte révèle un épaississement de la corde vocale gauche étendu au vestibule dans son 1/3 moyen.
	2	GHORZI	LARBI	HOMME	65	BENI-SAF	0770672557	CCVVVV	pecheur retraité	OUI	NON	tabagique chronique(30 paquets mois-ethyisme chronique-pneumothorax en 2008-bronchopneumopathie chronique obstructive(40 paquets mois)	dysphonie chronique évoluant depuis 04 mois associée à une dyspnée progressive ayant nécessité sa trachéotomie à l'admission. L'examen clinique retrouve une tuméfaction
	3	BOUFRAINE	AEK	HOMME	61	remchi	0771747653	azeedif	employé dans une usine d'argile retraité	OUI	NON	tabagique chronique-diabète insulino dépendant-hypertension artérielle.	dysphonie chronique évoluant depuis 09 mois associée à une dyspnée progressive et toux grasse ayant nécessité une trachéotomie à l'admission. L'examen clinique dont la laryngoscopie

Figure 37 : Aperçu des patients enregistrés

7. Les algorithmes de l'interface :

7.1. L'algorithme de la TFD CT:



7.1.1. Explication :

- L'échantillonnage :

L'échantillonnage dans le domaine temporel consiste à multiplier le signal par un peigne de Dirac

$$\hat{f}(t) = f(t) \cdot \omega_{t, T_e} \quad (V.1)$$

$$\hat{f}(t) = f(t) \cdot \sum_{k \in \mathbb{Z}} \delta(t - kT_e) \quad (V.2)$$

$$\hat{f}(t) = \sum_{k \in \mathbb{Z}} f(k \cdot T_e) \cdot \delta(t - kT_e) \quad (V.3)$$

la multiplication temporelle est une convolution fréquentielle:

$$\hat{F}(f) = F(f) * \Omega_{f, F_e} \quad (V.4)$$

- Limitation de la durée du signal :

la limitation de la durée de cette suite, généralement imposée par la mémoire des calculateurs, s'obtient en multipliant la suite $f(t)$ par une porte rectangulaire $\pi_{T_0}(t)$, de hauteur unité et de durée T_0 , suffisamment grande pour que son spectre puisse être approximativement assimilé à une impulsion de Dirac. Cette seconde opération nous fournit une suite tronquée contenant N échantillons :

$$\hat{f}_{T_0}(t) = \hat{f}(t) \cdot \pi_{T_0}(t) = \sum_{k=0}^{N-1} f(k \cdot T_e) \cdot \delta(t - kT_e) \quad (V.5)$$

Sur le plan spectral on approxime la TF de $\pi_{T_0}(t)$ par une Dirac δ_f ce qui donne :

$$\begin{aligned} \hat{F}_a(f) &= \delta_f * \mathcal{F}\{\hat{f}_{T_0}(t)\} \\ &= \mathcal{F}\{\hat{f}_{T_0}(t)\} \\ &= \hat{F}_a(f) \\ \hat{F}_a(f) &= F_e \sum_{n \in \mathbb{Z}} F_a(f - n \cdot F_e) \end{aligned} \quad (V.6)$$

- La périodisation du signal :

La périodisation de la suite $\hat{f}_{T_0}(t) = \sum_{k=0}^{N-1} f(k \cdot T_e) \delta(t - kT_e)$ s'obtient en la convoluant avec un peigne de Dirac de période T_0 ce qui donne :

$$\hat{f}_{T_{0p}}(t) = [\sum_{k=0}^{N-1} f(k \cdot T_e) \cdot \delta(t - kT_e)] * [\sum_{l=-\infty}^{+\infty} T_0 \cdot \delta(t - lT_0)] \quad (V.7)$$

$$D'où \quad \hat{F}_a(f) = F_e \sum_{n \in \mathbb{Z}} F_a(f - n \cdot F_e) \cdot \sum_{l=-\infty}^{+\infty} F_0 \cdot \delta(f - lF_0)$$

Le nombre d'échantillons contenu dans la suite temporelle est :

$$N = T_0 / T_e.$$

Le nombre d'échantillon contenu dans la suite spectrale est :

$$N = \frac{1}{T_e} / \frac{1}{T_0} = \frac{1}{T_e} T_0 = T_0 / T_e$$

$$N = T_0 / T_e.$$

Le nombre d'échantillons contenu dans chacune des périodes temporelle est égal au nombre d'échantillons contenu dans chacune des périodes spectrales.

7.1.2. Relation de définition de TFD CT :

La transformée de Fourier discrète de la suite d'échantillons infinie est :

$$\begin{aligned} \mathcal{F}\{\hat{f}_{T_0}(t)\} &= \mathcal{F}\{\sum_{k \in \mathbb{Z}} f(k \cdot T_e) \cdot \delta(t - kT_e)\} \\ &= \sum_{k \in \mathbb{Z}} f(k \cdot T_e) \cdot e^{-j2\pi f k T_e} \end{aligned} \quad (\text{V.8})$$

La transformée d'une durée T_0 de la même suite est définie par la même relation mais la somme ne s'étend plus qu'à N échantillons :

$$\mathcal{F}\{\hat{f}_{T_0}(t)\} = \sum_{k=0}^{N-1} f(k) \cdot e^{-j2\pi f k T_e} \quad (\text{V.9})$$

où le spectre obtenu après périodisation est discret, ses échantillons étant séparés par l'intervalle $f_0 = \frac{1}{T_0}$, la fréquence ne prend que les valeurs $f = f_0 \cdot n = \frac{n}{T_0}$ avec $n=0,1,2,\dots,(N-1)$. il en résulte :

$$\mathcal{F}\{\hat{f}_{T_0}(t)\} = F\left(\frac{n}{T_0}\right) = \sum_{k=0}^{N-1} f(k) \cdot e^{-j2\pi f k T_e} \quad n=0, 1, 2, \dots (N-1).$$

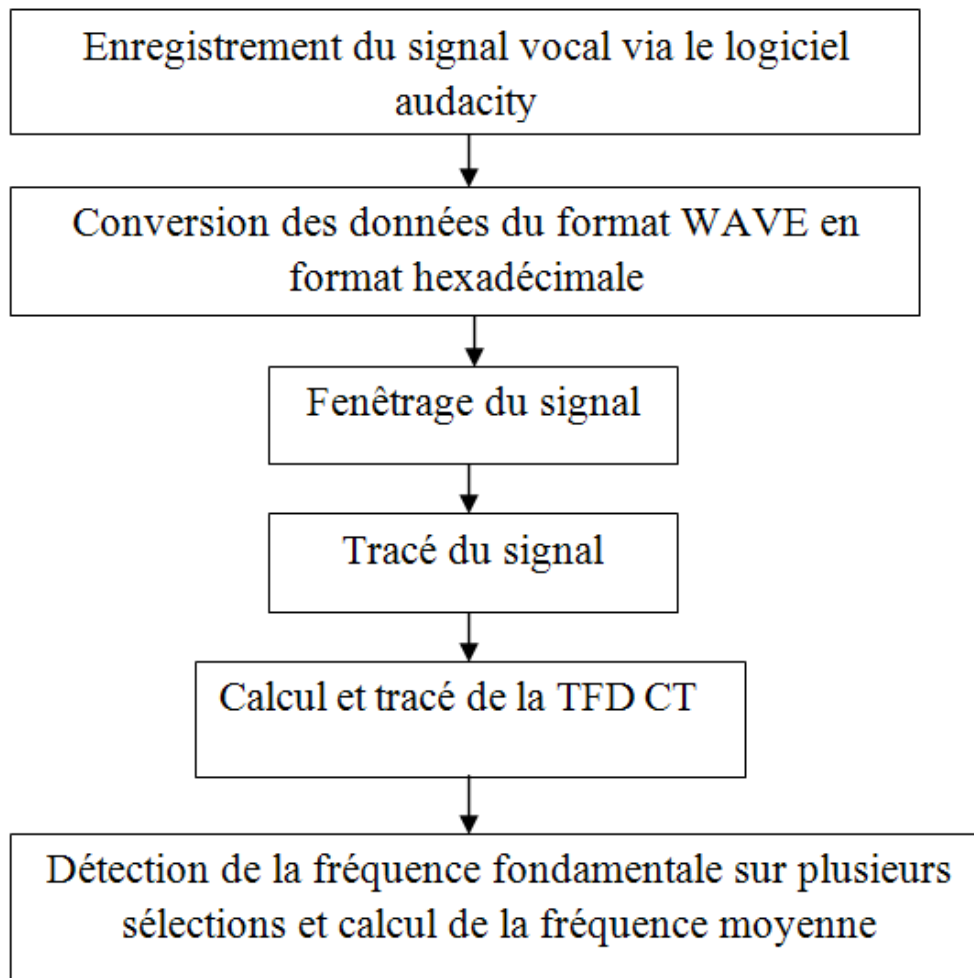
$$F\left(\frac{n}{T_0}\right) = \sum_{k=0}^{N-1} f(k) \cdot e^{-j2\pi n k / N} \quad \text{V.10}$$

On pose généralement pour simplifier les calculs $T_e=1$, d'où la forme usuelle de la transformée de Fourier discrète :

$$F(n) = \sum_{k=0}^{N-1} f(k) \cdot e^{-j2\pi \frac{n}{N} k} \quad (\text{V.11})$$

$$N = T_0 / T_e = T_0$$

7.2. l'algorithme de l'interface :



8. Conclusion :

Dans ce chapitre nous avons exposé la conception et l'implémentation de notre base de données à travers un site web et nous avons expliqué le principe de la TFD CT et l'algorithme de notre interface dans le chapitre suivant nous allons présenter les résultats trouvés.

Chapitre VI :

Evaluation clinique des résultats

Chapitre VI : Evaluation clinique des résultats

1. Introduction :

Dans ce chapitre nous nous sommes beaucoup plus intéressés à l'évaluation clinique de nos résultats de calculs de la fréquence fondamentale pour des sujets sains et malades avec une discussion et interprétation des résultats.

2. Présentation de l'interface graphique :

La figure 38 représente l'interface graphique implémentée sous environnement Visual Basic composée d'une barre de menus (option, tracé du signal, TFD) et d'une commande faisant appel à la base de données.

Le menu option offre la possibilité de l'acquisition du signal sous format WAVE et la conversion WAVE-HEXADECIMALE.

Le menu TFD offre la possibilité de tracé la TFD CT, le zoom de la TFD CT et le tracé de la moyenne des fréquences fondamentales.



Figure 38 : Capture d'écran de l'interface.

La figure 39 représente une capture écran de l'interface graphique du logiciel d'acquisition du signal vocal Audacity appelé à partir de notre interface.

Chapitre VI : Evaluation clinique des résultats

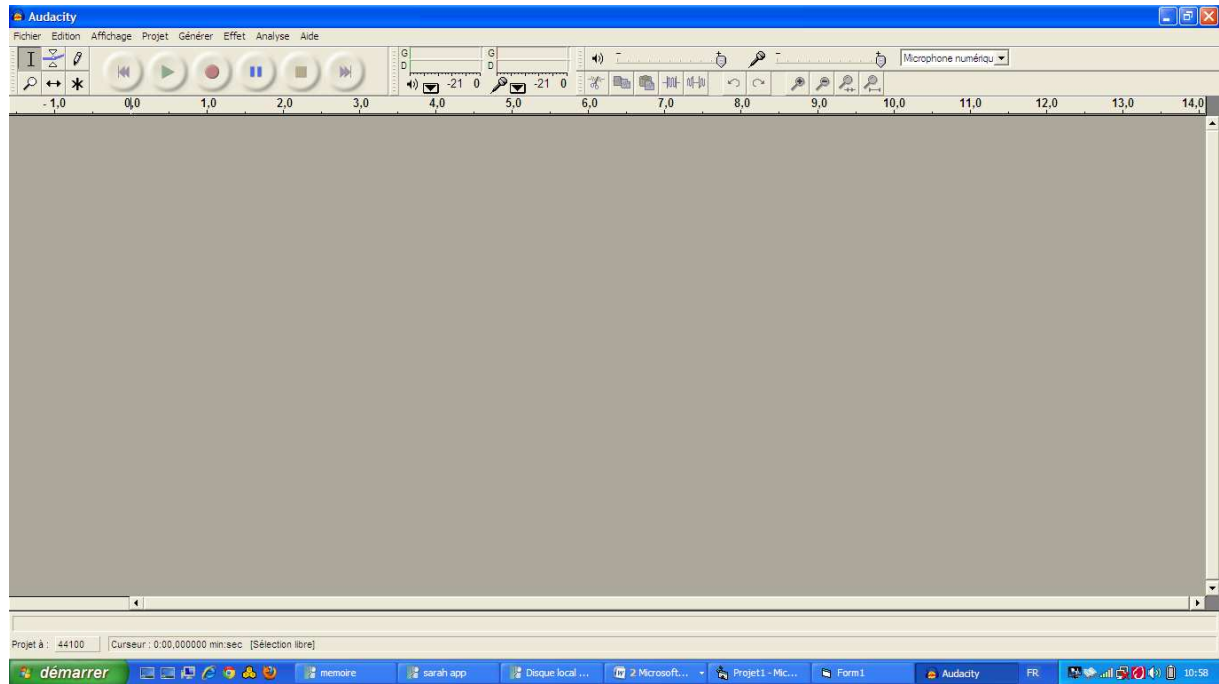


Figure 39 : Capture d'écran de logiciel d'enregistrement.

La figure 40 représente une capture écran du logiciel Watex de conversion de signal vocal du format WAVE au format HEXADECIMALE.

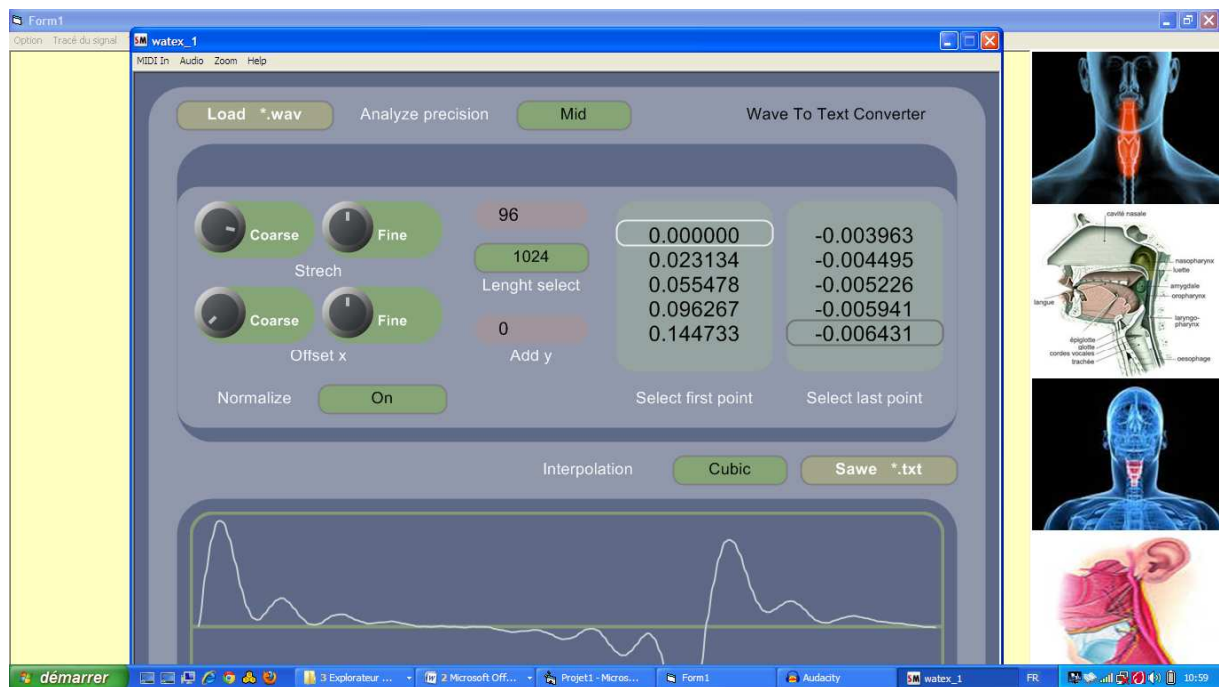


Figure 40 : Capture d'écran de logiciel de conversion.

La figure 41 représente l'appel de notre site internet à partir de l'adresse local localhost.

Chapitre VI : Evaluation clinique des résultats

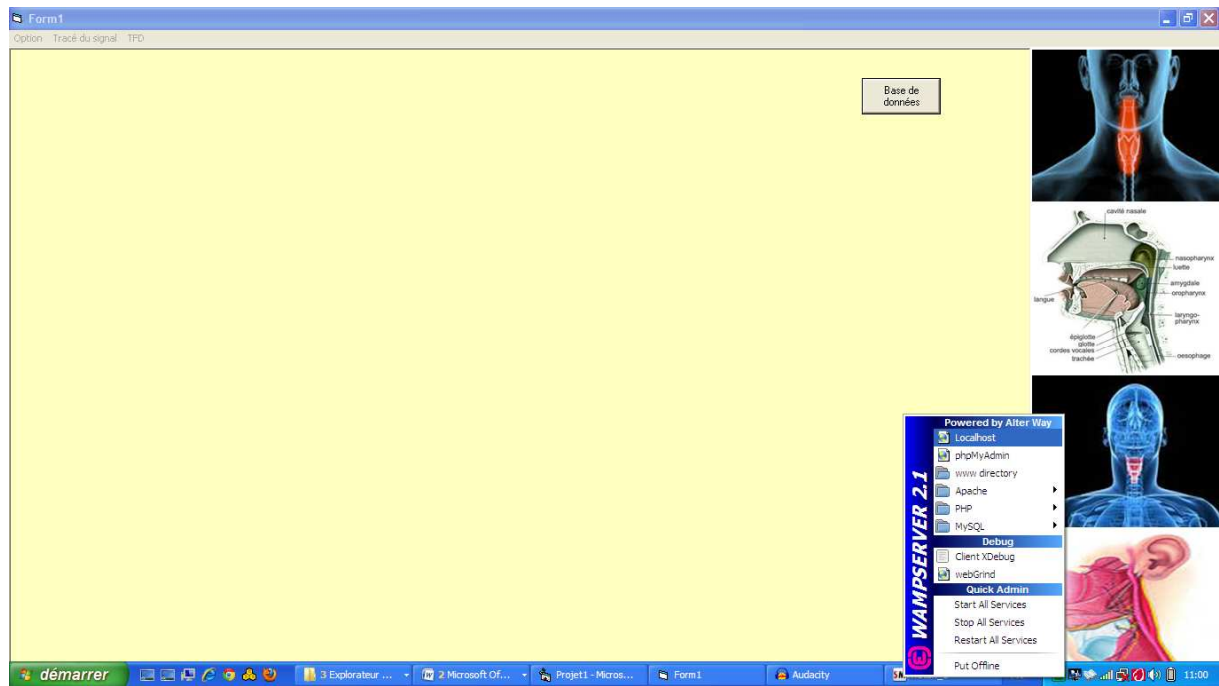


Figure 41 : Capture d'écran de l'appelle à de la base de données.

La figure 42 représente l'accueil de notre serveur web WampServer

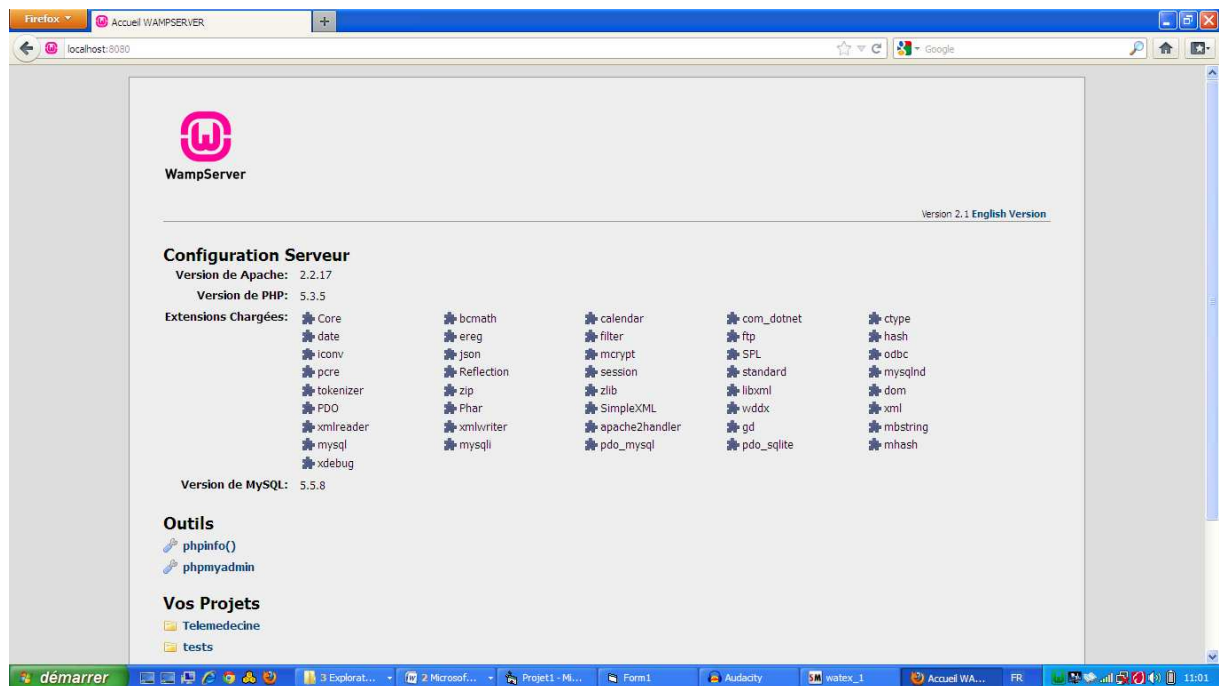


Figure 42 : Capture d'écran d'accueil de WampServer.

La figure 43 représente la page d'accueil de notre site web qui donne accès à la base de données.

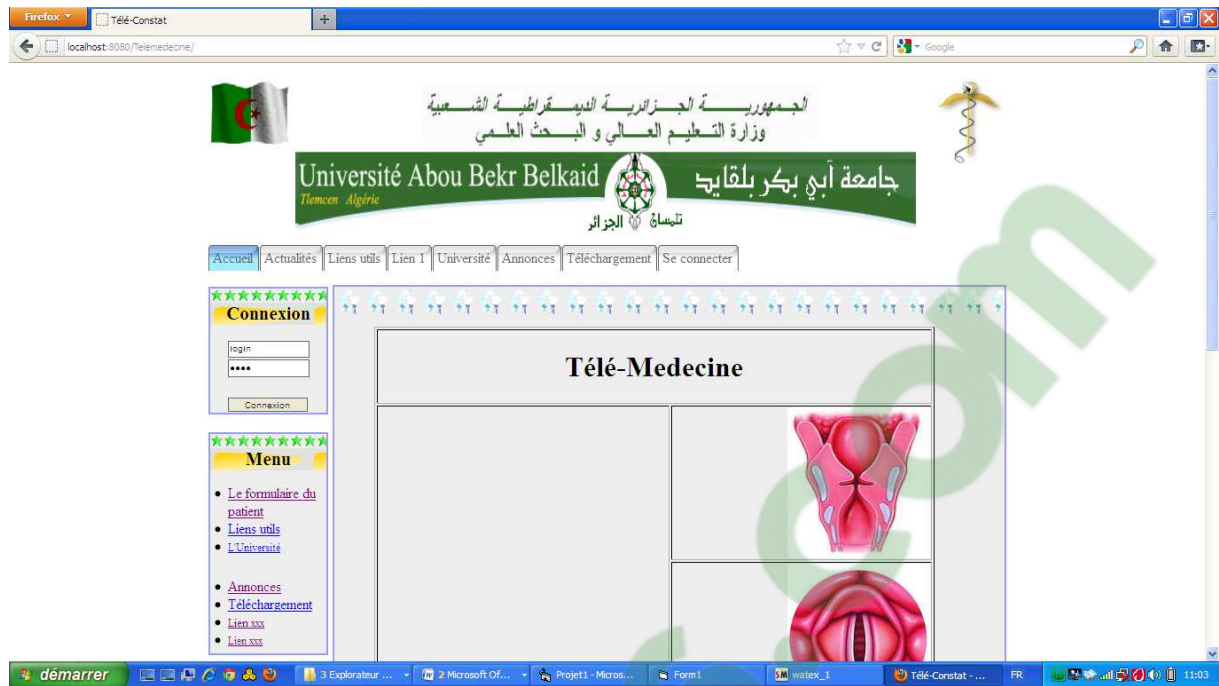


Figure 43 : Capture d'écran de la page d'accueil de notre site internet.

3. Validation clinique des résultats :

3.1. sujets sains :

3.1.1. le premier sujet (masculin) :

➤ Tracé temporel du signal vocal :

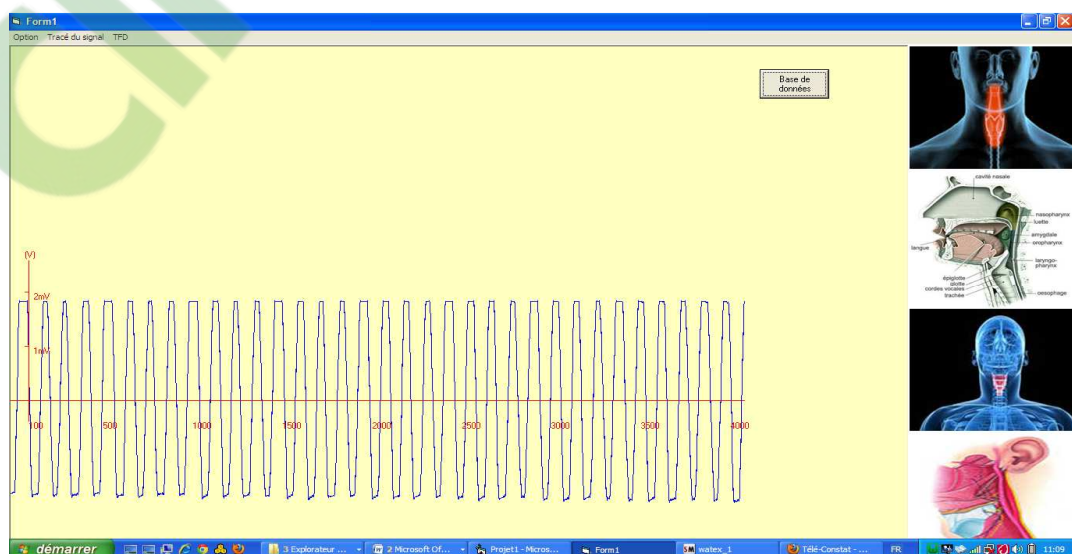


Figure 44 : Capture d'écran de tracé temporel du signal vocal du 1er sujet sain.

Chapitre VI : Evaluation clinique des résultats

➤ Tracé de la TFD CT :

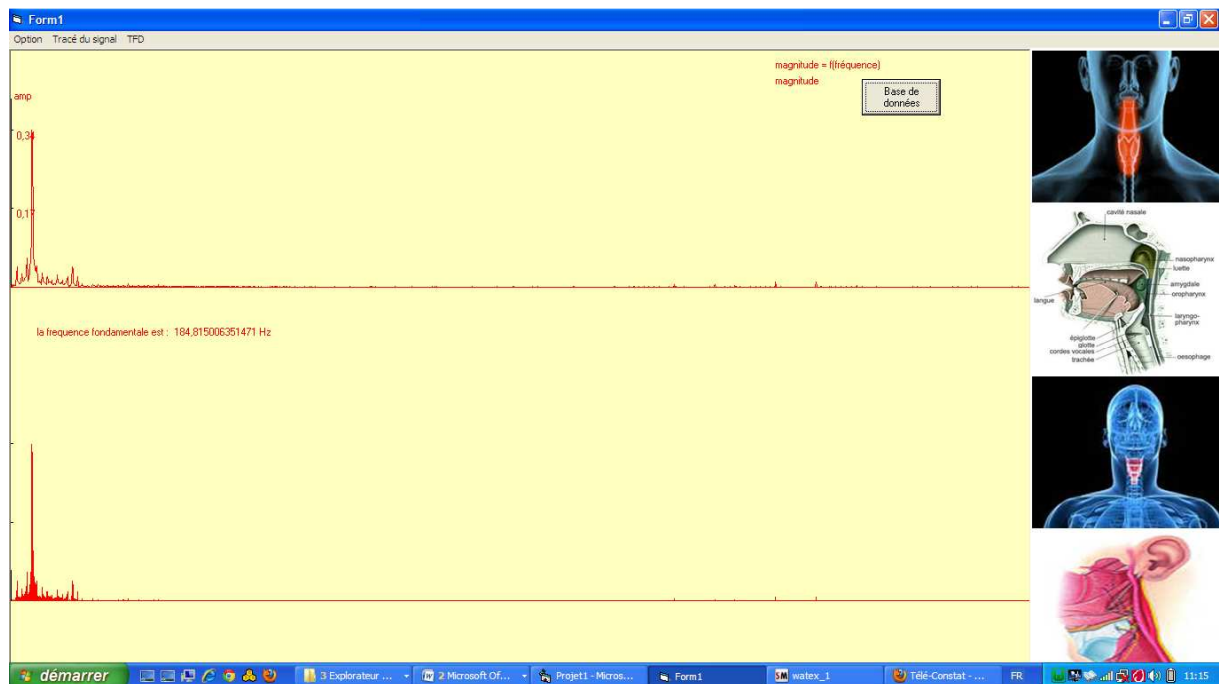


Figure 45 : Capture d'écran de la TFD CT du signal vocal du 1^{er} sujet sain.

➤ Tracé de zoom de la TFD :

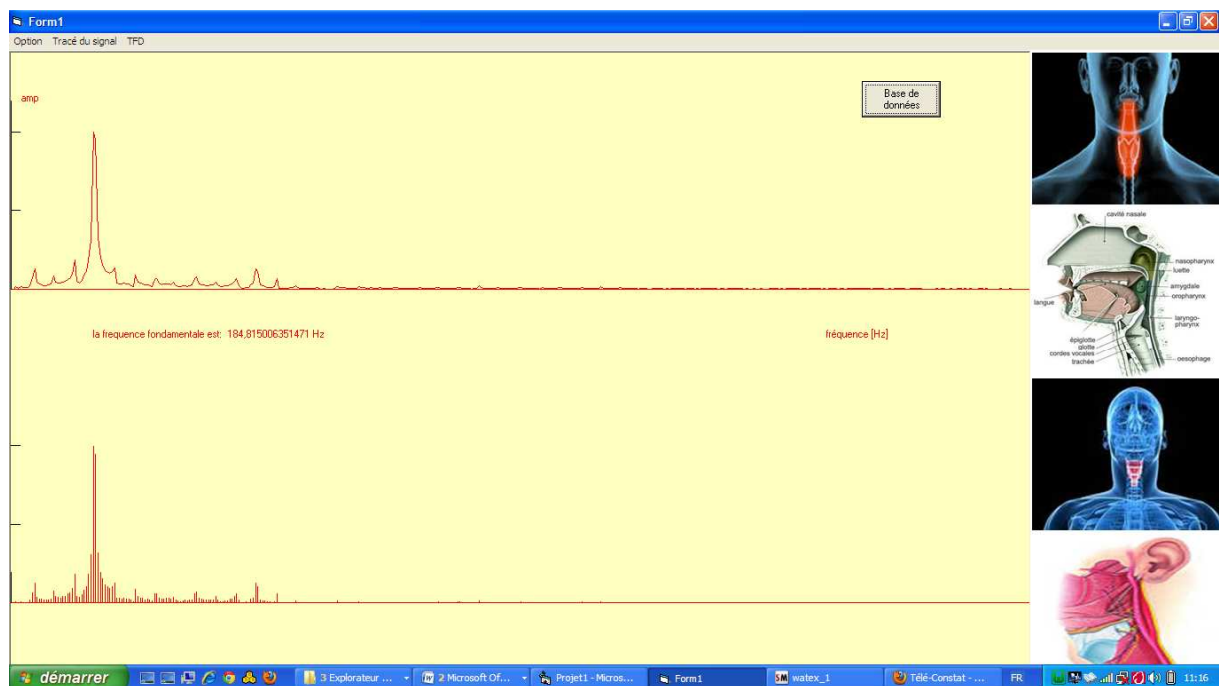


Figure 46 : Capture d'écran de zoom de la TFD CT du signal vocal du 1^{er} sujet sain

Les fréquences fondamentales des 6 sélections :

La première sélection : 184.81HZ

La 2^{ème} sélection : 211.21 HZ

Chapitre VI : Evaluation clinique des résultats

La 3^{ème} sélection : 179.53 HZ

La 4^{ème} sélection : 184.81 HZ

La 5^{ème} sélection : 211.21 HZ

La 6^{ème} sélection : 184.3 HZ

➤ Tracé de la moyenne des fréquences fondamentale :

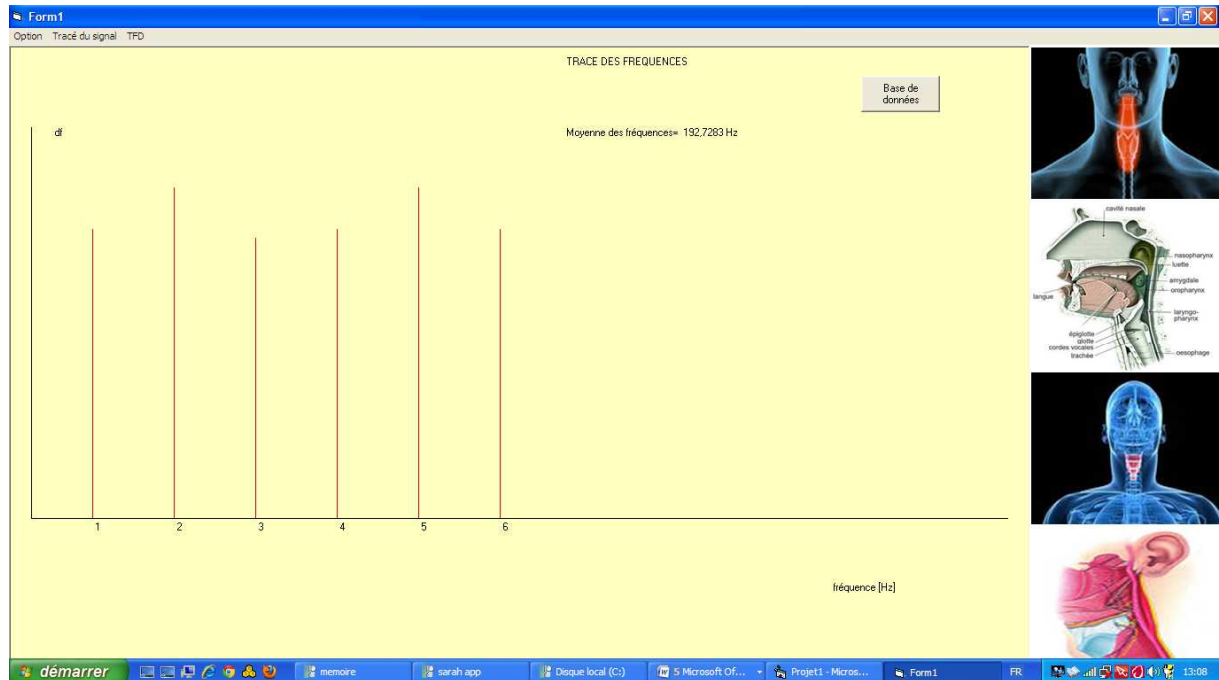


Figure 47 : Capture d'écran de la moyenne des fréquences fondamentales du signal vocal du 1^{er} sujet sain.

La fréquence fondamentale moyenne est : 192.72 HZ

3.1.2. le deuxième sujet (femme) :

➤ Tracé temporel du signal vocal :

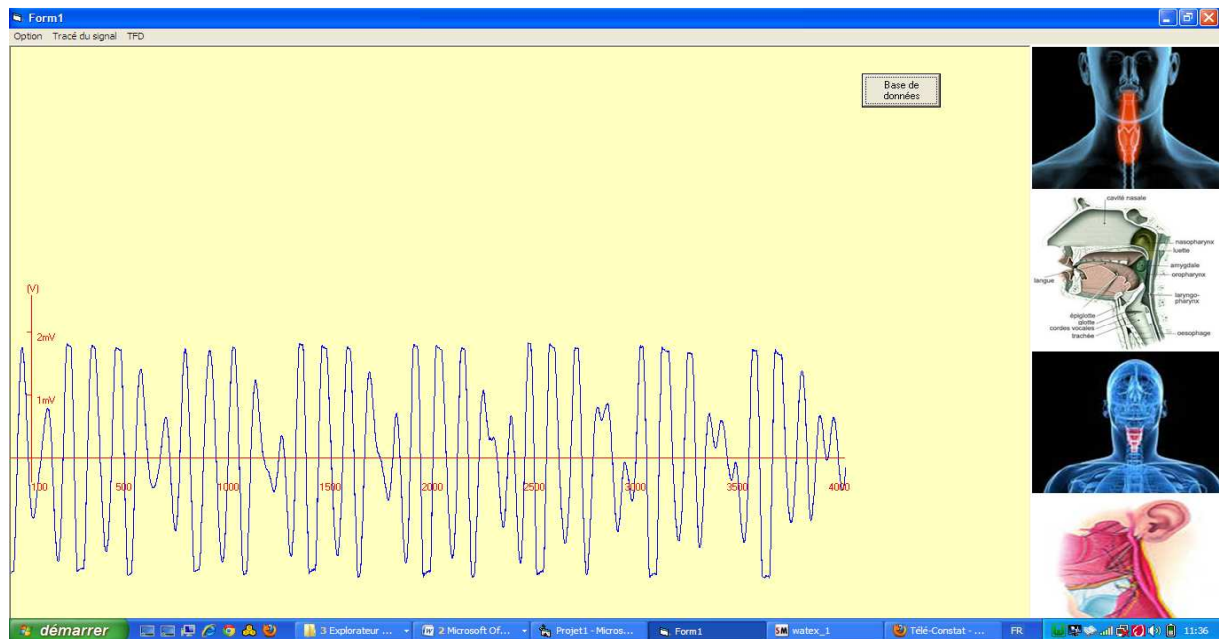


Figure 48 : Capture d'écran de tracé temporel du signal vocal du 2eme sujet sain.

➤ Tracé da la TFD CT :



Figure 49 : Capture d'écran de la TFD CT du signal vocal du 2^{eme} sujet sain

➤ Tracé de zoom de la TFD :

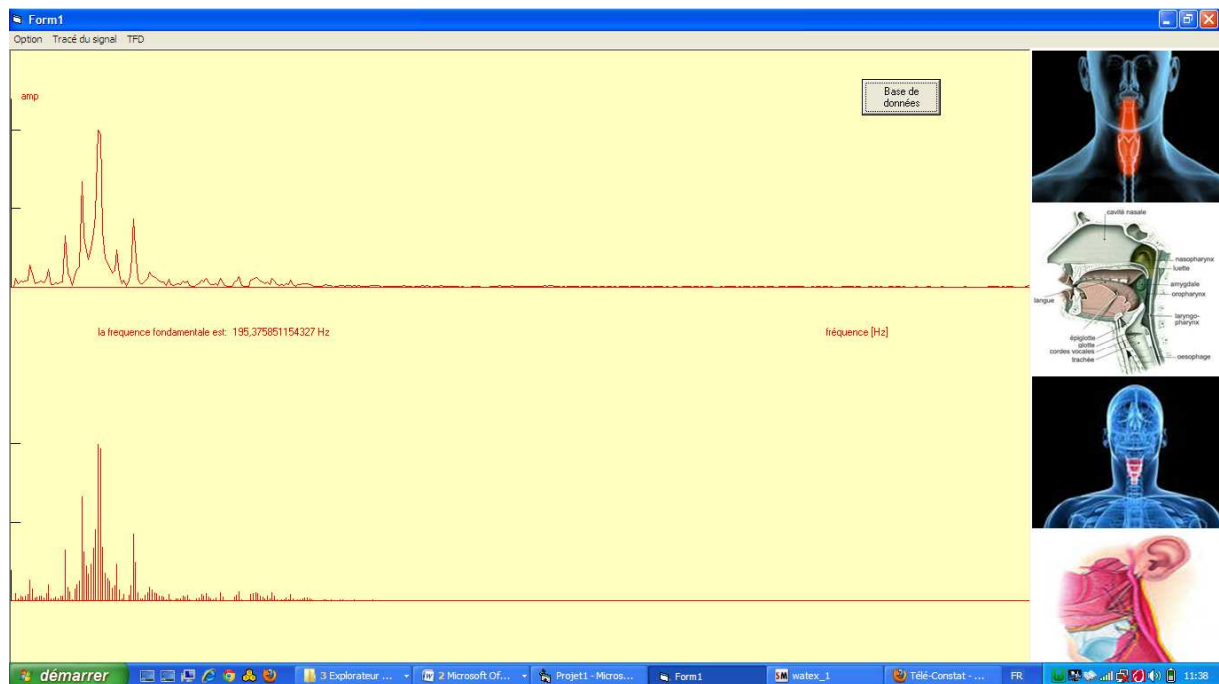


Figure 50 : Capture d'écran de zoom de la TFD CT du signal vocal du 2^{ème} sujet sain.

Les fréquences fondamentales des 6 sélections :

La 1^{ère} sélection : 200.65 HZ

La 2^{ème} sélection : 211.21 HZ

La 3^{ème} sélection : 184.8 HZ

La 4^{ème} sélection : 184.81 HZ

La 5^{ème} sélection : 184.81 HZ

La 6^{ème} sélection : 195.37 HZ

➤ Tracé de la moyenne des fréquences fondamentales :

Chapitre VI : Evaluation clinique des résultats

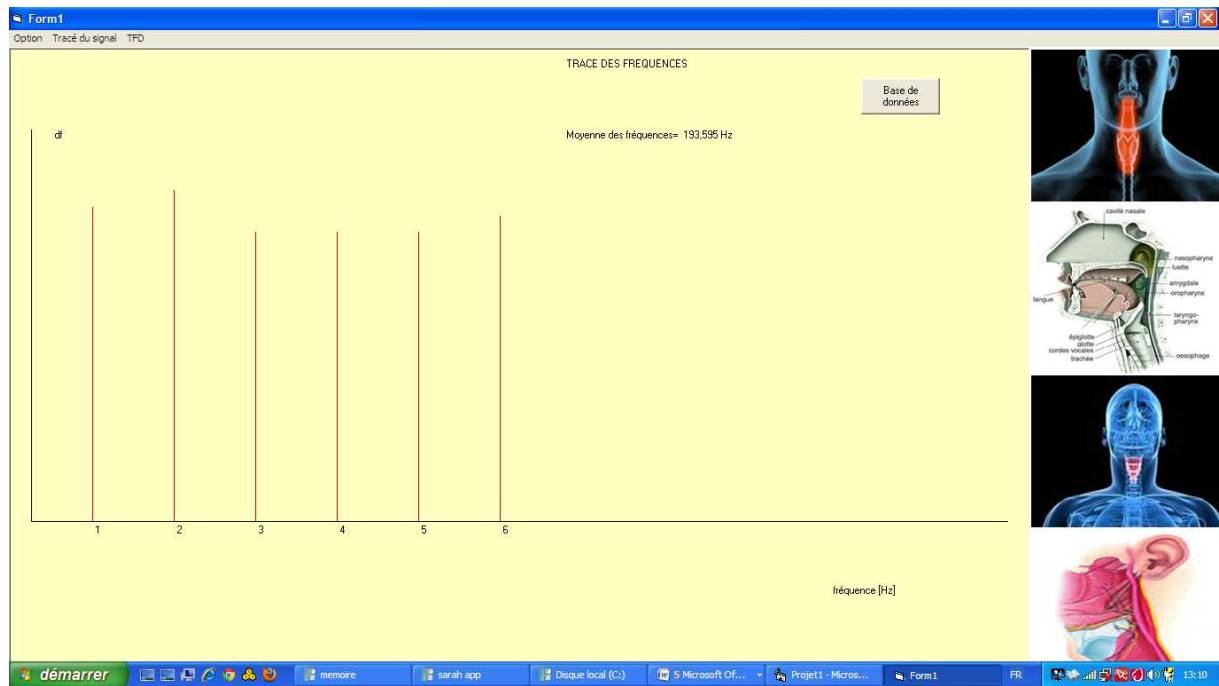


Figure 51 : Capture d'écran de la moyenne des fréquences fondamentales du signal vocal du 2^{ème} sujet sain.

La moyenne des fréquences fondamentales est : 193.605 HZ

3.1.3. Le 3^{ème} sujet (femme) :

➤ Tracé temporel du signal vocal :

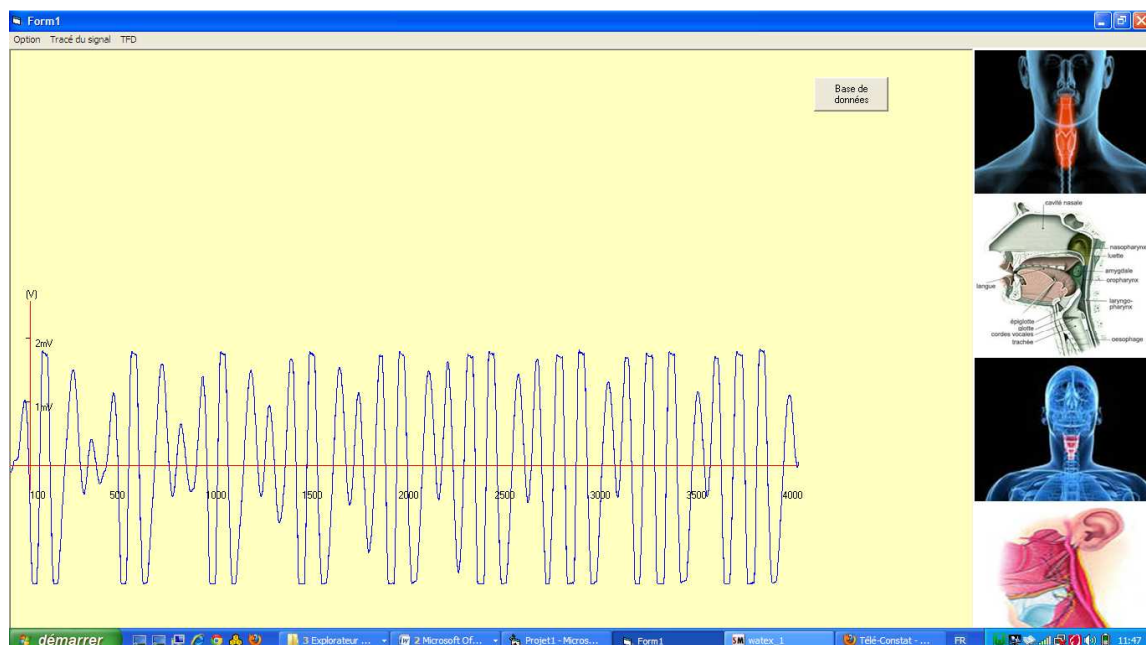


Figure 52 : Capture d'écran de tracé temporel du signal vocal du 3^{ème} sujet sain.

Chapitre VI : Evaluation clinique des résultats

➤ Tracé de la TFD CT :



Figure 53 : Capture d'écran de la TFD CT du signal vocal du 3^{ème} sujet sain.

➤ Tracé de zoom de la TFD :

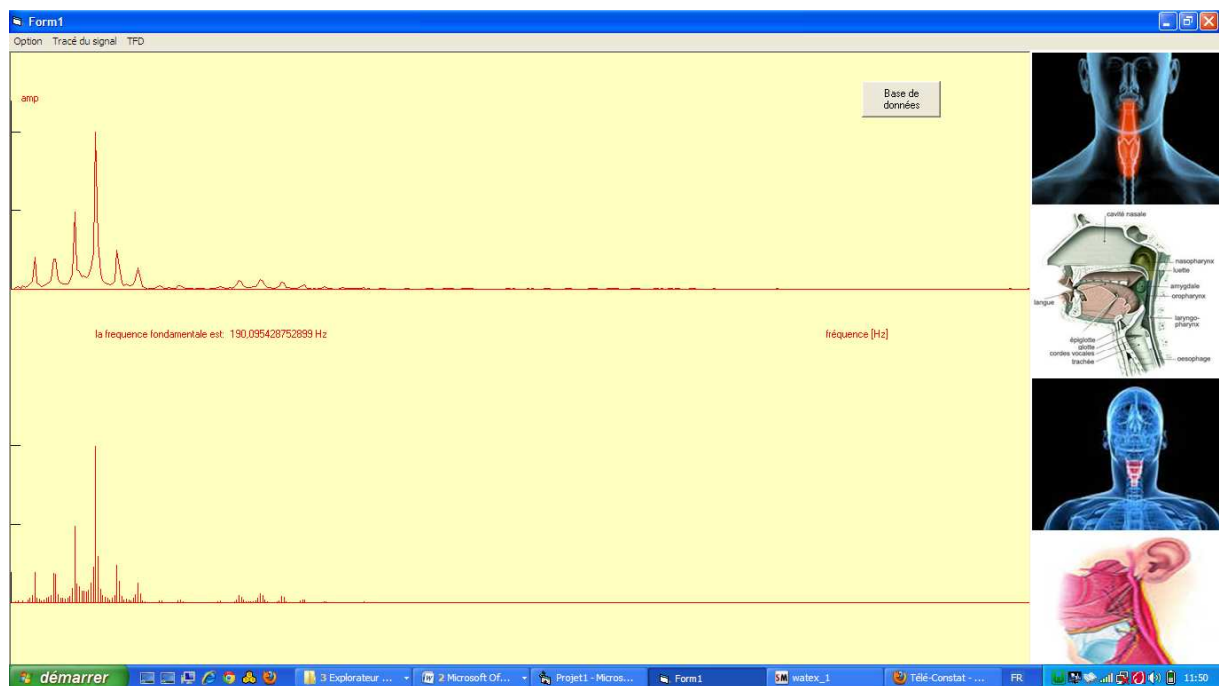


Figure 54 : Capture d'écran du zoom de la TFD CT du signal vocal du 3^{ème} sujet sain.

Les fréquences fondamentales des 6 sélections :

La 1^{ère} sélection : 190.09 HZ

La 2^{ème} sélection : 184.8 HZ

Chapitre VI : Evaluation clinique des résultats

La 3^{ème} sélection : 184.8 HZ

La 4^{ème} sélection : 216.49 HZ

La 5^{ème} sélection : 184.81 HZ

La 6^{ème} sélection : 211.21 HZ

➤ Tracé de la moyenne des fréquences fondamentales :

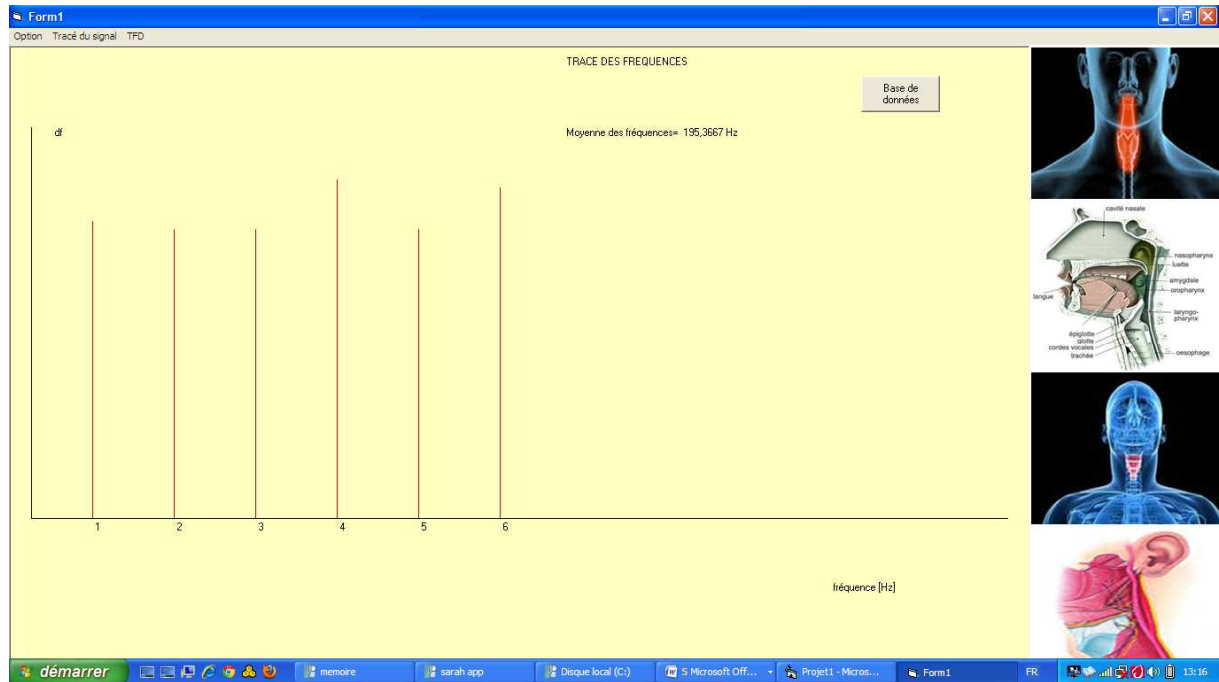


Figure 55 : Capture d'écran de la moyenne des fréquences fondamentales du signal vocal du 3^{ème} sujet sain.

La fréquence fondamentale moyenne est : 195.36 HZ

3.2. Sujets atteints d'une pathologie des cordes vocales :

3.2.1. le premier sujet (masculin) atteint d'un cancer du larynx :

➤ Tracé temporel du signal vocal :

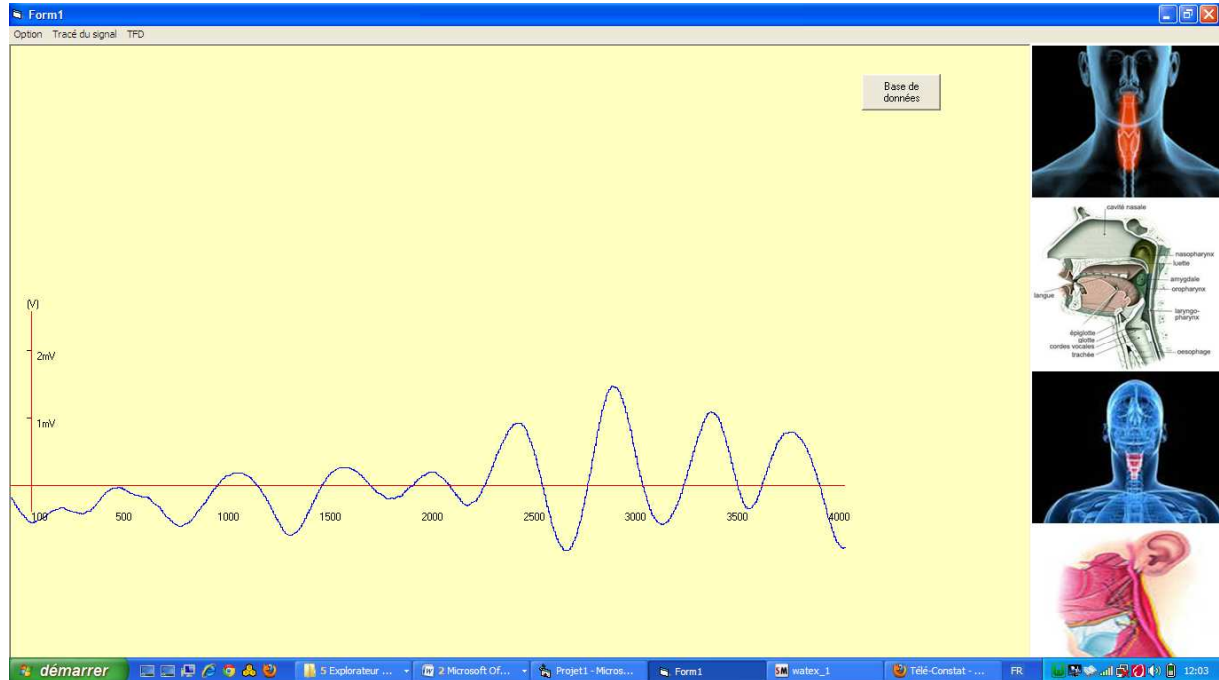


Figure 56 : Capture d'écran de tracé temporel du signal vocal du 1^{er} sujet pathologique.

➤ Tracé de la TFD CT :

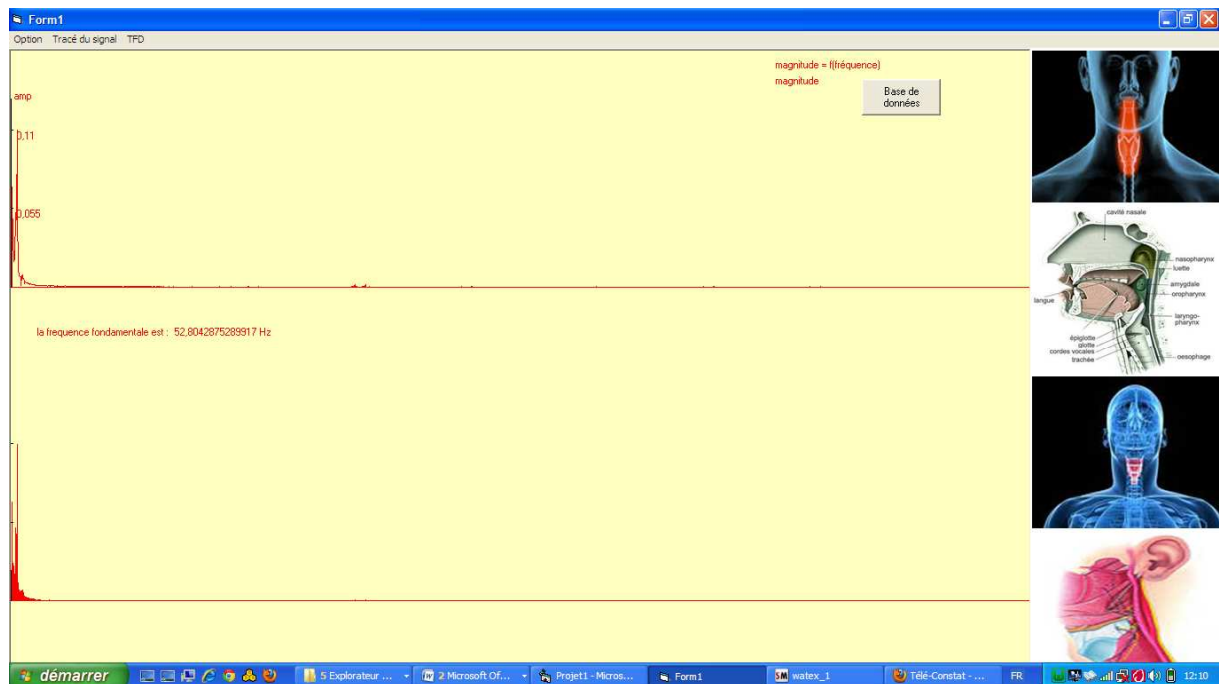


Figure 57 : Capture d'écran de la TFD CT du signal vocal du 1^{er} sujet pathologique.

➤ Tracé de zoom de la TFD :

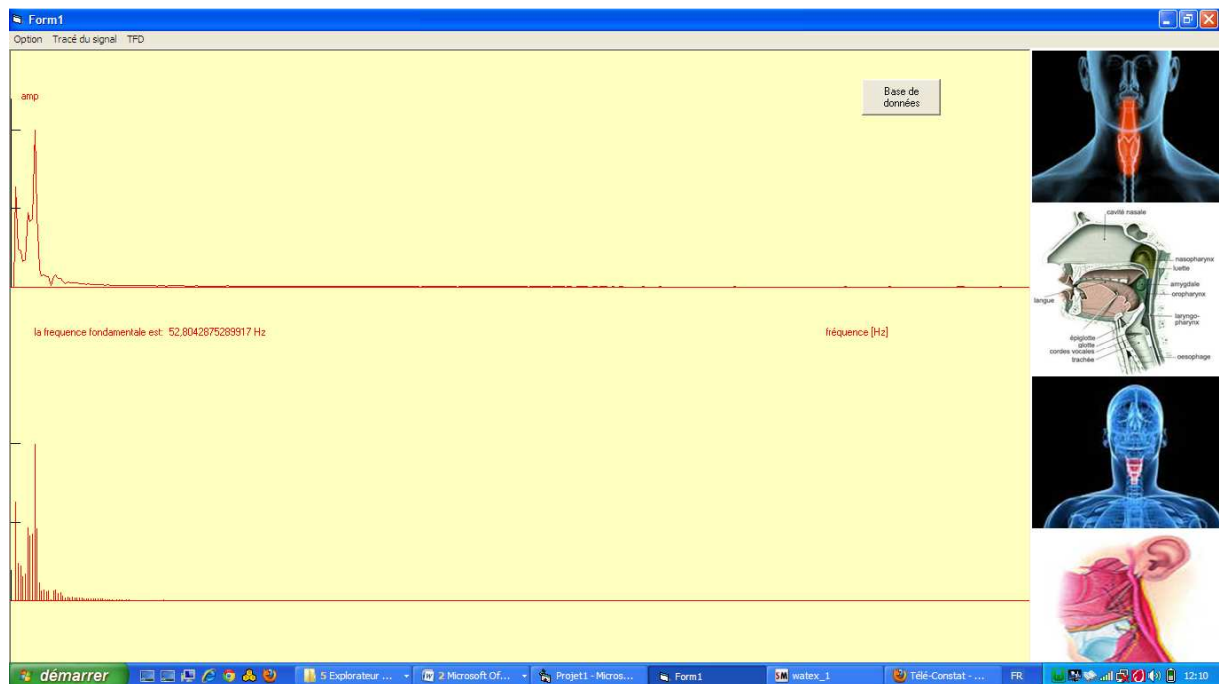


Figure 58 : Capture d'écran du zoom de la TFD CT du signal vocal du 1^{er} sujet pathologique.

Les fréquences fondamentales des 6 sélections :

La 1^{re} sélection : 47.52 HZ

La 2^{me} sélection : 52.8 HZ

La 3^{me} sélection : 52.8 HZ

La 4^{me} sélection : 47.58 HZ

La 5^{me} sélection : 58.08 HZ

La 6^{me} sélection : 58.08 HZ

➤ Tracé de la moyenne des fréquences fondamentales :

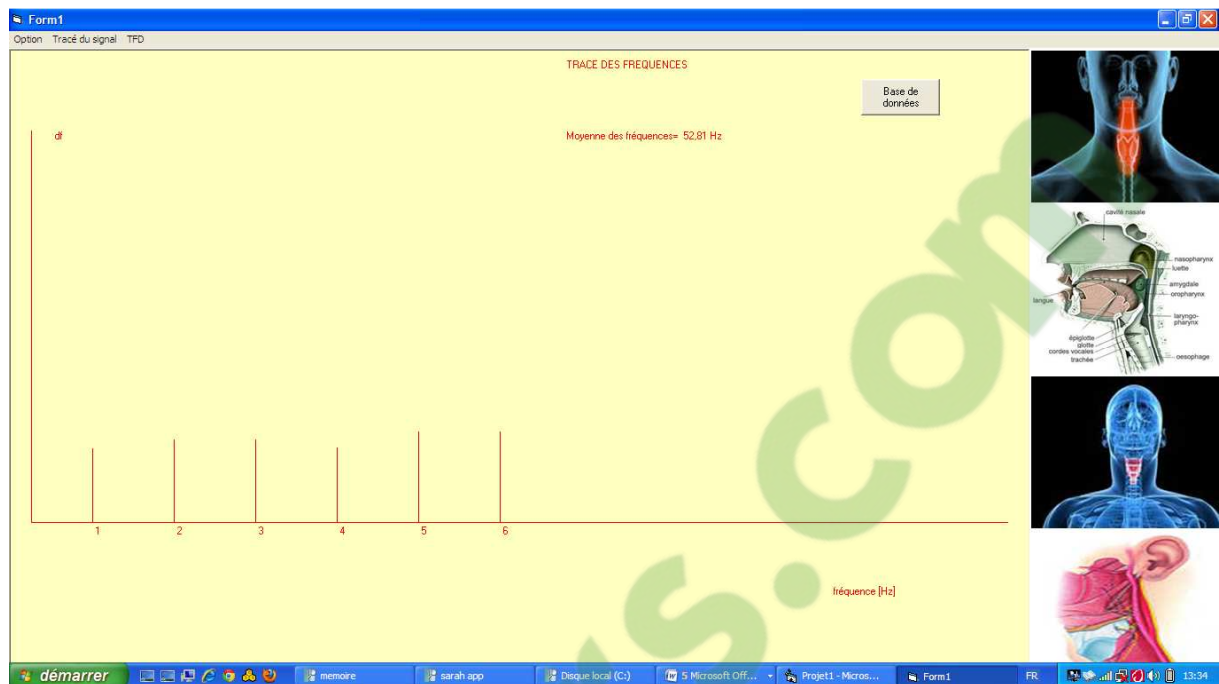


Figure 59 : Capture d'écran de la moyenne des fréquences fondamentales du signal vocal du 1^{er} sujet pathologique.

La fréquence fondamentale moyenne est : 52.81 HZ

3.2.2. le 2^{eme} sujet (masculin) atteint d'un polype inflammatoire :

➤ Tracé temporel du signal vocal :

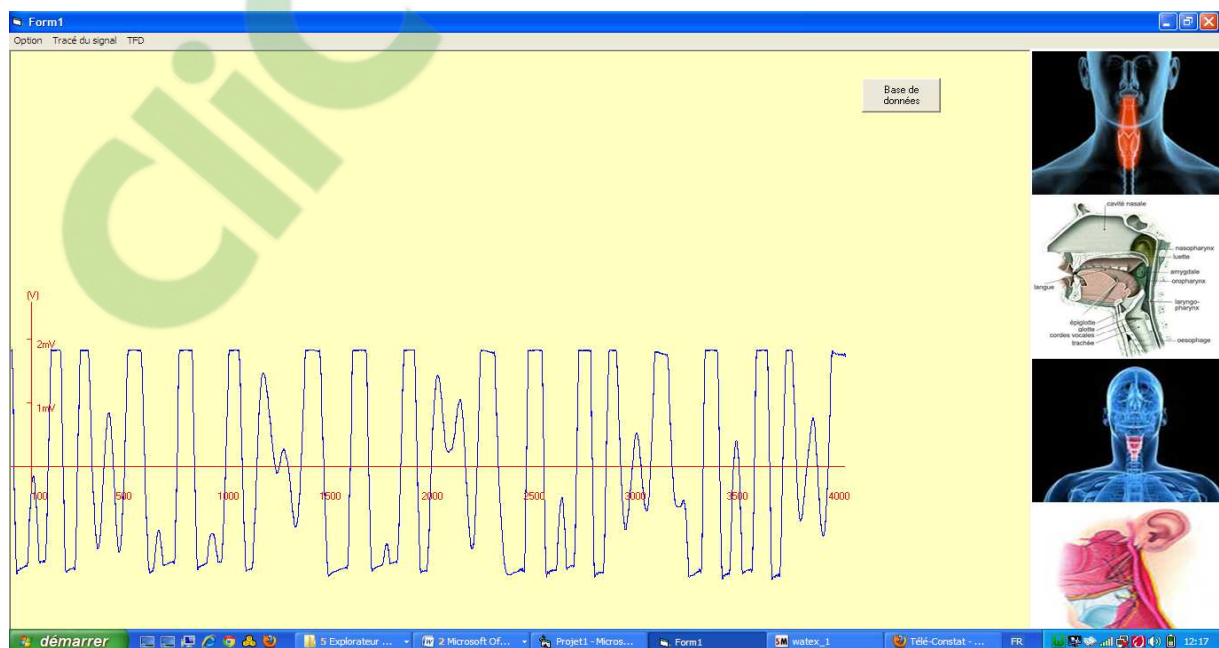


Figure 60 : Capture d'écran de tracé temporel du signal vocal du 2^{eme} sujet pathologique.

Chapitre VI : Evaluation clinique des résultats

➤ Tracé de la TFD CT :



Figure 61 : Capture d'écran de la TFD CT du signal vocal du 2^{ème} sujet pathologique.

➤ Tracé de zoom de la TFD :

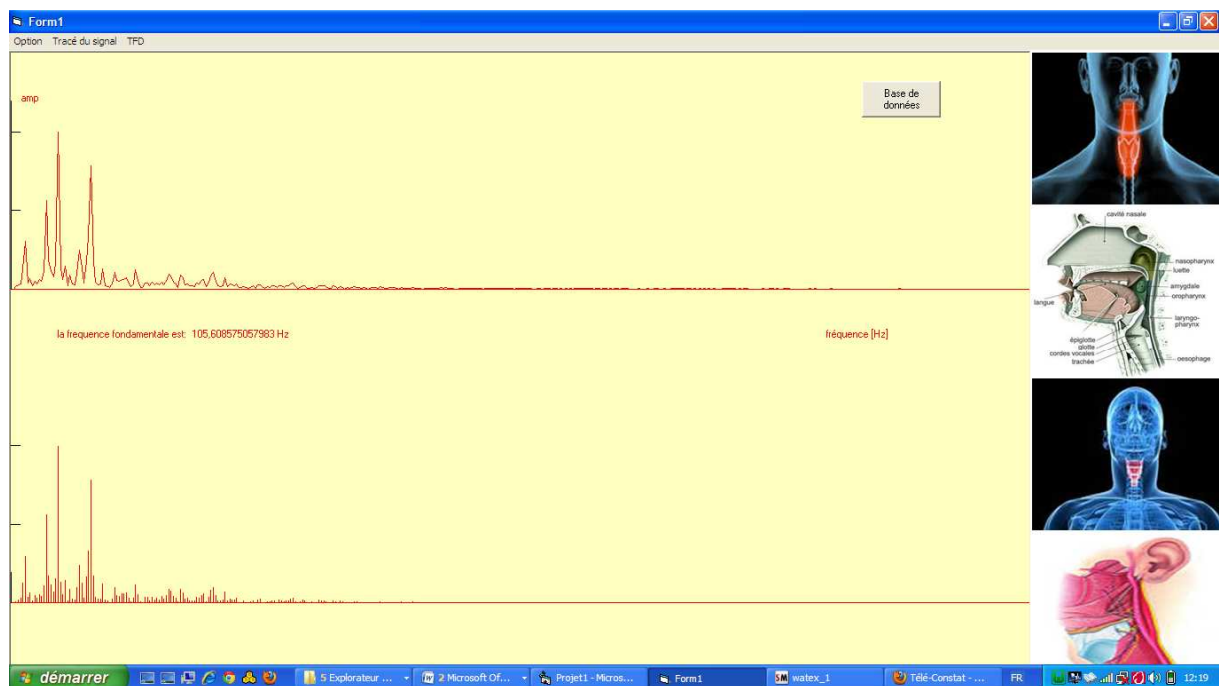


Figure 62 : Capture d'écran du zoom de la TFD CT du signal vocal du 2^{ème} sujet pathologique.

Les fréquences fondamentales des 6 sélections :

La 1^{ère} sélection : 121.44 HZ

La 2^{ème} sélection : 105.6 HZ

Chapitre VI : Evaluation clinique des résultats

La 3^{ème} sélection : 110.88 HZ

La 4^{ème} sélection : 100.32 HZ

La 5^{ème} sélection : 137.29 HZ

La 6^{ème} sélection : 137.29 HZ

➤ Tracé de la moyenne des fréquences fondamentales:

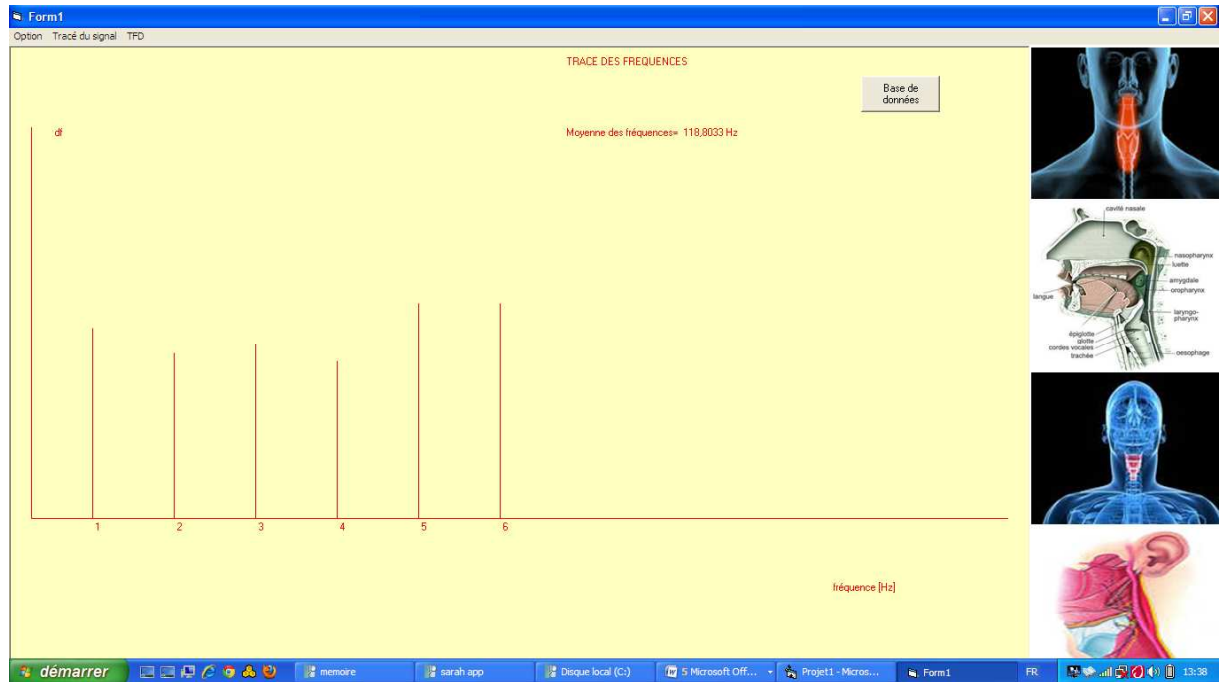


Figure 63 : Capture d'écran de la moyenne des fréquences fondamentales du signal vocal du 2^{ème} sujet pathologique.

La fréquence fondamentale moyenne est : 118.8 HZ

3.2.3. Le 3^{ème} sujet (masculin) atteint d'un cancer du larynx :

3.2.3.1. Premier enregistrement :

➤ Tracé temporel du signal vocal :

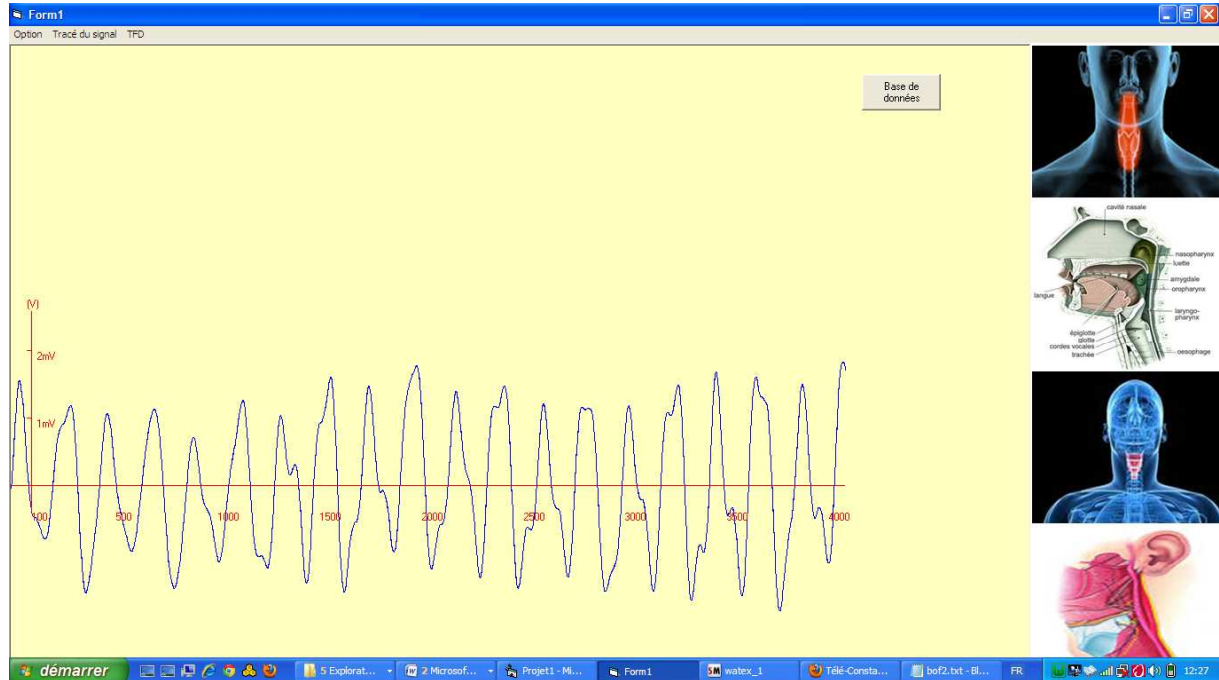


Figure 64 : Capture d'écran de tracé temporel du signal vocal du 3^{ème} sujet pathologique (1^{er} enregistrement).

➤ Tracé de la TFD CT :

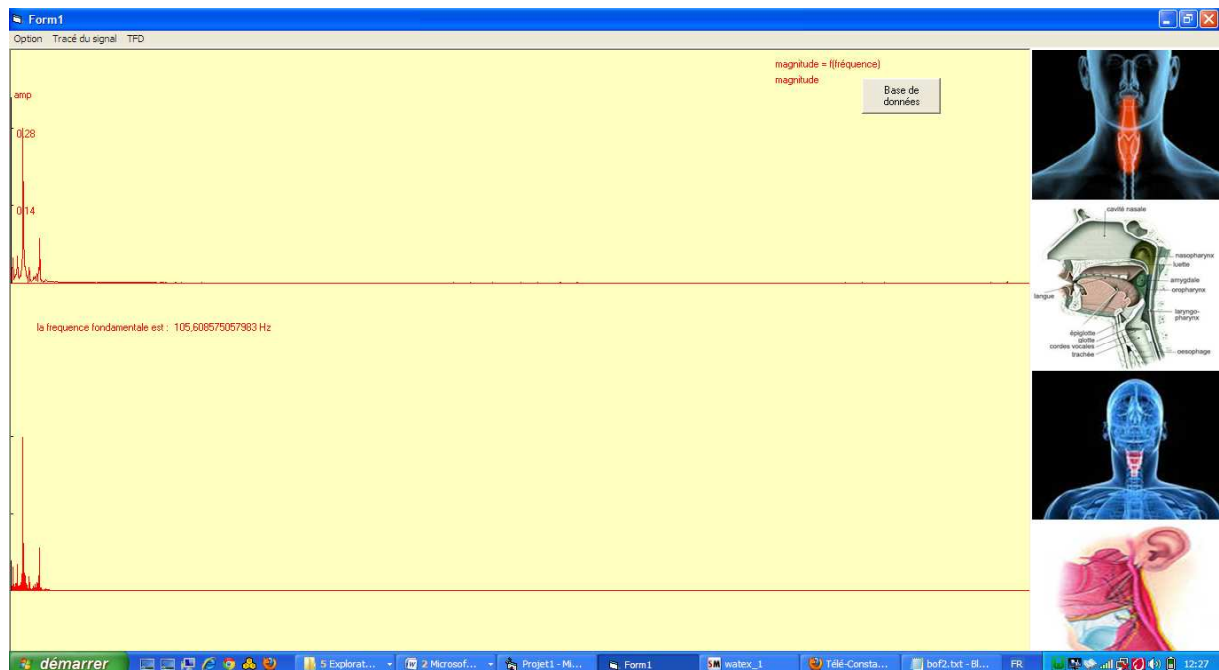


Figure 65 : Capture d'écran de la TFD CT du signal vocal du 3^{ème} sujet pathologique (1^{er} enregistrement).

➤ Tracé de zoom de la TFD :

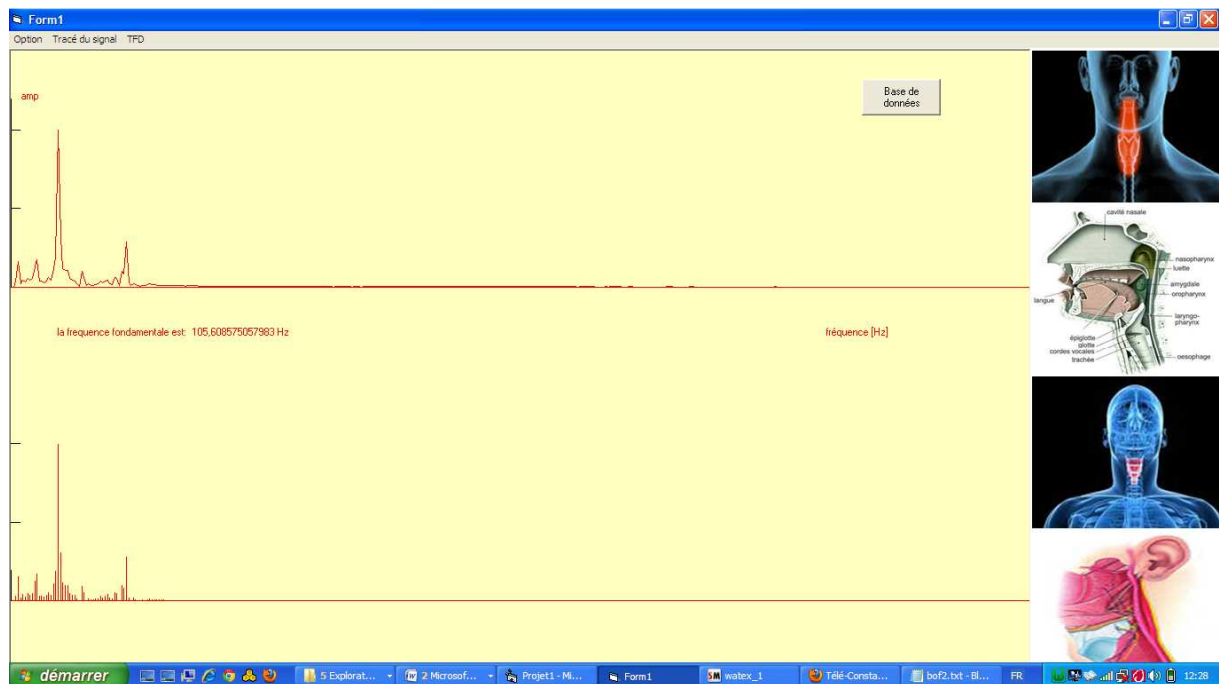


Figure 66 : Capture d'écran du zoom de la TFD CT du signal vocal du 3^{ème} sujet pathologique (1^{er} enregistrement).

Les fréquences fondamentales des 6 sélections :

La 1^{ère} sélection : 105.6 HZ

La 2^{ème} sélection : 137.29 HZ

La 3^{ème} sélection : 126.73 HZ

La 4^{ème} sélection : 110.88 HZ

La 5^{ème} sélection : 121.44 HZ

La 6^{ème} sélection : 105.6 HZ

Chapitre VI : Evaluation clinique des résultats

➤ Tracé de la moyenne des fréquences fondamentales:

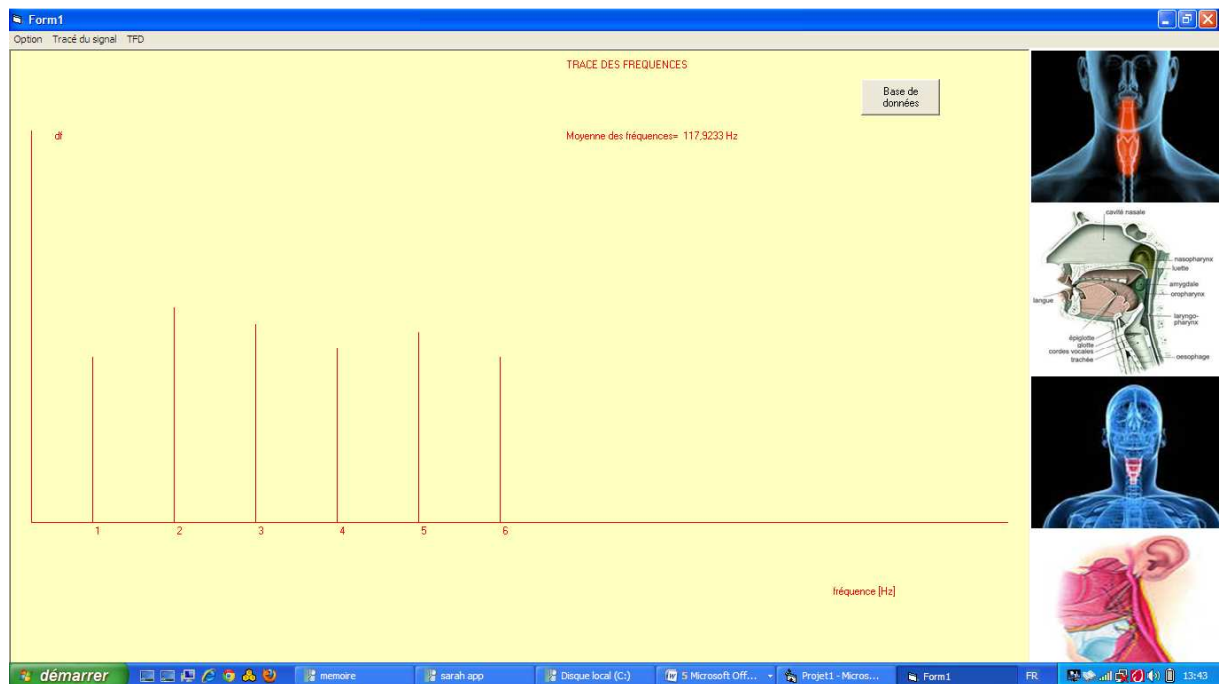


Figure 67 : Capture d'écran de la moyenne des fréquences du signal vocal du 3^{ème} sujet pathologique (1^{er} enregistrement).

La fréquence fondamentale moyenne est : 117.92 HZ.

3.2.3.2. 2^{ème} enregistrement (2 mois après) :

➤ Tracé temporel du signal vocal :

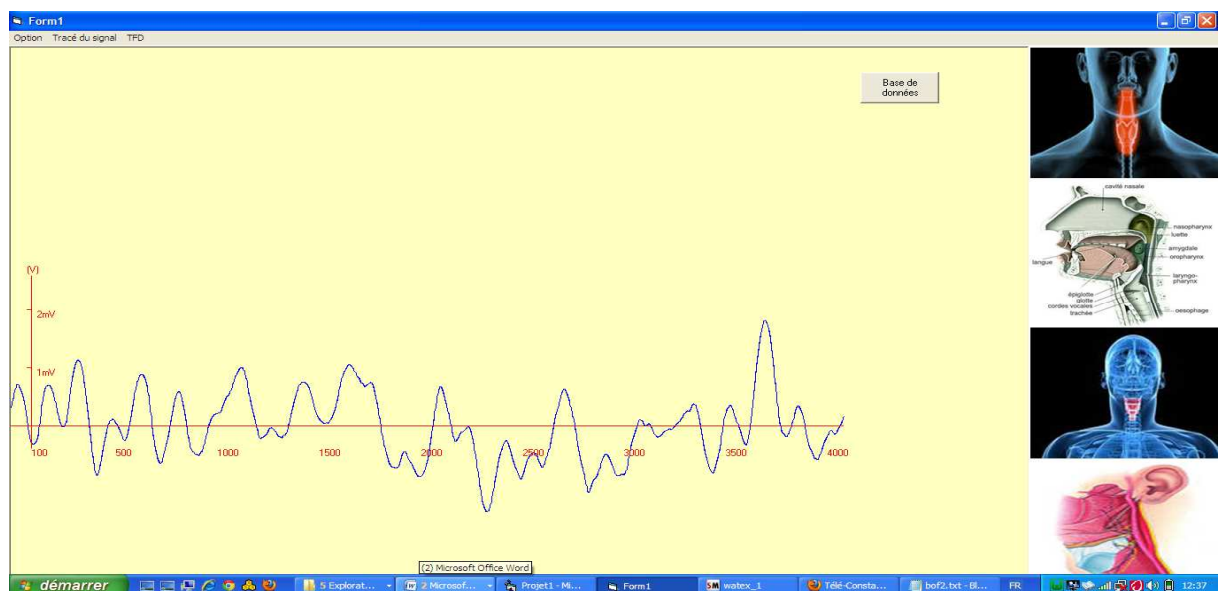


Figure 68 : Capture d'écran de tracé temporel du signal vocal du 3^{ème} sujet pathologique (2 mois après le 1^{er} enregistrement).

Chapitre VI : Evaluation clinique des résultats

➤ Tracé da la TFD CT :

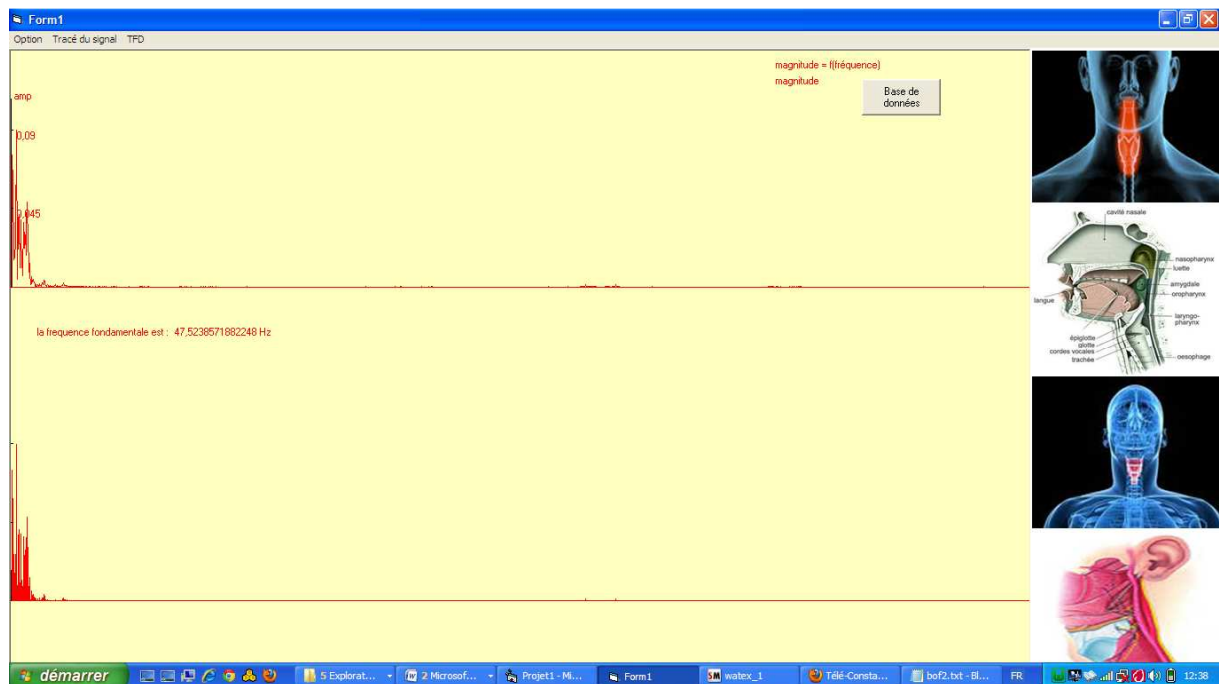


Figure 69 : Capture d'écran de la TFD CT du signal vocal du 3^{ème} sujet pathologique (2 mois après le 1^{er} enregistrement).

➤ Tracé de zoom de la TFD :

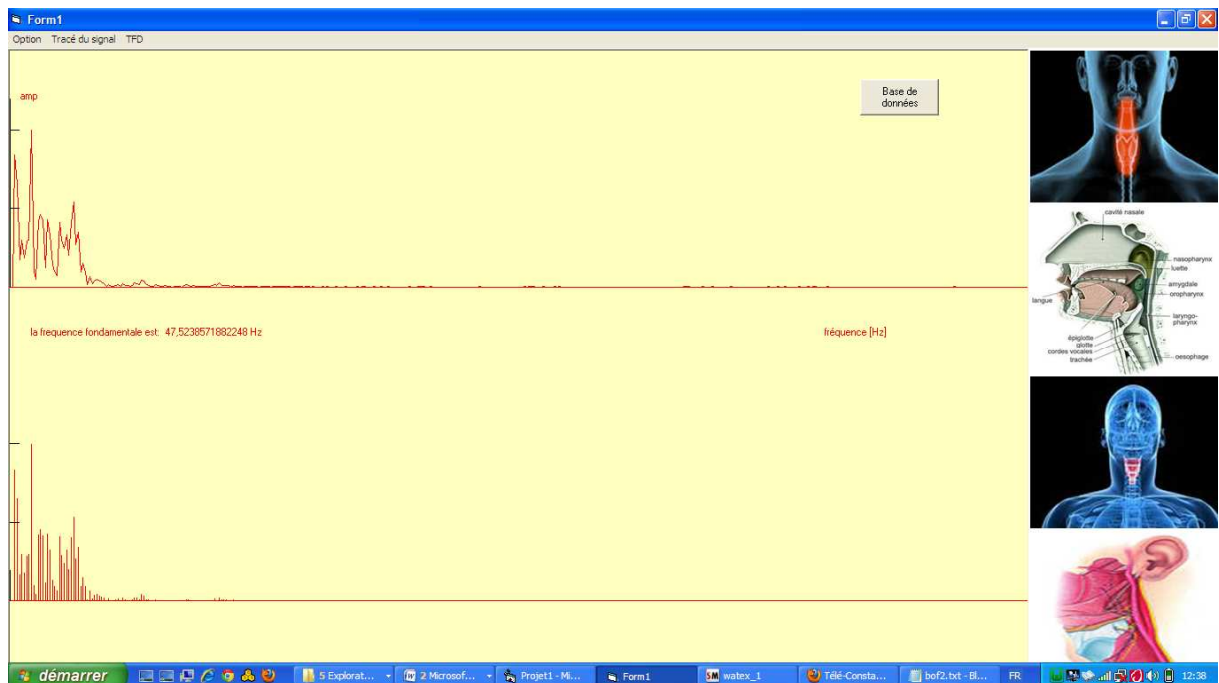


Figure 70 : Capture d'écran du zoom de la TFD CT du signal vocal du 3^{ème} sujet pathologique (2 mois après le 1^{er} enregistrement).

Chapitre VI : Evaluation clinique des résultats

Les fréquences fondamentales des 6 sélections :

La 1^{ère} sélection : 58.08 HZ

La 2^{ème} sélection : 47.52 HZ

La 3^{ème} sélection : 47.52 HZ

La 4^{ème} sélection : 42.24 HZ

La 5^{ème} sélection : 42.24 HZ

La 6^{ème} sélection : 47.52 HZ

La fréquence fondamentale moyenne est : 47.51 HZ.

➤ Tracé de la moyenne des fréquences fondamentales :

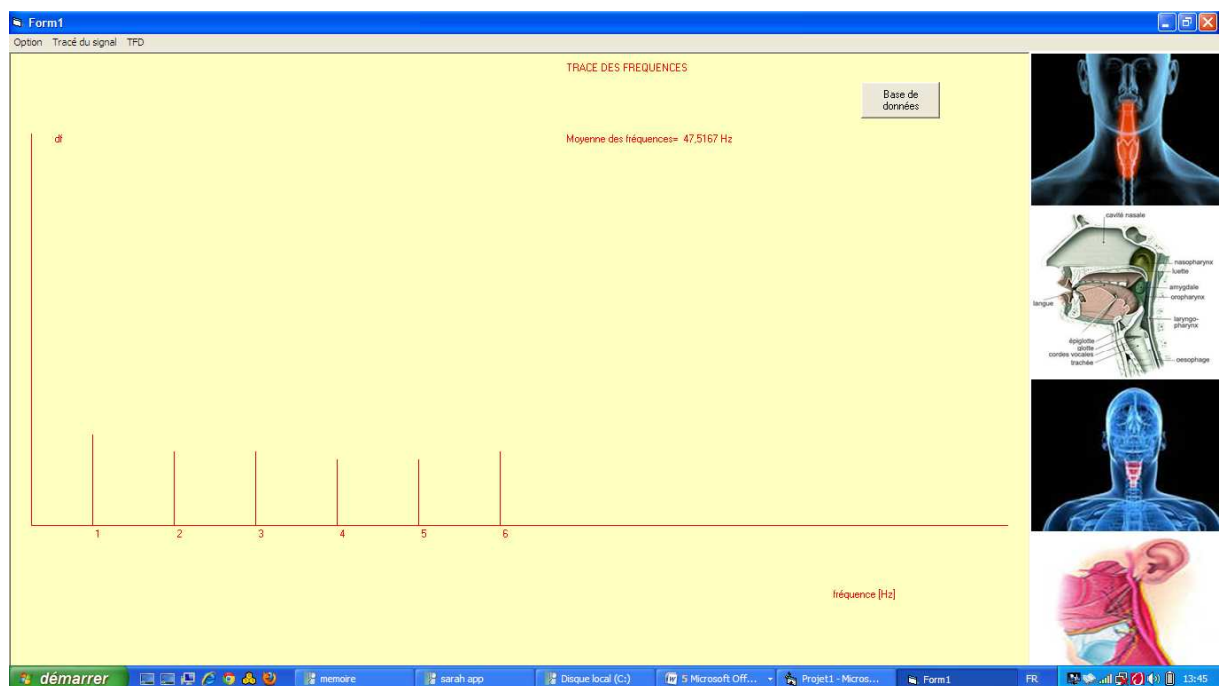


Figure 71 : Capture d'écran de la moyenne des fréquences fondamentales du signal vocal du 3ème sujet pathologique (2 mois après le 1er enregistrement).

4. Récapitulation des résultats :

	<i>Sujets sains</i>			<i>Sujets pathologiques</i>			
<i>Les fréquences fondamentales (HZ)</i>	<i>1^{er} sujet</i>	<i>2^{eme} sujet</i>	<i>3^{eme} sujet</i>	<i>1^{er} sujet</i>	<i>2^{eme} sujet</i>	<i>3^{eme} sujet 1^{er} enregistrement</i>	<i>3^{eme} Sujet 2^{eme} enregistrement (2 mois après)</i>
<i>1^{er} sélection (HZ)</i>	184.81	200.65	190.09	47.52	121.44	105.6	58.08
<i>2^{eme} sélection (HZ)</i>	211.21	211.2	184.8	52.8	105.6	137.29	47.52
<i>3^{eme} sélection (HZ)</i>	179.53	184.8	184.8	52.8	110.88	126.73	47.5
<i>4^{eme} sélection (HZ)</i>	184.81	184.8	216.49	47.58	100.32	110.88	42.24
<i>5^{eme} sélection (HZ)</i>	211.21	184.81	184.81	58.08	137.29	121.44	42.24
<i>6^{eme} sélection (HZ)</i>	184.8	195.37	211.21	58.08	137.29	105.6	47.52
<i>Moyenne des fréquences fondamentales (HZ)</i>	192.72	193.59	195.36	52.81	118.8	117.92	47.51

Tableau 4 : Les fréquences fondamentales des différents sujets sains et pathologiques.

5. Interprétation des résultats :

5.1. Sujets sains :

L'étude de la fréquence fondamentale du son émis par les sujets a travers le tracé du signal vocal relève une fréquence autour de 200 HZ.

5.2. sujets malades :

L'étude de la fréquence fondamentale du son émis par les sujets malades a travers le tracé du spectre du signal vocal relève:

- patient présentant un cancer du larynx : la fréquence fondamentale varie entre 47,51 HZ et 52,81 HZ.
- patient présentant des lésions bénignes du larynx type laryngites chroniques, polypes des cordes vocales... avec une fréquence fondamentale entre 117.92 HZ et 118.8 HZ.

6. Discussion diagnostique :

6.1. Les sujets sains :

Examen ORL dont la filière laryngé est strictement normale avec une bonne mobilité des cordes vocales.

La fréquence fondamentale du son émis calculée à travers le tracé du spectre du signal vocal est située autour de 200 HZ.

Le tracé temporel est quasi-périodique avec un spectre s'étalant vers les fréquences aigues.

6.2. Les sujets malades :

➤ Le premier sujet :

Patient âgé de 65 ans pêcheur de profession, tabagique chronique (25 paquets par mois) venant consulter pour une dysphonie chronique avec dyspnée et dont le bilan ORL révèle un cancer du larynx à un stade avancé et nécessitant un traitement chirurgical radical.

Son tracé temporel est irrégulier (aléatoire) et le spectre est limité dans les basses fréquences avec une fréquence fondamentale de 52.81 HZ.

➤ Le deuxième sujet :

Patient âgé de 45 ans tabagique chronique (30 paquets par mois) et alcoolique venant consulter pour une dysphonie chronique (trouble de la voix) évoluant depuis plus de 6 mois et dont le bilan ORL révèle un polype de la corde vocale dont l'examen dont l'examen ORL revenu d'un aspect inflammatoire bénin.

Son tracé temporel contient une certaine irrégularité avec aussi une similarité par rapport au tracé temporel des patients sains et un spectre contient des fréquences aigus mais moins que les patient sains sa fréquence fondamentale 118 .8 HZ.

➤ Le troisième patient:

Patient âgé de 61 ans retraité tabagique chronique, diabétique et hypertendu venant consulter pour une dysphonie chronique évoluant depuis de plus de 9 mois associe a une dyspnée laryngé sévère (gène respiratoire grave) ayant nécessité une trachéotomie a l'admission .le bilan ORL révèle un cancer du larynx nécessitant un traitement chirurgical radical.

Chapitre VI : Evaluation clinique des résultats

- Le premier enregistrement révèle des fréquences inférieures par rapport aux sujets sains aux alentours de 117.8 HZ.
- Le deuxième enregistrement (2 mois après) on note une diminution nette des fréquences similaire aux sujets malades et qui sont aux alentours de 47.51 HZ.

Son tracé temporelle est irrégulier avec un spectre évolutif dans le temps (au début le spectre contient des fréquences aigües mais moins que les patients sains) avec une fréquence fondamentale 117.8 HZ, après bilan ORL complet (2 mois) le spectre est devenu irrégulier avec une fréquence fondamentale basse 47.51 HZ).

7. conclusion :

Dans ce chapitre nous avons fait les tracés temporels et spectraux des sujets sains et des sujets malades et à partir du spectre nous avons calculé la fréquence fondamentale moyenne des 6 sélections choisies pour :

- le dépistage et la classification des dysphonies d'origine laryngées.
- le suivi (évolution dans le temps).

Conclusion générale et perspectives :

L'analyse et le traitement du signal vocal en vue du dépistage et du suivi des dysphonies vocales restent encore un domaine de recherche très vaste. Le but de ce travail a été de concevoir et d'implémenter une interface homme-machine dédiée à l'évaluation objective des dysphonies d'origine laryngée par le biais de la caractérisation spectro-temporelle du signal vocal acoustique.

Pour cela nous avons développé :

- Un système d'acquisition au moyen de la carte son pilotée par l'environnement logiciel Audacity permettant notamment une bande passante sonore s'étendant de 0 à 4 KHZ avec une fréquence d'échantillonnage de 8 KHZ et un enregistrement du signal vocal sous format WAVE.
- Une conversion du Format WAVE en format HEXADECIMAL sur 16 bits
- Un protocole expérimental basé sur l'émission d'un son voisé soutenu pendant 3 secondes en l'occurrence la voyelle 'a' la plus usitée dans ce genre d'exploration de la voix.
- Une application de traitement numérique de signal vocal par l'implémentation d'un algorithme de calcul de sa Transformée de Fourier Discrète à court terme (TFD-CT) et la détection de la fréquence fondamentale (F0) moyennée sur 6 trames. Celle-ci représente l'indice fréquentiel caractéristique du son aux alentours de 200 Hz chez le sujet adulte sain, résultat en concordance avec ceux retrouvés dans la bibliographie.
 - ✓ La fréquence fondamentale présente une grande variabilité chez les sujets présentant une pathologie des cordes vocales.
 - ✓ Cette variabilité de la FO paraît largement différente selon le type de la pathologie qu'elle soit inflammatoire ou tumorale.
- Ces enregistrements effectués in vivo et in situ au niveau du service ORL CHU Tlemcen nous ont permis d'entamer l'élaboration d'une base de données relative à la classification des pathologies laryngées selon le type inflammatoire ou tumoral en vue d'un dépistage précoce du cancer du larynx et sa prévention par l'étude épidémiologique des facteurs favorisants.

Chapitre VI : Evaluation clinique des résultats

Sur le plan fondamental nous envisageant comme perspective d'investigation l'étude approfondie des différents modèles de production de la parole en vue d'améliorer et de développer d'autres protocoles d'exploration de la voix.

Sur le plan expérimental il nous semble impératif d'élargir notre base de données en vue d'une validation clinique du système..

Et enfin développer tous les algorithmes relatifs à la caractérisation objective des dysphonies d'origine laryngées en particulier l'indice STD, les timbres, les formants, le jitter, le shimmer.

Nous tenons à remercier le personnel du service ORL, en particulier le Docteur Aounallah pour nous avoir accueilli, guidé, et soutenu par ses précieux conseils qui nous ont permis de comprendre la relation entre l'aspect clinique de la pathologie et les paramètres mesurables représentatifs signal vocal acoustique.

Clicours.COM

Bibliographie et Netographie

Bibliographie

- [1] : A.Drygajlo, « **Traitement de la parole** »-Partie 1-, école Polytechniques Fédérale de Lausanne EPFL, 1998.
- [2] : L.Benlaldj, « **Etude de la méthode PLP (Perceptual linear prediction) en reconnaissance automatique de la parole** », Thèse de magister en électronique, spécialité : Signaux et Systèmes, promotion 1999/2000.
- [3] : LAKHDAR Mohammed Hicham, HASSAM Ahmed, « classification des sons de la parole par la technique PLP » , Mémoire de fin d'étude pour l'obtention du diplôme d'ingénieur d'état en contrôle, Université ABOU BEKER BELKIAD TELEMEN ,2001.
- [4] : Calliope, la parole et son traitement automatique, Ed. Masson, 1989.
- [5] : T. Dutoit, « **Introduction au Traitement Automatique de la Parole** », Notes de cours, Faculté Polytechnique de Mons, 2000.
- [6] : Dr. ANDRZEJ DRYGAJLO, « **TRAITEMENT DE LA PAROLE** », “Interactive Multimodal Information Management (IM)2” , Laboratory of IDIAP (LIDIAP) Institute of Electrical Engineering (IEL) ,Swiss Federal Institute of Technology Lausanne (EPFL).
- [7]: Dr Camille Finck, chapitre sur « **L'évaluation fonctionnelle des dysphonie d'origine laryngé** », Service d'Oto-rhino-laryngologie, CHU Sart Tilman,Université de Liège, Liège, Belgium.
- [8] : Emmanuelle Guibert, « **Caractéristiques physiques et auditives d'un signal sonore** ».
- [9] : SUZANNE ROY, « **atelier de son et traitement sonore avec le logiciel audacity** »,11 mars 2003.
- [10] : SLIMANI hicham « **Conception et developement d'un systeme de teleconstat dedie a la police** », Mémoire de fin d'étude pour l'obtention du diplôme de master en télécommunication, Université ABOU BEKER BELKIAD TELEMEN ,2010.
- [11] : GUY ALMOUZNI, « **traitement de la parole** »,2011.

[12] : VAN LOO JONATHAN, «**Analyse du signal vocal : Détermination de F0 (fréquence fondamentale)** », 21/03/2008.

[13] : KARA Omar Ali, KOULA Bahmed, « **Gestion pédagogique des étudiants à travers un site web intranet.** » Mémoire de fin d'étude pour l'obtention du diplôme d'ingénieur d'état en Télécommunications, PROMOTION IGE 27, INSTITUT DES TÉLÉCOMMUNICATIONS ORAN, 2005/2006.

Neto-graphie

[1'] : <http://scgwww.epfl.ch/courses>

[2'] : <http://perso.aricia.fr/alluin/parole/prparole.htm9>

[3'] : http://www.espacefrancais.com/orthographe/alphabet_phonetique.html

[4'] : <http://www.ph-ludwigsburg.de/html/2b-frnz-s-01/overmann/baf3/phon/3k.htm>

[5'] : http://www.alluin.fr/article.php?id_article=13

[6'] : <http://fr.wikipedia.org/wiki/Formant>

[7'] : http://fr.wikipedia.org/wiki/Triangle_vocalique

[8'] : <http://www.vulgarisation-informatique.com/carte-son.php>

[9'] : <http://www.commentcamarche.net/contents/pc/carte-son.php3>

[10'] : <http://www.freesoundeditor.com/docwave.htm>

[11'] : <http://www.siteduzero.com/tutoriel-3-193603-preparer-son-ordinateur.html>

[12'] : <http://fr.wikipedia.org/> et <http://en.wikipedia.org/>

Liste des Figures :

FIGURE 1 : SCHEMA DE L'APPAREIL PHONATOIRE [3]	15
FIGURE 2 : CORDES VOCALES [1'].....	16
FIGURE 3 : LA PHONATION DES PHONEMES FRANÇAIS.....	18
FIGURE 4 : EXEMPLE DE SONS VOISE	18
FIGURE 5 : EXEMPLE DE SONS NON VOISE.....	19
FIGURE 6 : A. GLOTTE PENDANT LA RESPIRATION B. GLOTTE POUR LA PHONATION	19
GLOTTE 2. CORDES VOCALES 3. EPIGLOTTE 5. CARTILAGES	19
FIGURE 7 : LES PHONEMES DE LA LANGUE FRANÇAISE	21
FIGURE 8 : TRAPEZE ACOUSTIQUE DES VOYELLES ORALES DU FRANÇAIS.....	23
FIGURE 9 : LIEU D'ARTICULATION DES PHONEMES [4'].....	27
FIGURE 10 : LE MODELE PHYSIQUE DE LA PRODUCTION DE PAROLE.....	34
FIGURE 11 : REPRESENTATION TEMPORELLE D'UN SIGNAL ANALOGIQUE	41
FIGURE 12 : REPRESENTATION TEMPORELLE D'UN SIGNAL ECHANTILLONNE	41
FIGURE 13 : L-INTERFACE DE LOGICIEL AUDACITY	44
FIGURE 14 : LA BARRE D'OUTILS D'AUDACITY	45
FIGURE 15 : ENREGISTREMENT DU SON SUR AUDACITY	45
FIGURE 16 : LE MENU DES EFFETS D'AUDACITY.....	46
FIGURE 17 : LA BASE DE DONNEES DANS UN RESEAU	48
FIGURE 18 : SIGNAL SONORE PERIODIQUE, SOMME DE SONS PURS	51
FIGURE 19 : DEUX SONS PURS D'AMPLITUDE DIFFERENTE	52
FIGURE 20 : REPRESENTATION TEMPORELLE D'UN SON : SIGNAL GLOBAL ET ZOOM	53
FIGURE 21 : DEUX SPECTRES DE SIGNAUX HARMONIQUES DE MEME FREQUENCE FONDAMENTALE (220HZ)	54
FIGURE 22 : DEUX REPRESENTATIONS TEMPORELLES DE SIGNAUX HARMONIQUES DE MEME FREQUENCE FONDAMENTALE (220 HZ).....	55
FIGURE 23 : ENVELOPPE (TEMPORELLE) D'UNE NOTE JOUEE PAR UN INSTRUMENT	55
FIGURE 24 : CAPTURE D'ECRAN DE PHPMYADMIN.....	69
FIGURE 25 : CREATION D'UNE BD	69
FIGURE 26 : APRES LA CREATION DE LA BD	70
FIGURE 27 : APERÇU DES CHAMPS DE LA TABLE AVEC PHPMYADMIN	71
FIGURE 28: FONCTIONNEMENT DU PHP.....	74
FIGURE 29 : APERÇU DE LA PAGE D'ACCUEIL.....	76
FIGURE 30 : APERÇU DE LA PAGE DE LA CONNEXION (CONCERNANT L'ADMINISTRATION).....	77
FIGURE 31 : APERÇU DE LA PAGE D'ACCUEIL ADMINISTRATIVE (PAGE SECURISE)	77
FIGURE 32 : APERÇU DE LA PAGE FORMULAIRE DU PATIENT.....	78

FIGURE 33 : APERÇU DE LA PAGE VISUALISATION	79
FIGURE 34 : APERÇU DE LA PAGE COORDONNE D'UN PATIENT	80
FIGURE 35 : APERÇU DE LA PAGE ADMINISTRATEUR	81
FIGURE 36 : APERÇU DES PATIENTS RECHERCHES	81
FIGURE 37 : APERÇU DES PATIENTS ENREGISTRES	82
FIGURE 38 : CAPTURE D'ECRAN DE L'INTERFACE.....	88
FIGURE 39 : CAPTURE D'ECRAN DE LOGICIEL D'ENREGISTREMENT.....	89
FIGURE 40 : CAPTURE D'ECRAN DE LOGICIEL DE CONVERSION.....	89
FIGURE 41 : CAPTURE D'ECRAN DE L'APPELLE A DE LA BASE DE DONNEES.	90
FIGURE 42 : CAPTURE D'ECRAN D'ACCUEIL DE WAMPSEVER.....	90
FIGURE 43 : CAPTURE D'ECRAN DE LA PAGE D'ACCUEIL DE NOTRE SITE INTERNET.....	91
FIGURE 44 : CAPTURE D'ECRAN DE TRACE TEMPOREL DU SIGNAL VOCAL DU 1 ^{ER} SUJET SAIN.....	91
FIGURE 45 : CAPTURE D'ECRAN DE LA TFD CT DU SIGNAL VOCAL DU 1 ^{ER} SUJET SAIN.	92
FIGURE 46 : CAPTURE D'ECRAN DE ZOOM DE LA TFD CT DU SIGNAL VOCAL DU 1 ^{ER} SUJET SAIN.....	92
FIGURE 47 : CAPTURE D'ECRAN DE LA MOYENNE DES FREQUENCES FONDAMENTALES DU SIGNAL VOCAL DU 1 ^{ER} SUJET SAIN.....	93
FIGURE 48 : CAPTURE D'ECRAN DE TRACE TEMPOREL DU SIGNAL VOCAL DU 2 ^{EME} SUJET SAIN.....	94
FIGURE 49 : CAPTURE D'ECRAN DE LA TFD CT DU SIGNAL VOCAL DU 2 ^{EME} SUJET SAIN	94
FIGURE 50 : CAPTURE D'ECRAN DE ZOOM DE LA TFD CT DU SIGNAL VOCAL DU 2 ^{EME} SUJET SAIN.....	95
FIGURE 51 : CAPTURE D'ECRAN DE LA MOYENNE DES FREQUENCES FONDAMENTALES DU SIGNAL VOCAL DU 2 ^{EME} SUJET SAIN.....	96
FIGURE 52 : CAPTURE D'ECRAN DE TRACE TEMPOREL DU SIGNAL VOCAL DU 3 ^{EME} SUJET SAIN.....	96
FIGURE 53 : CAPTURE D'ECRAN DE LA TFD CT DU SIGNAL VOCAL DU 3 ^{EME} SUJET SAIN.	97
FIGURE 54 : CAPTURE D'ECRAN DU ZOOM DE LA TFD CT DU SIGNAL VOCAL DU 3 ^{EME} SUJET SAIN.	97
FIGURE 55 : CAPTURE D'ECRAN DE LA MOYENNE DES FREQUENCES FONDAMENTALES DU SIGNAL VOCAL DU 3 ^{EME} SUJET SAIN.....	98
FIGURE 56 : CAPTURE D'ECRAN DE TRACE TEMPOREL DU SIGNAL VOCAL DU 1 ^{ER} SUJET PATHOLOGIQUE.....	99
FIGURE 57 : CAPTURE D'ECRAN DE LA TFD CT DU SIGNAL VOCAL DU 1 ^{ER} SUJET PATHOLOGIQUE.....	99
FIGURE 58 : CAPTURE D'ECRAN DU ZOOM DE LA TFD CT DU SIGNAL VOCAL DU 1 ^{ER} SUJET PATHOLOGIQUE.	100
FIGURE 59 : CAPTURE D'ECRAN DE LA MOYENNE DES FREQUENCES FONDAMENTALES DU SIGNAL VOCAL DU 1 ^{ER} SUJET PATHOLOGIQUE.....	101
FIGURE 60 : CAPTURE D'ECRAN DE TRACE TEMPOREL DU SIGNAL VOCAL DU 2 ^{EME} SUJET PATHOLOGIQUE.	101
FIGURE 61 : CAPTURE D'ECRAN DE LA TFD CT DU SIGNAL VOCAL DU 2 ^{EME} SUJET PATHOLOGIQUE.....	102
FIGURE 62 : CAPTURE D'ECRAN DU ZOOM DE LA TFD CT DU SIGNAL VOCAL DU 2 ^{EME} SUJET PATHOLOGIQUE.	102
FIGURE 63 : CAPTURE D'ECRAN DE LA MOYENNE DES FREQUENCES FONDAMENTALES DU SIGNAL VOCAL DU 2 ^{EME} SUJET PATHOLOGIQUE.....	103
FIGURE 64 : CAPTURE D'ECRAN DE TRACE TEMPOREL DU SIGNAL VOCAL DU 3 ^{EME} SUJET PATHOLOGIQUE (1 ^{ER} ENREGISTREMENT).....	104

FIGURE 65 : CAPTURE D'ECRAN DE LA TFD CT DU SIGNAL VOCAL DU 3 ^{EME} SUJET PATHOLOGIQUE (1 ^{ER} ENREGISTREMENT).....	104
FIGURE 66 : CAPTURE D'ECRAN DU ZOOM DE LA TFD CT DU SIGNAL VOCAL DU 3 ^{EME} SUJET PATHOLOGIQUE (1 ^{ER} ENREGISTREMENT).....	105
FIGURE 67 : CAPTURE D'ECRAN DE LA MOYENNE DES FREQUENCES FONDAMENTALES DU SIGNAL VOCAL DU 3 ^{EME} SUJET PATHOLOGIQUE (1 ^{ER} ENREGISTREMENT).	106
FIGURE 68 : CAPTURE D'ECRAN DE TRACE TEMPOREL DU SIGNAL VOCAL DU 3 ^{EME} SUJET PATHOLOGIQUE (2 MOIS APRES LE 1 ^{ER} ENREGISTREMENT).....	106
FIGURE 69 : CAPTURE D'ECRAN DE LA TFD CT DU SIGNAL VOCAL DU 3 ^{EME} SUJET PATHOLOGIQUE (2 MOIS APRES LE 1 ^{ER} ENREGISTREMENT).	107
FIGURE 70 : CAPTURE D'ECRAN DU ZOOM DE LA TFD CT DU SIGNAL VOCAL DU 3 ^{EME} SUJET PATHOLOGIQUE (2 MOIS APRES LE 1 ^{ER} ENREGISTREMENT).	107
FIGURE 71 : CAPTURE D'ECRAN DE LA MOYENNE DES FREQUENCES FONDAMENTALES DU SIGNAL VOCAL DU 3 ^{EME} SUJET PATHOLOGIQUE (2 MOIS APRES LE 1 ^{ER} ENREGISTREMENT).....	108

Liste des Tableaux :

TABEAU 1 : LES SEMI VOYELLE DU FRANÇAIS.....	29
TABEAU 2 : LES SEMI VOYELLE DU FRANÇAIS.....	30
TABEAU 3 : FREQUENCES TYPIQUES DES FORMANTS.....	58
TABEAU 4 : LES FREQUENCES FONDAMENTALES DES DIFFERENTS SUJETS SAINS ET PATHOLOGIQUES.....	109

Annexe

Les voyelles(19) :

Les voyelles orales(12) :

phonèmes	Mot-clé
a	fa cu lté
ɛ	re g le
e	th e
i	fil l e
y	lu n ette
ø	fe u
œ	co eu r
ə	ce
u	po u le
o	ba te au
ɔ	po m me
ɑ	pa t e

Les voyelles nasales(4) :

phonèmes	Mot-clé
œ̃	un
ɛ̃	pa in
ã	gan t
õ	ba ll on

Les semi-voyelles(3) :

j	Fil l e
w	po i sson
ɥ	Hu i t

Les consonnes(18) :

Les consonnes occlusives(10) :

Phonèmes	Mot-clé
P	p ile
B	b ol
M	m ur
Dentales	
T	t able
D	d é
N	n oeud
ɲ	l igne

Vélares	
k	cadeau
N	parking

Les consonnes 6 fricatives / constrictives (6) :

f	flûte
v	valise
Dentales	
s	citron
z	maison
Palatales	
ʃ	chat
ʒ	jupe

Les consonnes vibrantes, constrictive(2) :

Phonèmes	Mot-clé
l	lampe
ʀ	roue

Résumé

Dans le cadre de ce mémoire nous développons un système dédié à la caractérisation objective des dysphonies chroniques d'origine laryngées. Le but de ce système est triple : diagnostic, thérapeutique et suivi de la maladie. Pour cela nous mettons à profit dans un premier temps l'enregistrement et l'archivage du signal vocal acoustique au moyen de l'environnement logiciel Audacity qui permet de délivrer un signal temporel sous le format WAVE. Notre contribution a consisté en l'implémentation sous environnement Visual Basic d'un algorithme permettant de réaliser l'analyse spectro-temporelle d'un signal vocal acoustique voisé en l'occurrence un « a » soutenu pendant trois secondes. Nous avons appliqué cet algorithme sur trois sujets sains et trois sujets pathologiques dont deux présentant un cancer du larynx et un présentant un polype de type inflammatoire des cordes vocales. 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 F_0 moyennée sur plusieurs trames du signal voisé.

Mots clés :

TFD CT, dysphonies vocales, sons voisés, jitter, shimmer, STD, formants, fréquence fondamentale.