

الجمهورية الجزائرية الديمقراطية الشعبية  
REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE

وزارة التعليم العالي والبحث العلمي

MINISTERE DE L'ENSEIGNEMENT SUPERIEUR  
ET DE LA RECHERCHE SCIENTIFIQUE



UNIVERSITE MOHAMED SEDDIK BENYAHIA JIJEL

Faculté des sciences et de la technologie

Département d'Electronique

N° :...../2023

## **MEMOIRE DE MASTER**

**DOMAINE: Sciences et Technologies**

**FILIERE: Télécommunications**

**SPECIALITE: Systèmes de Télécommunications**

### **Thème**

**Acquisition compressée de la parole pour les systèmes de  
télémédecine : application à la parole dysarthrique**

**Présenté Par : Melle Hanane ABDOU**

**Encadré Par : Pr. Abdellah KACHA**

**Melle Amina Manar ZENNIR**

**Date de soutenance: 25/06/2023**

#### **Jury de Soutenance**

**Président : Faycel Boukerroum**

**Grade Docteur**

**Univ MSB jijel**

**Encadreur : Abdellah Kacha**

**Grade Professeur**

**Univ MSB jijel**

**Examineur 1: Tahar Brahimi**

**Grade Professeur**

**Univ MSB jijel**

**Promotion : Juin 2022 /2023**

## ***Remerciements***

Nous souhaitons exprimer notre profonde gratitude envers notre encadrant de mémoire, Monsieur Abdellah Kacha, pour ses précieux conseils, son soutien constant et son expertise inestimable, qui ont grandement contribué à la réalisation de ce travail.

Nous tenons à remercier les membres du jury pour avoir accepté avec bienveillance de consacrer leur temps et leur expertise à l'évaluation de notre mémoire, aussi leur regard critique nous offre une perspective précieuse pour améliorer la qualité de notre recherche.

Nous exprimons également notre reconnaissance envers nos proches, nos familles et nos amis pour leur soutien moral et leur compréhension tout au long de ce processus de recherche, qui ont été une source précieuse de motivation.

Ce mémoire n'aurait pas été possible sans la contribution de tous ces acteurs. Leurs efforts et leur dévouement ont permis de mener à bien cette étude et d'en tirer des conclusions significatives.

Nous espérons que les résultats de cette recherche contribueront à l'avancement des connaissances dans le domaine de la télémédecine et serviront de base pour de futures études et applications cliniques.

Enfin, nous sommes profondément reconnaissants envers toutes les personnes qui ont joué un rôle essentiel dans la réalisation de ce travail.

## *Dédicaces*

À nos chers parents, dont les sacrifices, l'amour, la tendresse et le soutien infaillible ont jalonné notre parcours académique, nous exprimons notre profonde gratitude. Votre dévouement sans limite, vos encouragements constants et votre compréhension sans faille ont été des fondations solides sur lesquelles nous avons pu bâtir notre projet d'études. Vous êtes notre source d'inspiration et de motivation inépuisable.

À nos enseignants, encadrants et membres du personnel qui ont contribué à notre formation et à la réalisation de ce mémoire et qui nous ont accompagné tout au long de nos études.

À nos chers frères et sœurs, dont l'appui et les encouragements permanents et le soutien moral inébranlable ont été notre force, nous vous remercions. Votre présence constante et votre soutien indéfectible nous ont poussés à surmonter tous les obstacles rencontrés dans cette aventure universitaire exigeante.

À nos amies, qui ont été des piliers solides tout au long de notre parcours universitaire, nous vous sommes reconnaissants. Votre soutien inconditionnel, vos précieux conseils et votre soutien moral ont été d'une importance capitale pour notre réussite.

Nous dédions humblement ce travail, et nous sommes honorés de partager cette réalisation avec vous.

*Amina & Hanane*

## *Table des matières*

Remerciements .....	I
Dédicaces.....	II
Table des matières .....	III
Liste des figures.....	VII
Liste des tableaux .....	IX
Liste des abréviations .....	X
Introduction générale.....	1

### **CHAPITRE I : PRODUCTION DE LA PAROLE ET EVALUATION DES**

### **TROUBLES DE LA VOIX ..... 4**

<b>I.1 INTRODUCTION .....</b>	<b>5</b>
<b>I.2 LA PRODUCTION DE LA PAROLE .....</b>	<b>5</b>
I.2.1 APPAREIL PHONATOIRE.....	6
I.2.1.1 Soufflerie et vibration.....	7
I.2.1.2 Le larynx.....	7
I.2.1.3 Le corps sonore.....	8
I.2.2 LE SYSTEME ARTICULAIRE.....	9
<b>I.3 CATEGORIES DES SONS DE LA PAROLE .....</b>	<b>10</b>
I.3.1 CLASSIFICATION SELON LA SOURCE .....	10
I.3.1.1 Les sons voisés .....	10
I.3.1.2 Les sons non voisés .....	11
I.3.2 CLASSIFICATION PHONETIQUE .....	11
I.3.2.1 Les voyelles .....	11
I.3.2.2 Les consonnes.....	12
I.3.2.3 Les semi voyelles.....	12
<b>I.4 MODELISATION DE LA PRODUCTION DE LA PAROLE .....</b>	<b>12</b>
I.4.1 MODELE SOURCE-FILTRE .....	12
<b>I.5 CARACTERISTIQUE DU SIGNAL DE PAROLE .....</b>	<b>14</b>
I.5.1 LA FREQUENCE FONDAMENTALE (PITCH).....	15
I.5.2 SPECTRE FREQUENTIEL.....	15
I.5.3 L'ENERGIE.....	15
I.5.4 LES FORMANTS.....	16

<b>I.6 L'ANALYSE DE LA PAROLE POUR L'EVALUATION DES TROUBLES DE LA VOIX.....</b>	<b>16</b>
I.6.1 LES TROUBLES DE LA PAROLE ET DE LA VOIX .....	16
I.6.1.1 Troubles organiques.....	17
I.6.1.2 Troubles fonctionnelles .....	17
I.6.1.3 Troubles neurologiques .....	18
I.6.2 EVALUATION DES TROUBLES DE LA VOIX ET DE LA PAROLE .....	22
I.6.2.1 Evaluation subjective.....	22
I.6.2.2 Evaluation objective .....	23
(i). Indice de perturbation d'amplitude.....	23
(ii). Indices de perturbation des durées des cycles.....	24
(iii). Indices spectraux et cepstraux.....	25
<b>I.7 CONCLUSION.....</b>	<b>27</b>

## **CHAPITRE II : SYSTEMES DE TELEMEDECINE ..... 28**

<b>II.1 INTRODUCTION.....</b>	<b>29</b>
<b>II.2 DEFINITION DE LA TELEMEDECINE .....</b>	<b>29</b>
II.2.1 LA CONFIGURATION TYPIQUE D'UN SYSTEME DE TELEMEDECINE .....	30
<b>II.3 POURQUOI LE SYSTEME DE SANTE A-T-IL BESOIN DE LA TELEMEDECINE?.....</b>	<b>31</b>
<b>II.4 LES ACTES DE LA TELEMEDECINE .....</b>	<b>32</b>
II.4.1 LA TELE CONSULTATION.....	32
II.4.2 LA TELE SURVEILLANCE.....	32
II.4.3 LA TELE EXPERTISE.....	33
II.4.4 LA TELE RADIOLOGIE.....	33
II.4.5 LA TELE ASSISTANCE.....	33
II.4.6 LA TELE CHIRURGIE .....	34
II.4.7 LA TELE FORMATION .....	34
<b>II.5 AVANTAGES ET INCONVENIENTS DE LA TELEMEDECINE.....</b>	<b>34</b>
II.5.1 AVANTAGES.....	34
<b>II.6 INCONVENIENTS.....</b>	<b>35</b>
<b>II.7 LE SERVICE DE TELEMEDECINE POUR LES TROUBLES DE LA PAROLE ET DE LA VOIX.....</b>	<b>36</b>
II.7.1 CODAGE ET COMPRESSION DE LA PAROLE .....	37
II.7.2 CLASSIFICATION DES ALGORITHMES DE COMPRESSION .....	38
II.7.2.1 Compression sans perte .....	38

II.7.2.2 Compression avec perte .....	39
II.7.3 LA TECHNOLOGIE DE COMPRESSION MP3 .....	39
II.7.3.1 Principe et fonctionnement .....	39
II.7.3.2 Algorithmes de codage .....	40
<b>II.8 CONCLUSION.....</b>	<b>42</b>

**CHAPITRE III : ACQUISITION COMPRIMEE ..... 43**

<b>III.1 INTRODUCTION .....</b>	<b>44</b>
<b>III.2 BASE DE L'ACQUISITION COMPRIMEE.....</b>	<b>44</b>
<b>III.3 PRINCIPE DE L'ACQUISITION COMPRIMEE.....</b>	<b>47</b>
III.3.1 MODELE .....	47
III.3.2 INCOHERENCE.....	48
III.3.3 PARCIMONIE .....	50
III.3.4 DOMAINES DE TRANSFORMATION .....	51
(i). Transformée de Fourier discrète .....	51
(ii). Transformée en ondelettes discrètes.....	52
(iii). Transformée en cosinus discrète.....	52
III.3.5 TRAITEMENT DU SIGNAL.....	52
III.3.6 EXEMPLE D'ACQUISITION ET DE RECONSTRUCTION .....	54
<b>III.4 CONCLUSION.....</b>	<b>58</b>

**CHAPITRE IV : RESULTATS ET DISCUSSION..... 59**

<b>IV.1 INTRODUCTION .....</b>	<b>60</b>
<b>IV.2 BASE DE DONNEES.....</b>	<b>61</b>
<b>IV.3 COMPRESSION ET RECONSTRUCTION DU SIGNAL DE PAROLE PAR ACQUISITION COMPRIMEE.....</b>	<b>61</b>
IV.3.1 EXEMPLE D'ILLUSTRATION : .....	62
<b>IV.4 RESULTATS ET DISCUSSION .....</b>	<b>63</b>
IV.4.1 FREQUENCE FONDAMENTALE.....	64
IV.4.2 JITTER.....	66
IV.4.3 TREMOR.....	67
IV.4.4 SHIMMER .....	67

IV.4.5 RAPPORT HARMONIQUE SUR BRUIT (HNR) .....	69
IV.4.6 PROEMINENCE DU PIC CEPSTRAL (CPP) .....	70
<b>IV.5 CONCLUSION</b> .....	<b>71</b>
Conclusion générale .....	73

## *Liste des figures*

<b>Figure I-1</b> Anatomie de l'appareil phonatoire. ....	6
<b>Figure I-2</b> Anatomie du larynx. ....	7
<b>Figure I-3</b> Processus oro-nasale. ....	10
<b>Figure I-4</b> Modèle source filtre de la production de la parole. ....	13
<b>Figure II-1</b> Schématisation d'un système de télé expertise. ....	31
<b>Figure II-2</b> Chaîne de transmission du signal de parole pour les services de télémédecine .....	37
<b>Figure III-1</b> Processus de l'acquisition comprimée et de reconstruction du signal. ....	45
<b>Figure III-2</b> Le modèle mathématique de l'acquisition comprimée. ....	47
<b>Figure III-3</b> Échantillons aléatoires du signal original (en haut) et transformée en cosinus discrète inverse du signal (en bas). ....	56
<b>Figure III-4</b> Solution <b>I1</b> de l'équation $Ax = b$ . ....	56
<b>Figure III-5</b> Solution <b>I2</b> de l'équation $Ay = b$ . ....	57
<b>Figure IV-1</b> Illustration du système de télésanté pour la détection et l'évaluation de la dysarthrie. ....	62
<b>Figure IV-2</b> Signal reconstruit par acquisition compressée (en rouge) superposé au signal original (en bleu). (a) Locuteur sain. (b) Locuteur Parkinsonien. ....	64
<b>Figure IV-3</b> Boîte à moustaches de la fréquence vocale en Hz calculée à partir du signal original, et du signal reconstruit par minimisation <b>I1</b> et du signal MP3 décompressé pour les locuteurs contrôles et Parkinsoniens. (a) Fréquence vocale des Femmes. (b) Fréquence vocale des Hommes ....	65
<b>Figure IV-4</b> Boîte à moustaches du jitter calculé à partir du signal original, et du signal reconstruit par minimisation <b>I1</b> et du signal MP3 décompressé pour les locuteurs contrôles et Parkinsoniens. ....	66
<b>Figure IV-5</b> Boîte à moustaches du tremor calculé à partir du signal original, et du signal reconstruit par minimisation <b>I1</b> et du signal MP3 décompressé pour les locuteurs contrôles et Parkinsoniens. ....	67
<b>Figure IV-6</b> Boîte à moustaches du shimmer calculé à partir du signal original, et du signal reconstruit par minimisation <b>I1</b> et du signal MP3 décompressé pour les locuteurs contrôles et Parkinsoniens. ....	68



<b>Figure IV-7</b> Boîte à moustaches du rapport harmonique sur bruit calculé à partir du signal original, et du signal reconstruit par minimisation <b>I1</b> et du signal MP3 décompressé pour les locuteurs contrôles et Parkinsoniens.....	69
<b>Figure IV-8</b> Boîte à moustaches de la proéminence du pic cepstral (CPP) calculée à partir du signal original, et du signal reconstruit par minimisation <b>I1</b> et du signal MP3 décompressé pour les locuteurs contrôles et Parkinsoniens. ....	70

## *Liste des tableaux*

**Tableau IV.1** Valeurs des indices acoustiques calculés à partir du signal original, du signal reconstruit par minimisation de la norme ***l1*** et du signal MP3 décompressé... 63

## *Liste des abréviations*

<b>ASHA</b>	<b>American Speech and Hearing Association</b>
<b>GARBAS-I</b>	<b>Grade, Roughness, Breathiness, Asthenia, Strain, Instability</b>
<b>SLA</b>	<b>Sclérose Latérale Amyotrophique</b>
<b>TAP</b>	<b>Traitement Automatique de la Parole</b>
<b>ShdB</b>	<b>Shimmer Absolu</b>
<b>Shim</b>	<b>Shimmer Local</b>
<b>APQ3</b>	<b>Three-point Amplitude Perturbation Quotient</b>
<b>APQ5</b>	<b>Five-point Amplitude Perturbation Quotient</b>
<b>Jitta</b>	<b>Jitter Absolu</b>
<b>Jitt</b>	<b>Jitter Local</b>
<b>RAP</b>	<b>Relative Average Perturbation</b>
<b>PPQ5</b>	<b>Five-point Period Perturbation Quotient</b>
<b>HNR</b>	<b>Harmonic to Noise Ratio</b>
<b>CPP</b>	<b>Proéminence du Pic Cepstral</b>
<b>dB</b>	<b>Décibel</b>
<b>MP3</b>	<b>MPEG Audio Layer 3</b>
<b>MPEG</b>	<b>Moving Picture Experts Group</b>
<b>WBAN</b>	<b>Wireless Body Area Network</b>
<b>CS</b>	<b>Compressed Sensing</b>
<b>DFT</b>	<b>Discrete Fourier Transform</b>
<b>DWT</b>	<b>Discrete Wavelet Transform</b>
<b>DCT</b>	<b>Discrete Cosine Transform</b>
<b>RIP</b>	<b>Restricted Isometry Perturbation</b>
<b>CD</b>	<b>Compact Disc</b>

---

## *Introduction générale*

---

## *Introduction générale*

La voix est l'outil de communication le plus important dans la vie humaine. Pour certaines personnes telles que les enseignants, les chanteurs et les employés des centres d'appels, c'est l'outil sur lequel repose leur vie professionnelle. Cependant, certaines conditions médicales peuvent entraîner des troubles de la voix et/ou de la parole, rendant la communication difficile pour les personnes atteintes. La dysarthrie est l'un de ces troubles de la parole. Elle se caractérise par une altération de la production des sons de la parole, généralement causée par des problèmes musculaires ou neurologiques. Pour évaluer et traiter efficacement la dysarthrie, une évaluation précise et efficace de la parole et de la voix est nécessaire.

La demande sur les systèmes de santé du monde entier ne cesse de croître. De nouvelles méthodes de planification et de prestation des services de santé en réponse à la demande croissante pour ces services sont alors proposés. La télémédecine utilise les technologies de l'information et de la communication afin d'élargir l'accès des patients aux soins ainsi que d'économiser du temps et de l'argent. La télémédecine offre un moyen de diagnostiquer et de traiter les patients même lorsque le patient et le professionnel de la santé ne sont pas situés dans les mêmes lieux physiques.

Récemment, à la suite de la pandémie de COVID-19, en raison de la propagation rapide du virus et du nombre élevé de patients nécessitant une hospitalisation et un traitement médical, les services de santé ont dû faire face à une nouvelle situation concernant les soins aux patients [1]. Les cliniciens doivent assurer des soins médicaux à leurs patients même en situation de pandémie.

Le nombre de travaux de recherche qui ont étudié la fiabilité et l'efficacité du suivi et de l'évaluation de la voix et de la parole par télésanté reste faible [2-3]. En comparant les évaluations perceptives dans les modalités face à face et télésanté, seul quelques indices tels que la gravité de la dysarthrie, l'intelligibilité au niveau de la phrase et de la conversation, et la précision articulatoire ont été identifiés comme étant en bon accord entre modalités, ainsi qu'au sein et entre les évaluateurs [4-6].

Dans le contexte de l'application de la télémédecine pour l'évaluation des troubles de la voix et de la parole, quelques chercheurs ont considéré la faisabilité et la fiabilité de

l'évaluation perceptuelle et la précision des mesures acoustiques de la voix dans une variété de plateformes de télé pratique [7-9]. Récemment, un modèle consacré à l'évaluation de la voix par télésanté a été développé pour les patients des services d'oto-rhino-laryngologie (ORL) pendant une pandémie mondiale [10].

Les applications récentes de la télémédecine pour l'évaluation clinique de la voix dépendent de la technologie existante développée pour les communications cellulaires et Internet. La technique de compression du signal est un élément crucial de ces systèmes [11]. Le format MP3 développé par le groupe d'experts en images animées (MPEG) pour la compression audio est largement utilisé. Cependant, la compression du signal vocal au format MP3 présente de nombreuses limitations, en particulier pour la télémédecine, telles que la transmission sécurisée [11] et l'authenticité des données [13].

Une méthode d'acquisition et de compression simultanée du signal appelée acquisition comprimée (CS : compressed sensing) est devenue le centre des recherches actuelles en acquisition et compression de données [14]. Sous certaines conditions, le CS combine simultanément les étapes d'échantillonnage et de compression dans un processus unique avec des propriétés intéressantes [15] pour stocker et transmettre seulement quelques coefficients non nuls, puis l'optimisation non linéaire est utilisée pour récupérer le signal original à partir d'un petit nombre de mesures ce qui réduit considérablement le temps d'acquisition des données et la quantité de données stockées à transmettre. La technique CS possède des performances de compression élevées, d'un codage non adaptatif et d'une indépendance d'encodage et de décodage [16]. De plus, plusieurs travaux ont rapporté la capacité de CS à résoudre les problèmes de la transmission sécurisée et de l'authenticité des données [17-18].

L'effet de l'acquisition comprimée de la parole dans le cadre de l'évaluation des troubles de la voix par télésanté a été étudié dans [19]. Les résultats ont montré que pour certains indices acoustiques, l'acquisition comprimée présente des performances en termes de corrélation avec les scores de l'évaluation auditive meilleures que la compression MP3. Motivé par les résultats obtenus dans [19], on se propose d'utiliser l'acquisition comprimée pour la compression des signaux de parole dans les applications de télémédecine pour l'évaluation de la dysarthrie. Nous analysons les indices acoustiques connus pour leur pertinence à caractériser la parole dysarthrique.

Le mémoire est organisé comme suit :

**Le chapitre 1** présente des notions de base sur la production de la parole et les méthodes d'évaluation des troubles de la voix et de la parole.

**Le chapitre 2** est une introduction aux systèmes de télémédecine tout en donnant un intérêt particulier à la compression de la parole.

**Le chapitre 3** présente les principes fondamentaux de l'acquisition comprimée et illustre le processus de compression et de reconstruction du signal.

**Le chapitre 4** présente les résultats obtenus en utilisant une base de données incluant des signaux de parole produite par des locuteurs sains et des locuteurs dysarthriques.

Enfin, on termine par une conclusion générale qui résume les résultats obtenus.

---

# **Chapitre I**

**Production de la parole et évaluation des troubles de la voix**

---



## **I.1 Introduction**

La parole est un mode de communication qui implique la production de séquences sonores et de moments de silence. C'est un système unique utilisé par les êtres humains pour exprimer leur pensée à travers des sons articulés.

Cette capacité à structurer des systèmes de communication basés sur la variation de la pression sonore est exclusive à l'espèce humaine [20].

Ce chapitre se concentre sur les aspects clés de la production de la parole, notamment les caractéristiques de la voix, la classification des sons, le modèle source-filtre, les propriétés du signal de la parole, ainsi que les troubles de la voix et de la parole et leur évaluation. En explorant ces éléments, nous pourrions mieux comprendre comment les organes et les structures de l'appareil phonatoire contribuent à la formation des sons de la parole. Nous examinerons également comment les caractéristiques vocales telles que la hauteur, l'intensité et le timbre sont importantes pour la communication verbale. De plus, nous étudierons les différentes catégories de sons de la parole, tels que les sons voisés et non voisés, et les modèles source-filtre qui décrivent leur production. En analysant les propriétés et les caractéristiques du signal de la parole, nous découvrirons comment il est perçu et interprété par l'auditeur. Enfin, nous aborderons les troubles de la voix, en distinguant les troubles fonctionnels et organiques, et nous discuterons des méthodes d'évaluation subjective et objective utilisées pour évaluer ces troubles.

## **I.2 La production de la parole**

La production de la parole nécessite la contribution de divers organes en plus des éléments du système nerveux central et périphérique. Parmi ces organes, on compte le larynx, le pharynx, les poumons, la langue, les lèvres, les cavités buccale et nasale, ainsi que d'autres structures anatomiques. Cette section vise à fournir une description détaillée de chaque organe impliqué dans ce processus complexe de production de la parole, mettant en évidence leurs rôles spécifiques et leurs interactions pour assurer une articulation et une phonation adéquates.

Le système de production de la parole peut être divisé en deux grands sous-systèmes qui sont : l'appareil phonatoire et le système articulatoire [21].

## I.2.1 Appareil Phonatoire

L'appareil vocal, également connu sous le nom de système phonatoire, est composé de quatre éléments fondamentaux qui travaillent en étroite synergie pour produire des signaux acoustiques. Dans l'ordre où ils se développent :

**La soufflerie** est constituée des poumons et de la trachée artère qui fournissent l'air nécessaire à la production de la parole. Les muscles du thorax et de l'abdomen sont responsables de l'actionnement des poumons.

**Le vibreur est le larynx**, où se trouvent les cordes vocales qui génèrent les ondes acoustiques nécessaires à la production des sons vocaux.

**Le corps sonore** comprend divers résonateurs, principalement le pharynx et la bouche, qui influencent la qualité sonore et la résonance des sons produits.

**Le système articulateur** est composé d'éléments fixes et mobiles, tels que la langue, les lèvres et les mâchoires, qui permettent de modifier la forme de l'onde sonore produite par le vibreur.

Tous ces éléments fonctionnent en étroite coordination, contrôlés par le système nerveux central, assurant ainsi le synchronisme et la coordination nécessaires pour la production fluide de la parole. La figure I.1 [22], fournit une représentation visuelle de l'anatomie de l'appareil phonatoire, mettant en évidence les différents organes et leurs interactions. Elle complète et renforce la description détaillée présentée dans cette section

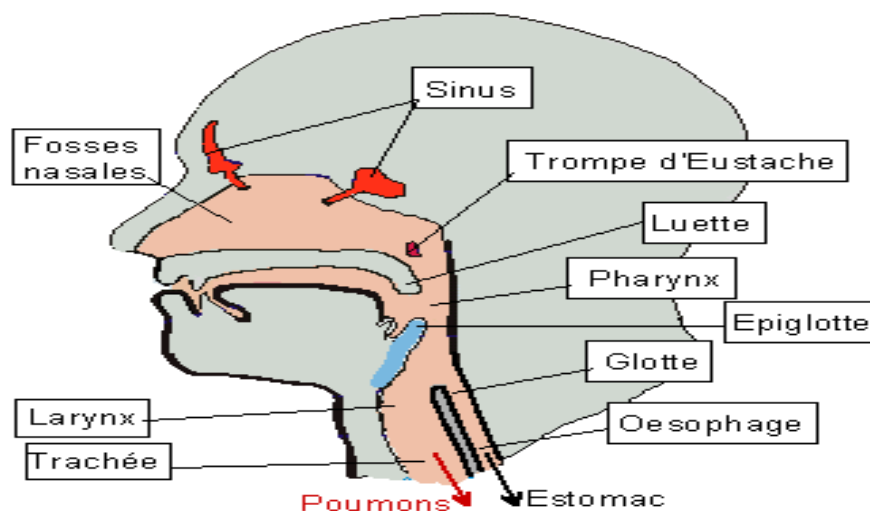


Figure I-1 Anatomie de l'appareil phonatoire.

### I.2.1.1 Soufflerie et vibration

L'air est essentiel pour la production de la voix. Le fonctionnement de notre appareil vocal peut être comparé à celui d'un instrument de musique, mais il est plus précisément similaire à un instrument à vent. En effet, notre système respiratoire agit comme une soufflerie en expulsant l'air provenant des poumons à travers la trachée. Cela crée le "souffle phonatoire" nécessaire à la production de la voix. Ce souffle peut être généré par l'abaissement de la cage thoracique ou par l'action des muscles abdominaux lors de la projection vocale [23].

### I.2.1.2 Le larynx

Le larynx, composé de muscles et de cartilages mobiles, est situé dans la partie supérieure de la trachée artère. Il joue un rôle crucial dans la modulation de la pression de l'air avant qu'il ne soit appliqué au conduit vocal. Les cordes vocales, qui sont en réalité deux lèvres symétriques au sein du larynx, peuvent se fermer complètement, formant ainsi une ouverture triangulaire appelée glotte. L'air circule librement à travers la glotte lors de la respiration, du chuchotement et de la production de sons non-voisés. En revanche, les sons voisés sont le résultat de vibrations périodiques des cordes vocales. Initialement, le larynx est complètement fermé, ce qui augmente la pression en amont des cordes vocales. Cette augmentation de pression entraîne l'ouverture des cordes vocales, réduisant ainsi la pression et permettant leur fermeture ultérieure [20].

La figure I.2 [24] présente une représentation visuelle détaillée de l'anatomie du larynx, mettant en évidence les muscles, les cartilages et les cordes vocales. Elle offre une perspective visuelle supplémentaire pour mieux comprendre l'organisation structurelle du larynx et son rôle dans la production de la parole.

#### Le larynx et cordes vocales

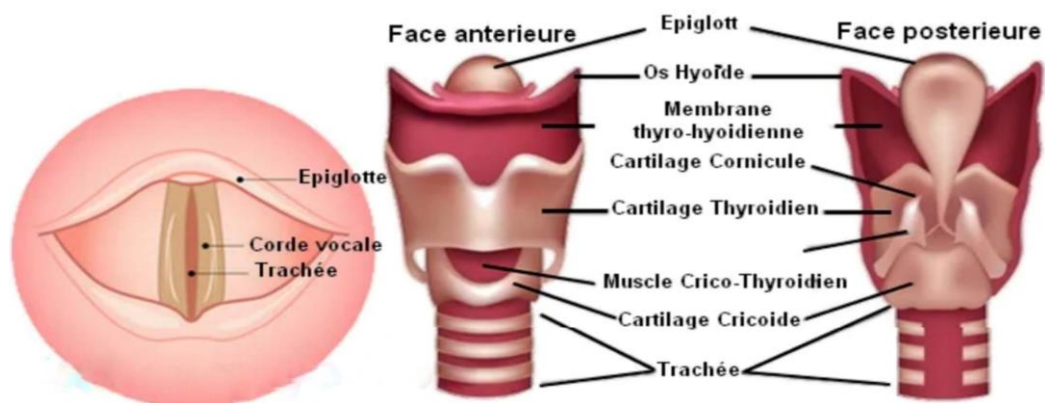


Figure I-2 Anatomie du larynx.

### **I.2.1.3 Le corps sonore**

Les résonateurs du système phonatoire jouent un rôle primordial dans le timbre de la voix. Contrairement aux caisses de résonance des instruments de musique traditionnels, ils se distinguent par leur capacité à modifier rapidement et de manière significative leur forme et leur volume grâce à un réseau musculaire complexe. Cette capacité unique permet à la voix humaine d'offrir une variété acoustique exceptionnelle. Les principaux résonateurs sont le pharynx, la cavité buccale (divisée en deux parties), la cavité labiale et la cavité nasale. Ces résonateurs sont interconnectés par des ouvertures ajustables en taille et tapissés de muqueuses. Contrairement à une croyance répandue, les sinus de la face, quant à eux, sont trop petits pour fonctionner comme des résonateurs et ne jouent aucun rôle reconnu dans la production de la voix [23].

#### **a) Le pharynx**

Le pharynx est une région anatomique située à l'intersection des voies respiratoires et du système digestif. Il joue un rôle crucial dans la déglutition, où les contractions musculaires et les réflexes permettent le passage des liquides et des aliments vers l'œsophage tout en préservant les voies respiratoires. De plus, le pharynx facilite le passage de l'air vers le larynx et les poumons, ce qui le rend également essentiel à la respiration.

#### **b) La cavité buccale**

La cavité buccale joue un rôle crucial dans la production de la parole en fournissant des points d'articulation essentiels. Elle se divise en deux parties : la mâchoire supérieure et la mâchoire inférieure.

Dans la mâchoire supérieure, la voûte palatine se compose de trois zones distinctes : les alvéoles, le palais dur et le palais mou, connu sous le nom de voile du palais. Chacune de ces zones est associée à des adjectifs spécifiques dans la description articulatoire, tels qu'alvéolaire, palatal et vélaire, qui décrivent les points de contact utilisés pour former les sons.

La mâchoire inférieure est principalement contrôlée par la langue, avec la pointe de la langue appelée "apex" et le dos de la langue. Les adjectifs apical et dorsal sont utilisés pour décrire ces parties de la langue et leur rôle dans la production des sons. Les lèvres et les dents contribuent également à la phonation, avec les termes labial et dental qui décrivent leur implication dans la formation des sons. Dans l'ensemble, ces éléments de la cavité buccale

permettent une variété de mouvements et de points d'articulation nécessaires pour produire les différents sons de la parole.

### **c) Les fosses nasales**

Les fosses nasales, également connues sous le nom de cavités nasales, désignent les deux espaces situés à l'intérieur du nez et séparés par une cloison cartilagineuse [25].

## **I.2.2 Le système articulaire**

L'appareil vocal comprend plusieurs organes mobiles essentiels à la production des sons de la parole, tels que le voile du palais, la mâchoire inférieure, la langue et les lèvres. Les mouvements de la mâchoire inférieure jouent un rôle important dans les variations de volume de la bouche. La langue, attachée à l'os hyoïde, est incroyablement mobile grâce à ses dix-sept muscles (huit paires et un impair). Elle s'appuie sur différents points du conduit pharyngo-buccal pour articuler les phonèmes. Ces points d'articulation comprennent :

- Les lèvres (articulations labiales ou bilabiales)
- Les dents (articulations dentales)
- Les alvéoles<sup>16</sup> (articulations alvéolaires)
- Le palais dur, ou partie osseuse de la voûte (articulations palatales)
- Le voile du palais ou « palais mou » (articulations vélares)
- La luette (articulations uvulaires)
- Le pharynx (articulations pharyngales)
- La glotte (articulations glottales)

Les lèvres sont extrêmement mobiles grâce à la présence de treize paires de muscles. Elles jouent un rôle essentiel dans la production des articulations labiales ou bilabiales, où les sons sont formés en rapprochant ou en séparant les lèvres.

Le voile du palais, une membrane qui prolonge le palais dur ou osseux, peut s'abaisser sous l'action des muscles pharyngo-staphylins. Cela permet d'ouvrir la communication entre les voies respiratoires et les fosses nasales, ce qui conduit à la nasalisation des phonèmes émis. Par exemple, les sons "o", "a" et "i" deviennent respectivement "on", "an" et "in". Lorsque le voile du palais se relève, sous l'action des muscles péristaphylins internes, le conduit pharyngo-nasal est fermé, et l'air expiré passe uniquement par la bouche.

Ainsi, les mouvements des lèvres et du voile du palais jouent un rôle crucial dans la modulation et la production des sons lors de la parole [23].

Dans la figure I.3 on peut apercevoir les organes mobiles tels que les lèvres et le voile du palais, ainsi que leurs rôles dans la modulation et la production des sons lors de la parole [26].

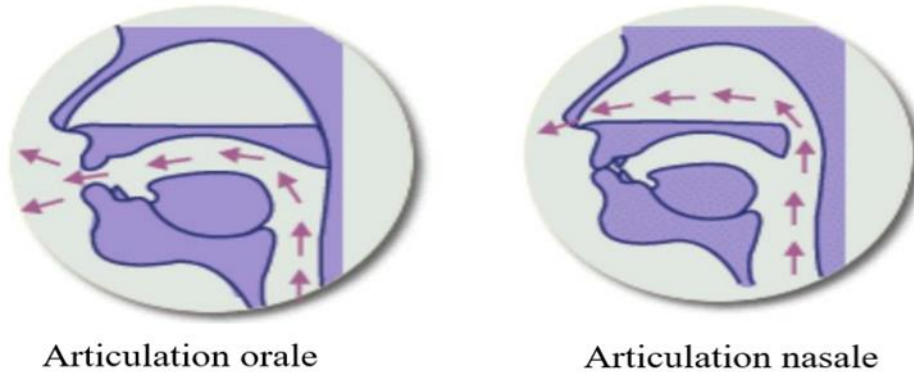


Figure I-3 Processus oro-nasale.

### I.3 Catégories des sons de la parole

#### I.3.1 Classification selon la source

Les sons de la parole peuvent être regroupés en deux catégories en fonction de la manière dont la source d'excitation à la sortie du larynx se comporte. Certains sons ont une nature périodique, ce qui signifie qu'ils présentent une régularité dans leur modèle d'ondes sonores. Ces sons sont appelés "voisés". D'autres sons ont une nature aléatoire, ce qui signifie qu'ils ne suivent pas de modèle régulier dans leurs ondes sonores. Ces sons sont qualifiés de "non voisés". Ainsi, la classification des sons en sons voisés et non voisés est basée sur la façon dont ils sont générés à partir de la source d'excitation dans la production de la parole [27].

##### I.3.1.1 Les sons voisés

Les cordes vocales, lorsqu'elles vibrent, génèrent les sons voisés tels que les voyelles, les semi-voyelles et les consonnes nasales. Les muscles du larynx interagissent avec les cartilages qui soutiennent les cordes vocales afin d'ajuster leur tension, ce qui influence la fréquence des vibrations. Cette fréquence fondamentale, connue sous le nom de F0, représente la fréquence de base des vibrations des cordes vocales lors de l'émission des sons voisés. Ainsi, en modulant la tension des cordes vocales, nous pouvons produire une gamme étendue de fréquences et de nuances dans notre voix [28].

### **I.3.1.2 Les sons non voisés**

Le deuxième mode d'excitation des sons est généré par différents bruits produits lorsque l'air passe à travers des points de rétrécissement du conduit vocal, ou par des bruits d'occlusion ou de plosion résultant de la fermeture ou de l'ouverture des lèvres, ainsi que des contacts de la langue avec le palais. Dans cette catégorie de sons, les cordes vocales ne vibrent pas [28].

Ces conditions spécifiques permettent de produire des sons tels que les consonnes fricatives (comme [s] et [f]), les occlusives (comme [p] et [k]). Ces bruits spécifiques sont formés par des actions précises des organes articulatoires dans le conduit vocal, ce qui contribue à la diversité des sons de la parole.

### **I.3.2 Classification phonétique**

Le phonème est la plus petite unité de son dans la parole, capable de changer le sens d'un mot par sa présence. Le nombre de phonèmes est généralement limité, souvent inférieur à 50. La classification des phonèmes en différentes catégories distinctes, appelées alphabets phonétiques, est essentielle pour caractériser les langues. Les phonèmes sont généralement regroupés en trois catégories principales : les voyelles, les consonnes et les semi-voyelles. Cette classification permet de mieux comprendre et représenter les sons utilisés dans les langues à travers des systèmes de symboles phonétiques spécifiques [27].

#### **I.3.2.1 Les voyelles**

Les voyelles sont des sons produits par un conduit vocal ouvert, sans obstruction, et sont caractérisées par des traits tels que la nasalité, le degré d'ouverture du conduit vocal, la position de la constriction principale et la protrusion des lèvres. Elles peuvent être classées en fonction de la position de la langue, de la forme des lèvres et de l'aperture. Les voyelles occupent le triangle vocalique et peuvent être labialisées ou non labialisées. Elles sont généralement sonores, mais certaines langues utilisent des voyelles sourdes. Elles sont essentielles pour former des syllabes, contrairement aux consonnes qui doivent être associées à une voyelle [27].

### **I.3.2.2 Les consonnes**

Les consonnes sont des sons qui précèdent ou suivent un élément vocalique. Elles sont classées en fonction de leur mode d'articulation, de leur lieu d'articulation, de leur voisement et de leur nasalisation. Les consonnes se divisent en trois principales sous-classes : les fricatives, les occlusives et les sonantes. Les fricatives sont produites par une turbulence aérodynamique due à un rétrécissement dans le conduit vocal, généralement au niveau du palais, des dents ou des lèvres. Les occlusives se caractérisent par une fermeture complète du conduit vocal suivie d'un relâchement brusque, créant une forte pression en amont de l'occlusion. Les occlusives peuvent être voisées ou non voisées, avec ou sans vibration des cordes vocales. Les consonnes sonantes comprennent les liquides (comme [l] et [r]) et les nasales. Les liquides sont difficiles à classer car ils ont une articulation similaire aux voyelles, mais avec une fermeture partielle du conduit vocal. Les consonnes sonantes perdent leur voisement lorsqu'elles sont suivies de consonnes sourdes [27].

### **I.3.2.3 Les semi voyelles**

Les semi-voyelles [j, ɥ, w], parfois appelées semi-consonnes ou glides, combinent certaines caractéristiques des voyelles et des consonnes. Elles peuvent être classées comme des consonnes sonantes. Leur position de base est similaire à celle des voyelles, mais elles se caractérisent par un relâchement soudain de cette position, produisant une friction typique des consonnes. Les fréquences initiales des formants d'une semi-voyelle sont proches de celles de la voyelle correspondante [27].

## **I.4 Modélisation de la production de la parole**

### **I.4.1 Modèle source-filtre**

Le modèle général du système de production de la parole, également connu sous le nom de modèle source-filtre, est une représentation conceptuelle qui décrit les différentes étapes impliquées dans la production des sons de la parole. Ce modèle postule l'existence de deux composantes principales : la source et le filtre. La source représente la source d'énergie acoustique, c'est-à-dire la vibration des cordes vocales, qui génère un signal sonore de base. Le filtre, quant à lui, représente la cavité buccale et les différentes articulations qui modifient le signal sonore en le filtrant et en lui donnant ses caractéristiques spécifiques. Ainsi, le modèle source-filtre permet de simuler comment les actions de la source (les cordes vocales) et du filtre (la cavité buccale) se combinent pour produire les différents sons de la parole. Ce



modèle offre une compréhension approfondie des mécanismes sous-jacents à la production de la parole et constitue une base essentielle pour de nombreuses applications telles que la synthèse vocale, la reconnaissance automatique de la parole et l'étude des troubles de la parole. Dans la figure I.4, nous explorerons plus en détail les éléments constitutifs du modèle source-filtre.

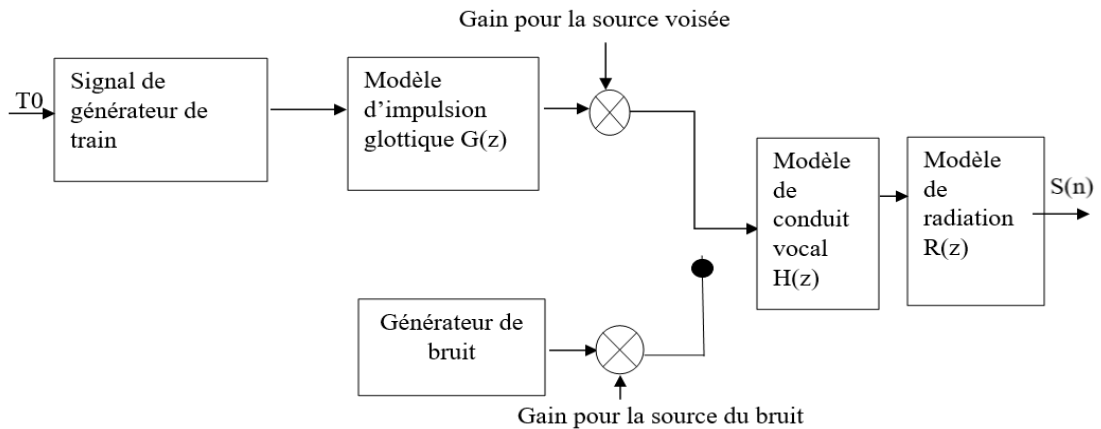
Dans le contexte des voyelles, la source sonore est représentée par un train d'ondes périodique. Ce train d'ondes, défini par la fonction  $G(z)$ , imite la forme du signal produit par les cordes vocales. Sa période, notée  $T_0$ , correspond à la durée d'un cycle complet des cordes vocales.

Le signal d'excitation utilisé pour les sons voisés est :

$$U(n) = \sum_k \delta(n - kT_0) \tag{I.1}$$

$T_0$  Est la durée du cycle vocal

Pour les sons non-voisés, le signal d'excitation  $U(n)$  est caractérisé par un bruit blanc ayant une moyenne nulle et une variance unitaire.



**Figure I-4** Modèle source filtre de la production de la parole.

Une modélisation courante du conduit vocal utilise un filtre tout pôle. Cette approche est basée sur la démonstration selon laquelle un filtre tout pôle peut représenter la fonction de transfert d'un système de tubes acoustiques en série, excités par une onde plane. Cette modélisation approxime de manière raisonnable le comportement du conduit vocal pour les sons non nasaux.

La fonction de transfert d'un filtre tout pôle peut être exprimée par :

$$H(z) = \frac{H_0}{1 + \sum_{k=1}^N b_k z^{-k}} = \frac{H_0}{\prod_{k=1}^P (1 - p_k z^{-1})} \quad (I.2)$$

Où  $H_0$  : représente un gain

$p_k, k = 1, \dots, p$  : représentent les pôles de la fonction transfert

Lorsqu'il s'agit de la parole voisée, la source périodique (provenant des cordes vocales) est représentée par un modèle qui peut varier en fonction de l'application, à condition que ce modèle soit de type "tout pôle".

Un modèle de la source glottique à deux pôles est donné par :

$$G(n) = (\alpha^n - \beta^n)u(n), \quad \beta < \alpha < 1, \quad \alpha \approx 1 \quad (I.3)$$

Où  $u(n)$  représente la séquence échelon unité

La fonction de transfert du rayonnement aux lèvres est modélisée par un filtre proche d'un filtre dérivateur exprimé par :

$$R(z) = 1 - z_0 z^{-1}, \quad z_0 \approx 1, \quad z_0 < 1 \quad (I.4)$$

Pour assurer la stabilité du filtre inverse  $1/R(z)$ , dans certaines applications, la valeur de  $z_0$  est fixée à une valeur légèrement inférieure à 1.

En résumé la transmittance globale entre le train d'impulsions de la figure et le signal émis peut être exprimée par [29] :

$$T(z) = \begin{cases} G(z)H(z)R(z) & \text{Pour les sons voisés} \\ H(z)R(z) & \text{Pour les sons non voisés} \end{cases} \quad (I.5)$$

## I.5 Caractéristique du signal de parole

Le signal de parole est un vecteur acoustique qui contient des informations complexes, variées et redondantes. Ces informations sont représentées par des caractéristiques appelées traits acoustiques. Chaque trait acoustique a une signification perceptuelle spécifique. Parmi ces traits acoustiques, on retrouve :

### **I.5.1 La fréquence fondamentale (pitch)**

La fréquence fondamentale, également appelée "pitch" en anglais, correspond à la hauteur perçue d'un son vocalisé dans la production de la parole. C'est la fréquence dominante qui définit si un son est perçu comme grave ou aigu. Elle est déterminée par la fréquence de vibration des cordes vocales. Les variations de la fréquence fondamentale diffèrent généralement entre les hommes, les femmes et les enfants :

**Chez les hommes**, la fréquence fondamentale est généralement plus basse, variant approximativement de 70 à 250 Hz.

**Chez les femmes**, la fréquence fondamentale est plus élevée que chez les hommes. Les variations typiques vont d'environ 150 à 400 Hz.

**Chez les enfants**, la fréquence fondamentale se situe dans une gamme plus élevée. On observe couramment des variations de l'ordre de 200 à 600 Hz.

### **I.5.2 Spectre fréquentiel**

Le spectre fréquentiel joue un rôle crucial dans la perception du timbre de la voix. Le timbre est une caractéristique distinctive qui permet d'identifier une personne uniquement en écoutant sa voix. Il dépend de la corrélation entre la fréquence fondamentale et les harmoniques, qui sont les multiples de cette fréquence. En d'autres termes, le timbre est façonné par la structure spectrale du signal vocal, où les harmoniques apportent des caractéristiques particulières à la voix d'une personne. Cette corrélation unique entre la fréquence fondamentale et les harmoniques contribue à la singularité et à la reconnaissance individuelle des voix [30].

### **I.5.3 L'énergie**

L'énergie du signal de parole correspond à son intensité sonore et mesure la quantité d'énergie acoustique qu'il contient. Elle varie en fonction des différents segments de la parole. Les segments "voisés", associés aux voyelles et à certaines consonnes, ont généralement une énergie plus élevée en raison de l'activation des cordes vocales et de la production de sons mélodiques. En revanche, les segments "non voisés", tels que les consonnes plosives ou fricatives, ont une énergie plus faible car ils ne dépendent pas d'une

activation continue des cordes vocales et sont plus bruités. Ainsi, l'énergie du signal de parole varie en fonction des composantes sonores et des types de sons produits [30].

#### **I.5.4 Les formants**

Le son que nous percevons dans la parole est différent de celui produit par les cordes vocales. Ce que nous entendons est le résultat d'un processus de filtrage appliqué à une onde complexe. Les cavités supra-glottiques, situées au-dessus des cordes vocales, agissent comme des résonateurs et ont la capacité d'amplifier ou d'atténuer certaines fréquences. La forme, la taille et la composition de ces cavités déterminent les fréquences qui seront mises en évidence et celles qui seront atténuées. Cela se manifeste par un changement de configuration lors de la prononciation de différentes voyelles, ce qui crée des variations de timbre. Le conduit vocal agit comme un filtre qui transforme la source sonore en phonèmes, et les formants, qui sont les pics de la fonction de transfert du filtre, représentent les zones de renforcement maximal. En pratique, seuls quelques formants jouent un rôle central dans la perception de la parole. Ainsi, les voyelles sont souvent décrites en utilisant les deux premiers formants F1 et F2, qui peuvent être représentés par un triangle acoustique [31].

### **I.6 L'analyse de la parole pour l'évaluation des troubles de la voix**

L'analyse de la parole joue un rôle essentiel dans l'évaluation des troubles de la voix, offrant des informations précieuses pour le diagnostic et le suivi des patients. Elle permet d'identifier et de caractériser les altérations de la voix, qu'elles soient d'origine fonctionnelle ou organique. Grâce à l'avancement des techniques d'analyse et de traitement du signal, de nombreuses méthodes ont été développées pour extraire des mesures objectives et quantitatives à partir de l'échantillon vocal.

#### **I.6.1 Les troubles de la parole et de la voix**

Les troubles de la parole et de la voix sont des altérations qui affectent la communication orale. Ils peuvent se manifester par des difficultés à produire les sons, à articuler correctement les mots ou à moduler la voix. Ces troubles peuvent avoir diverses causes, notamment des problèmes neurologiques, structurels ou fonctionnels. Ils peuvent entraîner des difficultés à communiquer efficacement, tant sur le plan professionnel que personnel. Ces troubles de la parole englobent diverses altérations de la production verbale [32].

Ces troubles peuvent être classés en deux grandes catégories : les troubles organiques et les troubles fonctionnels/neurologiques

### **I.6.1.1 Troubles organiques**

Les dysfonctionnements organiques de la voix sont généralement causés par des altérations morphologiques de l'anatomie du larynx, en particulier au niveau de la glotte. Parmi les troubles vocaux organiques courants, on retrouve les nodules, les kystes et les laryngites [33].

### **I.6.1.2 Troubles fonctionnelles**

Les dysphonies fonctionnelles sont des troubles de la voix qui résultent d'une mauvaise utilisation du système vocal, bien que celui-ci soit anatomiquement et physiologiquement normal. Parmi ces dysphonies on trouve :

La paralysie du nerf laryngé récurrent qui entraîne une immobilité partielle ou complète d'une ou des deux cordes vocales, ce qui affecte la phonation. Le nerf laryngé récurrent innerve tous les muscles intrinsèques du larynx, à l'exception du muscle cricothyroïdien qui est innervé par le nerf laryngé supérieur. Il existe deux types de paralysie : temporaire et dégénération des fibres nerveuses.

Dans le cas de la paralysie temporaire, une fois la cause sous-jacente éliminée, l'activité musculaire peut reprendre. Cependant, dans la paralysie due à la dégénération des fibres nerveuses, les muscles restent inactifs à moins que la régénération des fibres ne soit complète. La paralysie peut être complète, où toutes les unités motrices sont paralysées, ou partielle, avec certaines unités motrices intactes et d'autres paralysées.

La phonation pendant la paralysie du nerf laryngé récurrent dépend du degré de paralysie. Les symptômes comprennent une diminution de la tension des muscles du larynx, ce qui peut entraîner une voix faible et voilée. La paralysie laryngée se manifeste par une diminution de la dureté des muscles du larynx.

La gestion de la paralysie laryngée peut nécessiter des interventions médicales ou chirurgicales, selon la cause et la gravité de la paralysie. Il est important d'évaluer attentivement la paralysie du nerf laryngé récurrent afin de déterminer le meilleur plan de traitement pour améliorer la fonction vocale et la qualité de vie du patient [34].

### **I.6.1.3 Troubles neurologiques**

Les troubles neurologiques de la voix sont des altérations de la production vocale qui résultent de problèmes au niveau du système nerveux. Ces troubles peuvent être causés par des lésions, des dysfonctionnements ou des anomalies dans les voies nerveuses qui contrôlent les muscles et les fonctions vocales, parmi ces troubles on trouve la dysarthrie.

#### **(i) Définition de la dysarthrie**

Selon l'American Speech and Hearing Association (ASHA), les troubles de la communication englobent les altérations de la réception, de l'émission, du traitement et de la compréhension des concepts et des systèmes de symboles verbaux, non verbaux et graphiques. Parmi ces troubles, on trouve les troubles de la voix, du langage, de l'audition et de la parole.

La dysarthrie, en tant que trouble de la parole, est considérée comme un trouble de communication. Initialement, la dysarthrie était décrite comme un trouble articulaire, mais on a reconnu l'importance de facteurs neurophysiologiques, psychologiques, phonétiques et pathologiques. Bien que le terme "dysarthro-pneumo-phonie" ait été proposé pour décrire les altérations non articulaires, il n'est pas couramment utilisé. En 1975, Darley a défini la dysarthrie comme un trouble moteur de la parole causé par des lésions du système nerveux central et/ou périphérique. Actuellement, le terme "dysarthrie" englobe les troubles moteurs de la parole d'origine neurologique acquis, à l'exception de l'apraxie, et exclut les troubles mécaniques qui peuvent également affecter la parole. Les dysarthries sont variées et entraînent différentes altérations et perturbations, ce qui a conduit à plusieurs classifications pour mieux les comprendre [35].

#### **(ii) Classification de la dysarthrie**

La classification de Darley (Darley et al., 1969b, a) est l'une des plus utilisées pour classer les dysarthries. Elle repose sur des caractéristiques perceptives observées dans la parole dysarthrique et a été établie à partir d'une étude perceptive impliquant l'évaluation de 38 paramètres de la parole chez 212 patients. Les résultats de l'étude ont révélé huit classes dysarthriques, basés sur des corrélations significatives et physiologiquement pertinentes entre les critères évalués. Les six classes de dysarthrie identifiées sont les suivantes :

**Dysarthrie flasque** : causée par des atteintes des motoneurones périphériques, de la jonction neuromusculaire ou des muscles impliqués dans la production de la parole. Elle peut résulter d'atteintes au niveau de la moelle épinière, du bulbe rachidien, des nerfs crâniens ou des nerfs spinaux.

**Dysarthrie spastique** : causée par des lésions bilatérales des voies reliant les structures hémisphériques aux noyaux du tronc cérébral contrôlant les effecteurs de la parole. Elle est souvent associée au syndrome pseudo-bulbaire.

**Dysarthrie ataxique** : causée par une atteinte du cervelet ou des voies cérébelleuses. Les causes peuvent être dégénératives, vasculaires, démyélinisantes, traumatiques, néoplasiques, inflammatoires, toxiques ou métaboliques.

**Dysarthrie hypokinétique** : causée par une atteinte des noyaux gris centraux du système nerveux, souvent associée à la maladie de Parkinson.

**Dysarthrie hyperkinétique** : causée par une atteinte des noyaux gris centraux et caractérisée par des mouvements involontaires, tels que dans la maladie de Huntington.

**Dysarthries mixtes** : impliquant une combinaison de composantes centrales et périphériques, souvent observée chez les patients atteints de sclérose latérale amyotrophique (SLA).

Plus tard, en 2005, Duffy a ajouté deux autres classes à la classification de Darley : les dysarthries liées à une atteinte unilatérale du premier neurone moteur et les dysarthries d'étiologie indéterminée.

Cette classification des dysarthries permet de regrouper les patients en fonction des caractéristiques de leur parole et des mécanismes physiopathologiques supposés sous-jacent [35].

### **(iii) La dysarthrie dans la maladie de Parkinson**

La maladie de Parkinson est une affection dégénérative du système nerveux central qui se caractérise par la diminution progressive des neurones dopaminergiques dans les noyaux gris centraux. Cette dénervation dopaminergique nigrostriatale entraîne un dysfonctionnement chronique des circuits neuronaux responsables du contrôle des mouvements, ce qui se traduit par des troubles moteurs, dont les troubles de la parole sont fréquents. Les patients atteints de la maladie de Parkinson peuvent présenter une dysarthrie hypokinétique, marquée par une réduction des mouvements articulatoires, une modulation

prosodique altérée et une parole monotone. Cette altération de la production de parole a un impact significatif sur l'autonomie et la qualité de vie des patients, ainsi que sur leur capacité à communiquer efficacement. L'évaluation des troubles de la parole dans la maladie de Parkinson repose principalement sur des méthodes d'analyse perceptives et instrumentales. Les échelles d'évaluation, telles que l'échelle GRBAS-I, permettent d'évaluer la dysphonie, la dysprosodie, la dysarthrie et l'intelligibilité de la parole en prenant en compte les différentes fonctions impliquées. Une approche pluridisciplinaire, impliquant des professionnels de la santé tels que des neurologues, des orthophonistes et des thérapeutes du langage, est nécessaire pour évaluer de manière exhaustive les troubles de la parole et proposer des interventions thérapeutiques personnalisées visant à améliorer la communication et la qualité de vie des patients parkinsoniens [36].

#### **(iv) Les éléments de l'évaluation de la dysarthrie de la maladie de Parkinson**

La dysphonie parkinsonienne se réfère aux anomalies de fonctionnement du larynx, affectant la hauteur, l'intensité et le timbre de la voix. Plusieurs études ont été menées pour examiner ces anomalies en détail (Robert et Spezza, 2005). En ce qui concerne la hauteur, les résultats des analyses perceptives et acoustiques sont contrastés. Cependant, la majorité des études rapportent une augmentation de la fréquence fondamentale (F0) moyenne chez les patients parkinsoniens lors de la tenue d'une voyelle, de la lecture ou de la parole spontanée. Il est important de noter que ces résultats peuvent être influencés par des facteurs tels que l'âge, le sexe, la durée de la maladie, les traitements antiparkinsoniens et la variabilité interindividuelle et intra-individuelle, ainsi que par les différentes méthodes d'évaluation utilisées.

En ce qui concerne l'intensité, les résultats des analyses perceptives et acoustiques montrent également une légère réduction de l'intensité moyenne. Cette réduction peut être attribuée à une incompétence phonatoire due à une diminution de la pression sous-glottique.

Les anomalies du timbre de la voix, telles que l'aspect soufflé ou éraillé, ainsi que le tremblement vocal, sont constamment observés dans les études perceptives. Les mesures acoustiques de voyelles tenues ont permis de mettre en évidence une augmentation des variations de la F0 (jitter) et de l'intensité (shimmer) d'un cycle à l'autre, ainsi que des variations à plus long terme liées au tremblement, entraînant une réduction du rapport



signal/bruit. Les études en laryngoscopie ont également révélé des anomalies au niveau du larynx, telles qu'un défaut d'accolement des cordes vocales, une possible hypertonie des bandes ventriculaires et un tremblement au niveau des cordes vocales ou du tractus vocal sus-glottique (Jiang et al., 1999 ; Yurceturk et al., 2002).

La dysprosodie parkinsonienne, caractérisée par une insuffisance prosodique, est probablement le signe le plus spécifique des troubles de la parole dans la maladie de Parkinson. Selon la description de la Mayo Clinic (Darley et al., 1975), 6 des 10 dimensions les plus déviantes de la parole parkinsonienne relèvent de la dysprosodie, notamment la monotonie de hauteur et d'intensité, la réduction de l'accentuation, les pauses inappropriées, les accélérations brèves et les variations de débit.

La monotonie de hauteur se manifeste par une réduction de la dynamique de fréquence et de la gamme tonale, ce qui altère la mélodie de la voix avec une diminution des contrastes de hauteur. La monotonie de l'intensité est également souvent rapportée, bien qu'elle soit plus difficile à objectiver à l'aide d'analyses acoustiques instrumentales.

En ce qui concerne le rythme, le débit de parole, la durée des pauses et d'autres aspects temporels, tels que la festination de la parole, la palilalie (répétition involontaire de syllabes ou de mots) et les pseudo-bégaiements, sont des anomalies couramment décrites dans la maladie de Parkinson. Des études ont montré que le débit de parole a tendance à être plus lent chez les patients parkinsoniens, avec des pauses plus longues et des interruptions à l'intérieur des mots et des syntagmes. Ces disfluences et anomalies rythmiques perturbent la cohésion des unités linguistiques et peuvent réduire l'intelligibilité de la parole.

En ce qui concerne les troubles articulatoires, mis à part l'imprécision générale des consonnes, les anomalies articulatoires dans la maladie de Parkinson se concentrent principalement sur les consonnes occlusives. Ces consonnes peuvent être perçues comme des consonnes constrictives, indiquant une fermeture insuffisante du conduit vocal. Les analyses acoustiques révèlent un bruit de friction de faible intensité plutôt qu'un silence normalement associé à l'occlusion des consonnes. Ce phénomène est appelé spirantisation et son intensité est corrélée au degré de dysarthrie chez les patients parkinsoniens. Bien qu'il puisse y avoir une tendance à la rhinolalie chez certains patients, en général, il n'y a pas de

différence significative entre les groupes de patients parkinsoniens et les sujets témoins en ce qui concerne la rhinolalie.

En ce qui concerne les voyelles, la réduction de la surface du triangle vocalique chez les patients parkinsoniens peut contribuer à la dédifférenciation des voyelles, en plus des phénomènes de spirantisation et de coarticulation, ce qui entraîne une réduction des contrastes acoustiques dans la parole parkinsonienne [37].

## **I.6.2 Évaluation des troubles de la voix et de la parole**

L'évaluation des troubles de la voix et de la parole peut être réalisée à l'aide de deux approches complémentaires : les méthodes subjectives et les méthodes objectives. Les méthodes objectives se basent sur des mesures et des analyses quantitatives de paramètres acoustiques, physiologiques et perceptuels de la voix et de la parole. Elles fournissent des données précises et objectives sur les caractéristiques vocales et articulatoires des individus.

D'autre part, les méthodes subjectives font appel à des évaluations perceptuelles et à des jugements qualitatifs réalisés par des experts ou des observateurs formés. Ces méthodes prennent en compte l'appréciation globale de la qualité vocale et de l'articulation, ainsi que l'impact de ces troubles sur la communication et la qualité de vie des patients.

### **I.6.2.1 Évaluation subjective**

L'évaluation subjective de la voix peut être réalisée de deux manières : par le praticien à travers une analyse perceptive et par le patient lui-même à travers une auto-évaluation par questionnaire.

L'analyse perceptive permet au praticien de juger de la qualité de la voix et de ses caractéristiques perceptives en utilisant différentes échelles, dont l'échelle GRBASI développée par Hirano (1981) et complétée par Dejonckere (1996).

Chaque paramètre est évalué par le praticien selon une échelle de sévérité comprenant quatre niveaux, allant de 0 (normal / absence) à 3 (présence maximale). Ainsi, un score GRBASI plus élevé indique une altération vocale jugée plus importante par le professionnel.

L'évaluation perceptive est indispensable pour juger de la qualité de la voix, mais elle présente des biais liés à la variabilité inter-évaluateur et intra-évaluateur. Les évaluations instrumentales "objectives" sont complémentaires et compensent cette importante subjectivité [38].

### **I.6.2.2 Evaluation objective**

L'évaluation objective se base sur les techniques de traitement automatique de la parole (TAP) afin d'extraire des indices acoustiques pertinents permettant de déterminer les caractéristiques de la voix et de fournir des informations sur l'état du larynx du locuteur. Les cliniciens privilégient les indices acoustiques qui sont corrélés aux caractéristiques perceptuelles de la voix. Cependant, il est important de souligner que l'accord entre les mesures acoustiques et les évaluations perceptuelles peut être limité lorsque l'évaluation perceptive manque de fiabilité. Il existe différentes méthodes d'évaluation objective de la voix pathologique [34]

#### **(i). Indice de perturbation d'amplitude**

- **Shimmer**

Le shimmer est un indice de perturbation de l'amplitude de la voix. Il mesure les variations de l'amplitude d'un cycle vocal à l'autre, fournissant des informations sur l'irrégularité et la stabilité de la voix. C'est un paramètre utilisé pour évaluer la qualité de la voix et détecter d'éventuels problèmes vocaux. On distingue quatre mesures liées au shimmer fournies par le logiciel Praat d'analyse de la parole [39].

- ✓ *Shimmer absolu (ShdB)* : est calculé en prenant le logarithme absolu moyen en base 10 de la différence entre les amplitudes des périodes vocales consécutives. Ce résultat est ensuite multiplié par 20 et exprimé en décibels (dB). Il permet de quantifier les variations d'amplitude entre les cycles vocaux.
- ✓ *Shimmer local (Shim)* : est calculé en prenant la différence absolue moyenne entre les amplitudes des périodes vocales consécutives. Cette valeur est ensuite divisée par l'amplitude moyenne et exprimée en pourcentage. Le shimmer local permet de quantifier les variations d'amplitude relatives entre les cycles vocaux.
- ✓ *L'Amplitude à trois points Quotient de perturbation (apq3)* : Il représente la moyenne des différences d'amplitude entre les valeurs de fréquence fondamentale à trois points spécifiques d'un cycle vocal.
- ✓ *L'amplitude à cinq points Quotient de perturbation (apq5)* : Il est calculé en prenant la différence absolue moyenne entre les amplitudes de cinq points équidistants le

long d'une période vocale, divisée par l'amplitude moyenne et exprimée en pourcentage.

Dans le cadre de notre étude, nous avons utilisé le shimmer local (Shim).

## **(ii). Indices de perturbation des durées des cycles**

### **Jitter**

Le jitter est un indice acoustique qui mesure les variations des durées des cycles (période fondamentale) de la voix. Il évalue la stabilité temporelle de la voix en calculant la différence moyenne entre les intervalles de temps des cycles de vibration des cordes vocales. Le jitter est exprimé en pourcentage et permet de détecter des variations anormales de la fréquence de la voix.

Plusieurs variantes du jitter sont disponibles dans le logiciel Praat. On trouve

- ✓ *Le jitter absolu (jitta)* représente la différence moyenne en temps entre les périodes consécutives du signal de la voix, exprimé en seconde.
- ✓ *Le jitter local ou relatif (jitt)* est calculé en prenant la différence absolue moyenne entre les périodes consécutives, puis en la divisant par la période moyenne, exprimée en pourcentage.
- ✓ *La perturbation moyenne relative (RAP)* est un indice qui mesure la variabilité relative entre les périodes consécutives d'un signal vocal. Il est calculé en prenant la différence absolue moyenne entre les périodes consécutives, puis en la divisant par la période moyenne, exprimée en pourcentage.
- ✓ *La perturbation de période à cinq points (PPQ5)* est obtenue en calculant la moyenne de la différence absolue entre les périodes consécutives, divisée par la période moyenne et multipliée par 100.

## Tremor

Le tremor vocal est une perturbation rythmique de la voix qui désigne les fluctuations involontaires basses fréquences des durées des cycles dont la bande de fréquence est située entre 2 et 10 Hz. Dans le cadre de ce travail, le tremor vocal est obtenu comme suit :

1. Calculer la série des durées des cycles en utilisant Praat
2. Interpoler de la série des durées des cycles
3. Soustraire la moyenne de la série des durées des cycles interpolés
4. Estimer le tremor vocal par filtrage entre 2 et 10 Hz de la série obtenue dans (3) au moyen de filtres de Savitzky-Golay.

### (iii). Indices spectraux et cepstraux

#### Rapport harmonique sur bruit

Le rapport harmonique sur bruit (HNR : harmonics-to-noise ratio) est un indice utilisé pour évaluer la proportion de composantes harmoniques par rapport au bruit dans un signal vocal. Un HNR élevé indique une plus grande proportion de composantes harmoniques, ce qui est associé à une voix claire et stable. En revanche, un HNR faible suggère une présence accrue de bruit et moins de composantes harmoniques, pouvant indiquer une voix perturbée ou des problèmes vocaux.

L'équation générale pour le calcul du rapport harmonique sur bruit (HNR) par le logiciel Praat est la suivante :

$$HNR = 10 \log_{10} \left( \frac{AC_v(T)}{AC_v(0) - AC_v(T)} \right) \quad (I.6)$$

$AC_v(0)$  Représente le coefficient d'autocorrélation.

$AC_v(T)$  Représente la composante d'autocorrélation correspondante à la période du pitch.

La différence  $AC_v(0) - AC_v(T)$  est l'énergie du bruit.

#### Proéminence du pic cepstral

Le cepstre est défini comme étant la transformée de Fourier inverse du logarithme du spectre d'amplitude [40]. Les composantes du cepstre sont appelées les queffrences. Pour les segments voisés, la région de hautes queffrences présente des pics cepstraux également

espacés (les harmoniques) provenant de la structure harmonique du logarithme du spectre d'amplitude. Des analyses cepstrales ont été utilisées pour évaluer la régularité du signal vocal.

La proéminence du pic cepstral (CPP : cepstral peak proeminence) est un indice acoustique utilisé pour l'évaluation de la qualité de la voix. Le calcul du CPP comprend les étapes suivantes :

1. Le spectre de puissance logarithmique est calculé pour chaque trame pondérée par la fenêtre de Hamming avec un chevauchement de 50%.
2. Adaptation d'une droite de régression linéaire au log-cepstre entre 1 ms et la quefrece maximale.
3. La hauteur du pic cepstral le plus proéminent par rapport à la ligne de régression est obtenue entre le minimum et le maximum attendu des quefrences vocales. Ceci est le pic cepstral local (par trame).
4. La proéminence du pic cepstral (CPP) globale est obtenue en faisant la moyenne des proéminences des pics cepstraux locales sur toutes les trames d'analyse.

Dans ce travail, nous avons utilisé le plugin CPP\_voiceDetection avec le logiciel Praat, un outil largement utilisé dans l'analyse acoustique [41].

Dans le cadre de ce travail, nous avons utilisé le plugin CPP\_voiceDetection avec le logiciel Praat, un outil largement utilisé dans l'analyse acoustique [41].

## **I.7 Conclusion**

En conclusion, nous avons exploré les différents éléments de la production de la parole, tels que l'appareil phonatoire, la classification des sons, la modélisation de la production de la parole et l'évaluation des troubles de la voix et de la parole. La coordination des éléments de l'appareil phonatoire est essentielle pour générer des signaux acoustiques de parole, et la classification des sons nous permet de mieux comprendre la diversité des sons produits. Le modèle source-filtre permet de modéliser la production de la parole, tandis que l'analyse des caractéristiques du signal de la parole permet d'évaluer les propriétés acoustiques. Nous avons identifié différentes catégories de troubles et les approches d'évaluation subjective et objective. Ces notions sont nécessaires pour développer des méthodes appropriées pour la détection et l'évaluation des troubles de la parole et de la voix.

---

# **Chapitre II**

**Systemes de télémédecine**

---



## **II.1 Introduction**

En raison des progrès technologiques et médicaux, l'espérance de vie humaine s'accroît, engendrant une pression accrue sur les systèmes de santé mondiaux, sollicités pour prodiguer des soins de qualité supérieure à un nombre croissant de patients. Cette demande croissante de services de santé incite les chercheurs à explorer de nouvelles modalités d'organisation et de prestation des soins. La télémédecine, en tant que champ de recherche émergent, favorise l'utilisation de systèmes multimédias afin d'accroître la disponibilité des soins pour les patients, tout en mettant en œuvre des stratégies d'économie de coûts et de temps. En d'autres termes, la télémédecine permet aux patients d'être examinés et traités alors que le professionnel de santé et le patient se trouvent à des emplacements distincts sur le plan physique. Grâce aux avancées technologiques de la télémédecine, les hôpitaux du futur seront en mesure de dispenser des services de santé à des patients du monde entier, en recourant à des systèmes multimédias et à la transmission de signaux à distance. La transmission, le stockage et le traitement des signaux et des images constituent les principaux piliers de la télémédecine [42].

Dans ce chapitre on va se concentrer sur les systèmes de télémédecine et leur utilisation dans le domaine des troubles de la parole et de la voix. Les troubles de la parole et de la voix peuvent avoir un impact significatif sur la qualité de vie des individus, affectant leur capacité à communiquer efficacement. On donne une définition des systèmes de télémédecine et leurs applications. Nous examinerons les actes médicaux réalisables à distance, les avantages et les défis de la télémédecine, ainsi que l'importance de la compression des signaux de parole pour une transmission efficace des données.

## **II.2 Définition de la télémédecine**

La télémédecine est un service lié à la santé qui utilise les technologies de télécommunication et de l'information électronique. Elle englobe l'ensemble des prestations conçues pour permettre aux patients et à leurs médecins ou prestataires de soins de santé d'interagir. Elle offre une large gamme d'utilisations, notamment des consultations en ligne, le contrôle à distance, les soins infirmiers à distance et la réadaptation physique et psychiatrique à distance. La télémédecine permet de faire de meilleurs choix en matière de soins de santé, d'améliorer la qualité et les performances des services d'urgence, de réduire le temps nécessaire pour établir un diagnostic et de réaliser des économies de coûts pour les

médecins et les patients en optimisant les procédures cliniques et en réduisant les frais de déplacement vers les hôpitaux.

La télémédecine a accru l'accès à des services de santé de haute qualité. Les patients ont désormais accès à des services cliniques plus personnalisés. Ils peuvent également consulter les meilleurs prestataires de soins simplement en utilisant des logiciels d'application vidéo, et les consultations peuvent être effectuées à distance. Les cliniciens disposent d'outils mieux adaptés pour la mise en réseau, le stockage des données, la gestion des rapports et la mise en valeur des compétences spécifiques de chacun, ce qui améliore la qualité de la pratique médicale. Les médecins passent moins de temps sur des affectations en milieu rural et peuvent fournir davantage de soins aux patients. La télémédecine permet également aux spécialistes des soins de santé privés de pratiquer et améliorer l'expérience des patients. Ces derniers n'ont plus à faire la queue pendant de longues périodes, et les médecins peuvent accéder aux informations des patients de manière plus pratique et efficace grâce aux dossiers électroniques, éliminant ainsi les temps d'attente globaux. De plus, les rendez-vous à distance permettent aux médecins de consacrer moins de temps à chaque patient, ce qui leur permet de traiter un plus grand nombre de patients [43].

### **II.2.1 La configuration typique d'un système de télémédecine**

Dans un système de télémédecine, il existe généralement trois composants principaux :

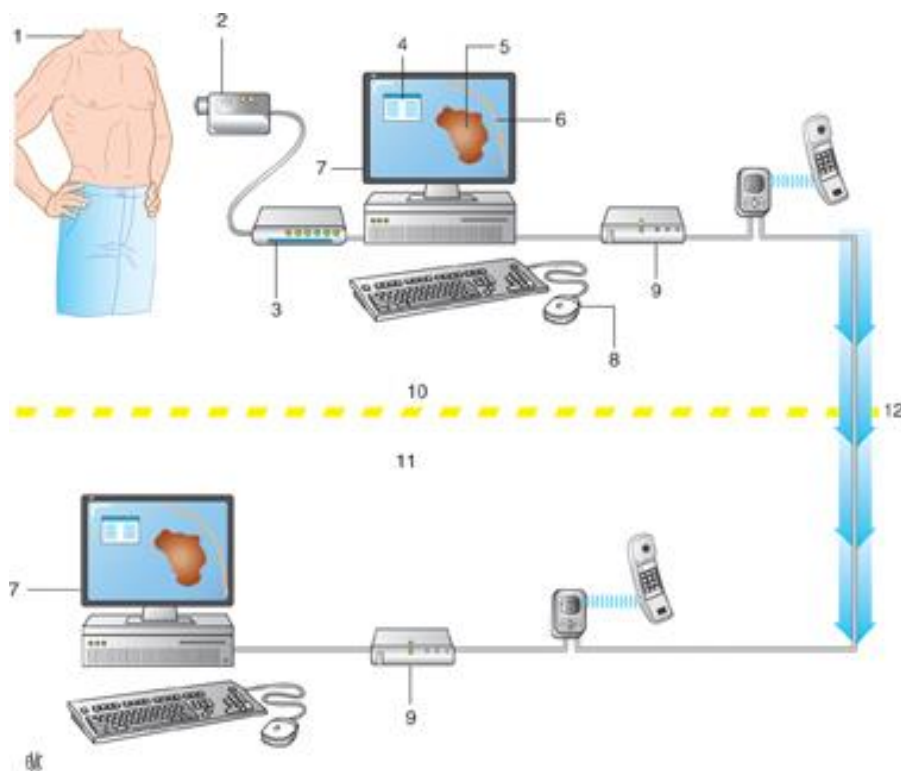
**Le patient :** Il s'agit de la personne qui bénéficie des services de télémédecine. Le patient peut être situé à distance de l'établissement de santé et interagit avec les professionnels de la santé via des technologies de communication.

**Le professionnel de la santé :** Il peut s'agir d'un médecin, d'un infirmier, d'un spécialiste ou d'un autre prestataire de soins de santé. Le professionnel de la santé est généralement situé dans un établissement de santé, mais il peut également être mobile ou travailler à distance.

**Le système de télémédecine :** Il comprend les équipements, les logiciels et les réseaux nécessaires pour faciliter la communication à distance entre le patient et le professionnel de la santé. Cela peut inclure des éléments tels que des caméras vidéo, des microphones, des haut-parleurs, des dispositifs de mesure à distance, des dossiers médicaux électroniques et des plateformes de communication sécurisées.

L'interaction entre le patient et le professionnel de la santé se fait généralement via une connexion Internet sécurisée. Le patient peut être équipé d'un dispositif de télémédecine, tel qu'un ordinateur, une tablette ou un smartphone, qui lui permet de communiquer avec le professionnel de la santé à distance.

Il est important de noter que la configuration précise d'un système de télémédecine peut varier en fonction des besoins et des technologies disponibles dans chaque situation spécifique. La figure II.1 [44], montre la configuration d'un système de télémédecine.



**Figure II-1** Schématisation d'un système de télé-expertise. 1. Sujet ; 2. caméra CCD couleur et éclairage ; 3. dispositif de commande de caméra ; 4. fenêtre d'information du patient et de l'image ; 5. lésion étudiée ; 6. peau environnante ; 7. micro-ordinateur

### II.3 Pourquoi le système de santé a-t-il besoin de la télémédecine?

La télémédecine est essentielle au système de santé pour plusieurs raisons. La hausse des coûts des soins de santé et la nécessité d'améliorer les traitements motivent de plus en plus d'hôpitaux à explorer les avantages de la télémédecine. Ils cherchent à favoriser un meilleur

contact entre les médecins et les patients éloignés et à optimiser l'utilisation des ressources médicales. La télémédecine permet également une meilleure connectivité, réduisant ainsi les réadmissions à l'hôpital et favorisant l'observance des plans de traitement prescrits. Elle facilite également la communication entre les médecins, leur permettant de former des réseaux de soutien et d'échanger leurs compétences pour fournir des services de santé de meilleure qualité. En utilisant des technologies telles que les visioconférences, la télémédecine améliore l'accessibilité des soins médicaux, tant pour les patients que pour les prestataires de soins de santé. Bien qu'il reste des défis techniques et des critiques à relever, la télémédecine complète et améliore l'expérience globale des patients [43].

## **II.4 Les actes de la télémédecine**

Pour apporter plus de clarté quant aux responsabilités médicales impliquées, il est crucial de simplifier la définition des actes médicaux dans le cadre de la télémédecine en prenant en compte les définitions adoptées par tous les pays, notamment les pays européens. Ces définitions viennent compléter celles déjà fournies par le Ministère de la Santé. Ainsi, la télémédecine se concentre principalement sur trois actes essentiels : la téléconsultation, la télé expertise et la télésurveillance [45].

### **II.4.1 La télé consultation**

La téléconsultation offre aux patients la possibilité d'accéder directement à distance à des consultations médicales ou paramédicales. Cela peut se faire via l'utilisation du téléphone, comme c'est le cas aux États-Unis, ou en utilisant une salle de consultation virtuelle. Il est également possible d'accompagner la téléconsultation d'une télémanipulation, par exemple pour réaliser une échographie à distance. L'objectif principal de la téléconsultation est d'établir un diagnostic et de permettre au professionnel médical de prescrire un traitement approprié. Dans certains cas, un professionnel de santé peut être présent aux côtés du patient pour assister le professionnel médical pendant la téléconsultation [46].

### **II.4.2 La télé surveillance**

La télésurveillance médicale consiste à surveiller à distance un patient, étendant ainsi le principe de surveillance observé en réanimation à un patient éloigné, notamment à son domicile. Un exemple concret est la surveillance de la dialyse à domicile pour les patients atteints d'insuffisance rénale chronique. Pour cela, il est nécessaire que le patient dispose

d'un équipement abordable permettant l'enregistrement et la transmission des données nécessaires au suivi médical. Ces données sont ensuite interprétées à distance par un professionnel médical, qui peut prendre des décisions concernant la prise en charge du patient. La télésurveillance médicale peut être automatisée ou réalisée par le patient lui-même, voire par un professionnel de santé [46].

### **II.4.3 La télé expertise**

La télé expertise a souvent été définie de manière limitée comme étant les échanges entre spécialistes pour obtenir un deuxième avis. Cependant, il serait préférable, par souci de simplification, d'élargir cette définition à tout acte diagnostique et/ou thérapeutique réalisé en l'absence du patient. Ainsi, l'acte médical de télé expertise peut-être décrit comme un échange entre deux ou plusieurs médecins qui collaborent pour établir un diagnostic et/ou un traitement en se basant sur les données cliniques, radiologiques ou biologiques présentes dans le dossier médical du patient [45].

### **II.4.4 La télé radiologie**

La télé radiologie peut être définie comme l'utilisation des technologies de l'information et de la communication pour transmettre à distance des images radiologiques et permettre leur interprétation par des radiologues éloignés du lieu où les images sont acquises. Cela facilite la collaboration entre les professionnels de la santé et permet d'améliorer l'accès aux services de radiologie dans les régions éloignées ou mal desservies [43].

### **II.4.5 La télé assistance**

La télé assistance médicale se réfère à la pratique où un médecin fournit une assistance à distance à un autre médecin ou à un professionnel de santé lors de la réalisation d'actes médicaux ou chirurgicaux. Elle peut également être utilisée dans des situations d'urgence pour assister à distance un secouriste ou toute personne apportant une aide médicale à une personne en danger, en attendant l'arrivée d'un médecin. Ainsi, la téléassistance médicale englobe la collaboration à distance entre professionnels de la santé pour optimiser les soins et assurer une assistance médicale efficace [45].

### II.4.6 La télé chirurgie

Manipulation de matériel médical (instruments chirurgicaux) contrôlée à distance par le praticien sur le patient (appelée aussi télémanipulation) [47].

### II.4.7 La télé formation

L'utilisation de l'informatique, en particulier pour la formation continue des médecins, joue un rôle essentiel. Cela comprend des aspects tels que les contacts professionnels via le réseau, l'accès à des informations médicales telles que les banques de données et l'imagerie médicale, le suivi des études épidémiologiques et des essais cliniques, la consultation de cours de formation et la participation à des visioconférences dans le cadre de l'enseignement à distance et des réunions.

Un aspect important de cette utilisation de l'informatique est la télésurveillance médicale à domicile, qui revêt une importance fondamentale pour l'amélioration de la qualité des soins et de la vie des personnes nécessitant une attention particulière ou des soins. Elle consiste à installer un dispositif dans le domicile d'une personne afin de recueillir des informations sur son état de santé, permettant ainsi au praticien de poser un diagnostic et de fournir une assistance à distance au patient.

Cette utilisation de l'informatique et des technologies de communication contribue à faciliter l'accès aux soins de santé, à améliorer la qualité des services médicaux et à offrir une meilleure prise en charge des patients, en leur permettant de bénéficier de suivis médicaux personnalisés et d'une assistance à distance lorsque cela est nécessaire [47].

## II.5 Avantages et inconvénients de la télémédecine

### II.5.1 Avantages

- Accès amélioré aux soins de santé : La télémédecine permet aux patients d'accéder à des services médicaux de qualité, en particulier dans les régions éloignées, rurales ou mal desservies par les infrastructures médicales traditionnelles. Cela permet de surmonter les barrières géographiques et d'améliorer l'accès aux soins, en particulier pour les populations vulnérables.
- Réduction des coûts et des déplacements : La télémédecine permet de réduire les coûts associés aux déplacements fréquents vers les établissements de santé, en particulier pour les consultations de suivi, les résultats de laboratoire et les

conseils médicaux. Cela peut également réduire les dépenses liées aux hospitalisations et aux soins d'urgence.

- Rapidité et efficacité des soins : Grâce à la télémédecine, les patients peuvent obtenir rapidement des consultations médicales, des diagnostics et des traitements sans délai d'attente important. Cela permet une prise en charge précoce des problèmes de santé, une meilleure gestion des maladies chroniques et une réduction des complications potentielles.
- Collaboration interdisciplinaire : La télémédecine favorise la collaboration entre les professionnels de la santé en permettant des consultations à distance et des échanges d'informations médicales. Cela facilite la prise de décisions conjointes, le partage des connaissances et l'expertise multidisciplinaire pour des diagnostics plus précis et des plans de traitement appropriés.

## **II.6 Inconvénients**

- Limitations technologiques : La télémédecine dépend de la disponibilité de technologies fiables telles que l'accès Internet haut débit, les équipements de télécommunication et les systèmes informatiques. Dans certaines régions, en particulier dans les zones rurales ou en développement, ces technologies peuvent être limitées, ce qui peut entraver l'adoption et l'utilisation de la télémédecine.
- Confidentialité et sécurité des données : La télémédecine implique le transfert et le stockage de données médicales sensibles. Il est essentiel de garantir la confidentialité et la sécurité de ces informations pour protéger la vie privée des patients. Des précautions doivent être prises pour prévenir les violations de données et les accès non autorisés.
- Limitations de l'examen physique : Bien que la télémédecine permette aux médecins de consulter les patients à distance, elle ne permet pas un examen physique complet. Certains aspects de l'évaluation médicale, tels que la palpation ou l'écoute des sons corporels, peuvent être limités, ce qui peut avoir un impact sur la précision du diagnostic.
- Barrières réglementaires et juridiques : La télémédecine peut être soumise à des réglementations et des restrictions légales, ce qui peut varier selon les pays et les États. Il peut y avoir des défis liés à la licence, à la responsabilité médicale, au

remboursement des services et aux pratiques transfrontalières, ce qui peut freiner l'adoption généralisée de la télémédecine.

## **II.7 Le service de télémédecine pour les troubles de la parole et de la voix**

L'évaluation clinique de la qualité de la voix et le diagnostic des pathologies du larynx peuvent bénéficier de l'analyse du signal de parole. Cette approche acoustique présente plusieurs avantages, tels que sa simplicité de mise en œuvre, son faible coût et son caractère non invasif. Elle permet également de documenter quantitativement le degré d'enrouement perçu par le clinicien.

L'objectif de l'analyse acoustique est d'extraire des indices pertinents pour caractériser les propriétés de la voix et fournir des informations sur l'état du larynx du locuteur. Différents indices acoustiques sont utilisés pour caractériser la parole des individus présentant des troubles de la voix.

La télémédecine présente un potentiel considérable en fournissant des services à domicile ou dans la communauté par le biais de vidéoconférences et d'activités thérapeutiques interactives sur ordinateur. Cette approche permet d'optimiser les résultats fonctionnels en facilitant la généralisation des effets du traitement dans le quotidien de la personne et en permettant un suivi à long terme des comportements de communication et de déglutition.

La télémédecine basée sur l'imagerie a été utilisée dans la prise en charge de divers troubles tels que les troubles neurogènes de la parole et du langage chez les adultes, le bégaiement, les troubles de la voix, les troubles de la parole et du langage chez les enfants, la laryngectomie et les troubles de la déglutition. Cependant, pour intégrer pleinement la télémédecine à la pratique de l'orthophonie, il est nécessaire de poursuivre le développement de ces applications, d'évaluer leur rentabilité, d'assurer la formation professionnelle appropriée et de développer d'autres thérapies informatisées [48].

La figure II.2 illustre la configuration d'un système de télémédecine pour la détection et l'évaluation des troubles de la parole et de la voix [19].



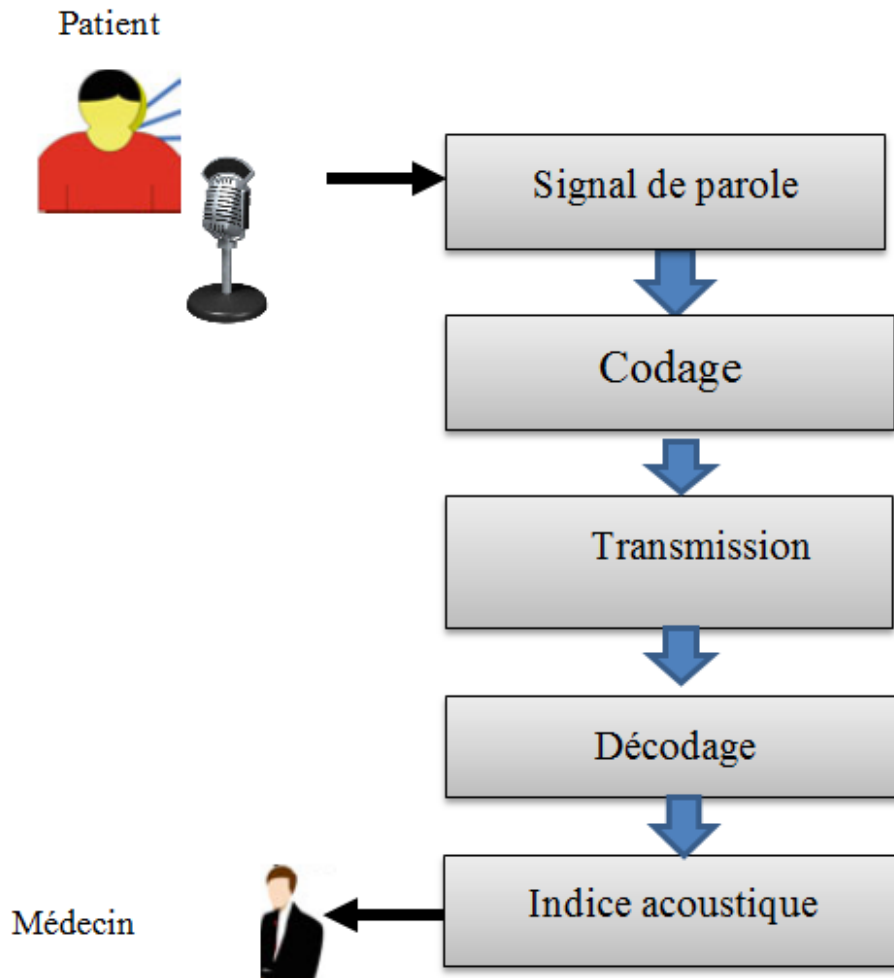


Figure II-2 Chaîne de transmission du signal de parole pour les services de télémédecine.

### II.7.1 Codage et compression de la parole

Au cœur du traitement du signal de la parole se trouve la nécessité de convertir le signal vocal en un signal électrique par le biais d'un microphone. La numérisation du signal vise à créer une représentation discrète d'un phénomène du monde réel. Ce processus se décompose en plusieurs étapes : l'échantillonnage, la quantification et le codage.

La première étape consiste en l'échantillonnage, où le signal vocal continu est échantillonné à intervalles réguliers pour obtenir une séquence de valeurs discrètes. Ensuite, lors de l'opération de quantification, chaque valeur échantillonnée est approximée à la valeur la plus proche d'un ensemble fini de niveaux de quantification prédéfinis. Cela permet de représenter numériquement l'amplitude du signal vocal à chaque instant.

Enfin, la dernière étape du processus de numérisation est le codage. Cette étape consiste à représenter les valeurs quantifiées en utilisant un format de codage approprié, ce qui permet de stocker et de transmettre efficacement les données numériques du signal vocal.

La compression de données, également appelée codage de source, vise à réduire la quantité de données à stocker, traiter ou transmettre, tout en préservant les informations nécessaires pour récupérer le signal lors du décodage (ou décompression).

Les différents algorithmes de compression se basent sur trois critères principaux :

**Taux de compression** : il s'agit du rapport entre la taille du fichier compressé et la taille du fichier initial. Il mesure l'efficacité de la compression en termes de réduction de la taille des données.

**Qualité de compression** : les algorithmes de compression peuvent être classés en deux catégories : avec ou sans perte. Les méthodes sans perte préservent intégralement les données d'origine, tandis que les méthodes avec perte entraînent une certaine perte d'information. La qualité de compression peut être évaluée en fonction du pourcentage de perte accepté.

**Vitesse de compression et de décompression** : ces critères mesurent le temps nécessaire pour compresser et décompresser les données. Ils peuvent varier en fonction de la complexité de l'algorithme de compression et des ressources disponibles [49].

## II.7.2 Classification des algorithmes de compression

Le critère de classification le plus pertinent est basé sur la perte des données. On peut distinguer deux types de compression qui seront présentés comme suit [50] :

### II.7.2.1 Compression sans perte

La compression sans perte, bien que moins performante que la compression avec perte, reste très utile dans certaines situations. Notamment, pour les fichiers de données où la moindre perte de données compromettrait leur intégrité et leur signification. Par exemple, dans le cas d'une feuille de comptabilité où la disparition d'un chiffre rendrait les données inutilisables. C'est pourquoi la compression sans perte conserve toute sa pertinence.

De plus, la compression sans perte est souvent utilisée en complément de la compression destructrice pour les fichiers multimédias. Après avoir appliqué une compression destructive, il est possible de réduire encore la taille du fichier compressé de 20 à 30% en

utilisant la compression sans perte. Ainsi, elle permet de réaliser des économies supplémentaires sans sacrifier davantage la qualité du contenu compressé [50].

### **II.7.2.2 Compression avec perte**

Les algorithmes de compression sans perte offrent une réduction significative de la taille des données en mémoire. Cependant, il est possible d'améliorer davantage ces taux de compression en utilisant des algorithmes dits avec perte ou destructifs. Cette approche devient d'autant plus intéressante pour les images haute résolution ou les fichiers audio de qualité CD, qui occupent beaucoup d'espace de stockage.

En utilisant des algorithmes de compression avec perte, il est possible de réduire considérablement la taille des fichiers multimédias tout en acceptant une certaine perte de qualité ou de précision. Cela permet d'économiser de l'espace de stockage sans compromettre complètement l'intelligibilité ou l'appréciation du contenu [50].

## **II.7.3 La technologie de compression mp3**

### **II.7.3.1 Principe et fonctionnement**

L'algorithme de compression audio MP3 (ou MPEG Audio Layer 3) est une technologie qui permet de réduire la taille des fichiers audio. Cette technologie a été développée en 1992 par le groupe d'experts Moving Picture Experts Group (MPEG), qui est chargé d'établir des normes internationales pour la compression, la décompression, le traitement et le codage de la vidéo et de l'audio.

En réalité, la technologie MP3 ne consiste pas simplement à compresser les données, mais plutôt à supprimer sélectivement certaines informations afin de simplifier et alléger le fichier audio. Son fonctionnement repose sur deux éléments clés : la capacité auditive humaine d'une part, et un modèle psycho acoustique d'autre part.

Ce modèle psycho acoustique permet de prendre en compte les caractéristiques de perception auditive de l'oreille humaine, en identifiant les parties du signal audio qui sont moins audibles ou moins importantes pour l'auditeur. Ainsi, ces parties moins perceptibles sont supprimées ou réduites de manière intelligente, ce qui permet de réduire la taille du fichier sans altérer significativement la qualité perçue par l'auditeur.

- **Suppression des fréquences inaudibles**

Il est important de noter que l'oreille humaine a une plage de fréquences limitée, capable de percevoir des sons dans la gamme de 20 Hz à 20 kHz. La technologie MP3 repose sur cette caractéristique en supprimant les données correspondant aux fréquences inaudibles pour l'oreille humaine. Cela permet de réduire la taille du fichier audio en éliminant les informations imperceptibles. L'objectif est d'alléger le fichier d'origine en éliminant ces données inaudibles, tout en conservant une qualité perceptive suffisante pour l'utilisateur. Ainsi, grâce à cette approche, les modifications apportées passent généralement inaperçues pour l'auditeur.

- **Suppression des fréquences masquées**

La compression au format MP3 repose également sur l'utilisation d'un modèle psychoacoustique qui prend en compte l'effet de masquage. Cet effet de masque se produit lorsque deux fréquences de différentes intensités coexistent, la fréquence la plus faible étant moins perceptible et masquée par la fréquence plus puissante. Pour mieux illustrer cela, imaginez-vous dans un jardin où vous entendez le chant des oiseaux. Si un avion passe au-dessus de votre tête (même à une altitude élevée), le son des oiseaux devient moins audible en raison de l'intensité plus élevée du son de l'avion. Il en résulte que l'on n'entend plus les oiseaux ou que leur chant semble avoir disparu. Cette observation montre qu'il n'est pas nécessaire de coder toutes les fréquences présentes dans un morceau pour que l'oreille humaine le perçoive correctement. Dans ce contexte, la technologie MP3 se concentre sur l'élimination des données correspondant à ces fréquences masquées que l'oreille humaine ne distingue pas. Ainsi, la taille du fichier est réduite sans que l'utilisateur puisse percevoir la différence.

### II.7.3.2 Algorithmes de codage

L'algorithme de compression MP3 utilise également la méthode de codage de Huffman, qui a été développée par Huffman en 1952. Cette méthode est largement utilisée dans divers algorithmes de compression, tels que la compression de fichiers texte, d'images et de sons. Elle repose sur l'utilisation d'un code à longueur variable en fonction de la probabilité d'apparition d'un événement, dans ce cas précis, la fréquence des données. Les fréquences qui apparaissent le plus souvent sont représentées par des codes plus courts (c'est-à-dire avec un nombre de bits plus faible). Ce codage de Huffman est appliqué à la fin du processus de compression et il permet une compression sans perte.

L'utilisation des différentes techniques de réduction mentionnées dépend de l'importance accordée à la taille finale d'une minute de son compressé, et donc du taux de compression nécessaire. Lorsqu'il s'agit d'encoder des fichiers en format MP3, on se réfère plutôt au taux de bits (bit rate) plutôt qu'au taux de compression. Le taux de bits représente le nombre de bits autorisés par seconde. Ainsi, si l'on souhaite compresser davantage un morceau de musique (pour réduire sa taille), il faudra choisir un taux de bits plus faible [51].

## **II.8 Conclusion**

La télémédecine a considérablement élargi les possibilités de prestation des soins de santé, offrant des solutions innovantes pour surmonter les barrières géographiques et améliorer l'accessibilité aux services médicaux. Dans ce chapitre, nous avons exploré les systèmes de télémédecine et leur application spécifique dans le domaine des troubles de la parole et de la voix.

Nous avons constaté que la télémédecine permet d'effectuer divers actes médicaux à distance, tels que la télé consultation, la télé surveillance, la télé expertise, la télé radiologie, la téléassistance, la télé chirurgie et la téléformation. Ces services offrent de nouvelles opportunités pour le diagnostic, le suivi et la prise en charge des troubles de la parole et de la voix, permettant aux patients d'accéder à des soins spécialisés sans les contraintes de distance et de déplacement.

Nous avons également examiné les avantages et les inconvénients de la télémédecine. Parmi les avantages, citons la facilité d'accès aux soins, la réduction des coûts, la flexibilité horaire, et la possibilité d'impliquer des experts à distance. Cependant, nous avons reconnu les défis associés à la sécurité des données, à la confidentialité et à la qualité de la connexion.

La télémédecine offre des opportunités considérables pour améliorer la prise en charge des troubles de la parole et de la voix. Elle ouvre de nouvelles perspectives pour la collaboration entre les professionnels de la santé, la diffusion des connaissances et l'accès aux soins pour les patients.

---

# Chapitre III

Acquisition comprimée

---

### III.1 Introduction

L'acquisition comprimée est une méthode novatrice de traitement du signal qui permet de capturer et de transmettre efficacement des données tout en réduisant considérablement la quantité d'informations nécessaires. Cette approche trouve de nombreuses applications dans divers domaines, y compris la santé, où elle peut jouer un rôle crucial dans le développement de systèmes de télémédecine avancés.

Dans le cadre de la télémédecine, la communication à distance entre les professionnels de la santé et les patients revêt une importance croissante, offrant des avantages tels que l'accès à des soins spécialisés pour les populations éloignées, la surveillance à domicile et le suivi à long terme des maladies chroniques. Cependant, la transmission en temps réel de données médicales, notamment de la voix, peut être un défi en raison des contraintes de bande passante, des limitations des réseaux de communication et de la nécessité de préserver la confidentialité des informations sensibles.

Dans ce chapitre nous présentons les principes fondamentaux de l'acquisition comprimée tout en donnant un exemple pour illustrer le processus de compression et de reconstruction du signal.

### III.2 Base de l'acquisition comprimée

Pour numériser un signal analogique à bande limitée, l'approche conventionnelle consiste à utiliser le théorème de Shannon-Nyquist : pour éviter toute perte d'information, le signal devrait être échantillonné à une fréquence supérieure ou égale au double de la fréquence maximale présente dans le signal. Si cette fréquence maximale est égale à  $f_{max}$ , alors la fréquence d'échantillonnage  $f_e$  devrait être supérieure ou égale à  $2f_{max}$ .

Cette fréquence minimale d'échantillonnage est appelée fréquence de Nyquist

$$f_N = 2f_{max} \quad (\text{III. 1})$$

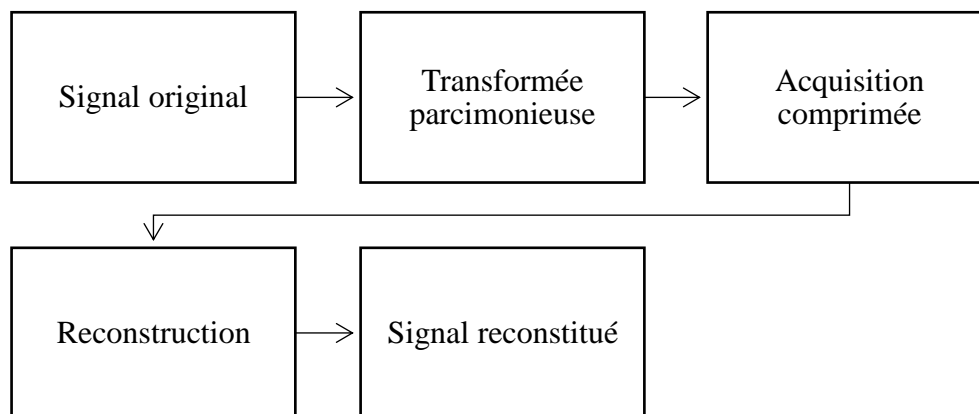
Pour certaines applications, comme le radar et les communications à large bande, l'application de ce théorème aboutit à des fréquences d'échantillonnage qui sont presque au-delà de la limite des capacités physiques des convertisseurs analogique-numérique.

Pour d'autres applications, comme le cas du WBAN (wireless body area network), même si elles ne sont pas directement confrontées à ce problème puisque la fréquence de Nyquist des signaux à mesurer n'est pas aussi élevée, les données numérisées devraient être compressées



pour économiser l'énergie, la bande passante du support de transmission et la mémoire de stockage. Pour contourner ce problème, l'acquisition comprimée (AC) (CS : compressed sensing) a été proposée. En plus de l'hypothèse de la bande limitée du signal, dans la plupart des cas, le signal à mesurer peut avoir une représentation parcimonieuse dans un domaine particulier. L'acquisition comprimée exploite cette information additionnelle pour capturer le signal à une fréquence en dessous de celle de Nyquist directement sous forme compressée et sans perte d'information [52].

La figure III.1 illustre le processus d'acquisition comprimée et de reconstruction du signal. Ce schéma met en évidence les étapes essentielles, telles que la transformée parcimonieuse, l'acquisition comprimée du signal ainsi que la reconstruction ultérieure du signal.



**Figure III-1** Processus de l'acquisition comprimée et de reconstruction du signal.

D'après le théorème Nyquist-Shannon, pour pouvoir récupérer le signal continu à partir de la version échantillonnée, il est nécessaire d'échantillonner le signal à une fréquence supérieure ou égale deux fois la fréquence maximale contenue dans le spectre du signal. Si cette condition n'est pas respectée, le phénomène de recouvrement spectral se produira rendant la récupération du signal inexacte.

En revanche, l'échantillonnage compressé, contrairement à l'approche classique, ne requiert pas cette condition, ce qui permet de réduire considérablement le temps et les coûts associés à l'échantillonnage. L'approche conventionnelle, qui consiste à échantillonner à grande vitesse et recompresser, présente deux inconvénients majeurs : elle entraîne un gaspillage important de ressources d'échantillonnage et génère des coûts matériels élevés. De plus, si des données compressées sont perdues, cela peut entraîner une récupération

inexacte du signal. En revanche, l'échantillonnage basé sur l'idée de l'échantillonnage compressé synchronise les processus d'échantillonnage et de compression, permettant de collecter et de comprimer le signal sans gaspiller de ressources [53].

## Représentation vectorielle du signal et normes

### Norme $l_1$

La mesure de la norme  $l_1$  est appelée "distance de Manhattan" ou "norme de taxi". En substituant 1 pour  $p$ , la norme  $l_1$  est la somme des magnitudes des éléments du vecteur/distance :

$$\|x\|_1 = (|x_1| + |x_2| + \dots + |x_n|) \quad (\text{III. 2})$$

En effet, il s'agit de la distance obtenue en se déplaçant uniquement dans la direction des dimensions de l'espace actuel. C'est comme si une personne se promenait dans un quartier de Manhattan où elle ne peut se déplacer que dans les directions fournies par les rues. C'est la raison pour laquelle on l'appelle "distance de Manhattan".

### Norme $l_2$

La norme  $l_2$  correspond à la distance la plus courte indiquée par un vecteur. Elle est également appelée norme euclidienne. En substituant 2 pour  $p$ , la norme  $l_2$  est la racine carrée de la somme des carrés des magnitudes des éléments du vecteur/distance :

$$\|x\|_2 = \sqrt{x_1^2 + x_2^2 + \dots + x_n^2} \quad (\text{III. 3})$$

### Norme $l_\infty$

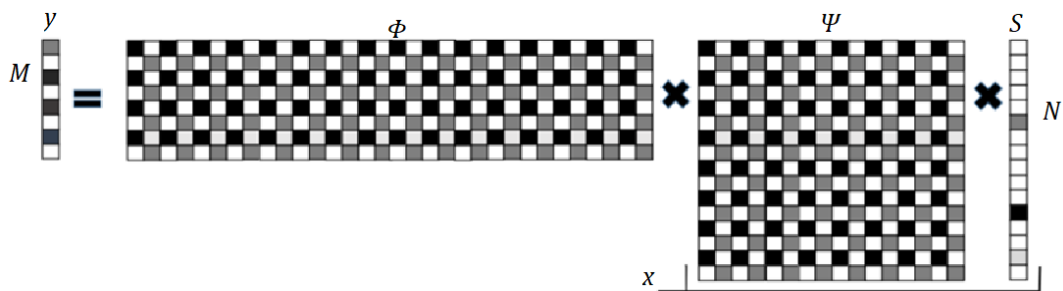
La norme  $l_\infty$  donne la plus grande valeur absolue dans le vecteur. En réalité, l'élément ayant la plus grande valeur absolue est le seul qui a un impact sur la distance de la norme  $l_\infty$ .

Ces normes, permettent de quantifier différentes caractéristiques des vecteurs et sont utilisées dans de nombreux domaines tels que les mathématiques, le traitement du signal et l'apprentissage automatique.

### III.3 Principe de l'acquisition comprimée

#### III.3.1 Modèle

L'acquisition comprimée est un nouveau cadre pour l'acquisition de signaux et la conception de capteurs. Il permet une réduction potentielle importante des coûts d'échantillonnage et de calcul pour les signaux qui ont une représentation parcimonieuse ou compressible. Alors que le théorème d'échantillonnage de Nyquist-Shannon stipule qu'un certain nombre minimum d'échantillons est nécessaire pour capturer parfaitement un signal arbitrairement limité en bande, lorsque le signal est parcimonieux dans une base connue, nous pouvons réduire considérablement le nombre de mesures à stocker. Ainsi, lors de l'échantillonnage de signaux parcimonieux, il est possible d'obtenir de meilleurs résultats que ceux suggérés par les résultats classiques. C'est l'idée fondamentale derrière le CS, au lieu d'échantillonner d'abord à un taux élevé puis de compresser les données échantillonnées, nous aimerions trouver des moyens de capter directement les données sous une forme compressée, c'est-à-dire à un taux d'échantillonnage inférieur. Pour rendre cela possible, le compressed sensing s'appuie sur deux principes : la parcimonie et l'incohérence [54].



**Figure III-2** Le modèle mathématique de l'acquisition comprimée.

Selon la figure III.2 [53], le modèle mathématique de l'acquisition comprimée peut être représenté de la manière suivante :

$$y = \Phi x \quad (\text{III. 4})$$

Dans cette équation,  $y$  représente le vecteur de mesures, qui est obtenu à partir d'un échantillonnage de  $x$  à l'aide de la matrice de mesure  $\Phi$ .  $x$  est le signal initial que l'on souhaite acquérir.

L'objectif de l'acquisition comprimée est de récupérer le signal initial  $x$  à partir des mesures compressées  $y$ , en utilisant la matrice de mesure  $\Phi$ . Cette approche est basée sur l'hypothèse que le signal initial  $x$  peut être représenté de manière parcimonieuse ou compressible dans une certaine base ou un dictionnaire [53].

Lorsque le signal initial  $x$  est parcimonieux, c'est-à-dire qu'il peut être représenté efficacement avec un petit nombre de coefficients non nuls dans une certaine base, il est possible de reconstruire le signal à partir d'un nombre réduit de mesures. Cela permet de réduire considérablement le coût d'échantillonnage et de stockage par rapport aux méthodes traditionnelles basées sur le théorème d'échantillonnage de Nyquist-Shannon.

Dans certains cas, le signal initial  $x$  peut ne pas être directement parcimonieux, mais il peut être sparsifié en utilisant une matrice de transformation  $\Psi$  et des coefficients parcimonieux  $s$ . Ainsi, l'équation peut être réécrite comme :

$$x = \Psi s \quad (\text{III. 5})$$

En combinant la matrice de mesure  $\Phi$  et la matrice de transformation  $\Psi$ , on obtient une matrice de détection  $A$  :

$$A = \Phi \Psi \quad (\text{III. 6})$$

Par conséquent, l'équation de l'acquisition comprimée peut être formulée comme suit :

$$y = A s \quad (\text{III. 7})$$

Dans certains cas, lorsque le signal initial  $x$  est décomposé de manière parcimonieuse en utilisant un dictionnaire  $D$  et un codage parcimonieux, l'équation peut être exprimée comme:

$$y = \Phi D * s \quad (\text{III. 8})$$

### III.3.2 Incohérence

Soit  $(\Phi, \Psi)$  une paire de bases orthogonales de  $R^N$ . La première base  $\Phi$  (également appelée matrice de mesure) est utilisée pour la détection ou la mesure du signal vocal, tandis

que la deuxième base  $\Psi$  (également appelée matrice de parcimonie ou de représentation) est utilisée pour représenter le signal vocal de manière parcimonieuse

L'incohérence entre la matrice de mesure  $\Phi$  et la base creuse  $\Psi$  est une caractéristique essentielle en acquisition comprimée. Elle exprime la faible corrélation mutuelle entre les éléments de  $\Phi$  et  $\Psi$ . La cohérence mutuelle est mesurée par la plus grande corrélation absolue entre les colonnes de  $\Phi$  et  $\Psi$ . Cette cohérence mutuelle peut être quantifiée par le coefficient  $\mu(\Phi, \Psi)$ , défini comme la valeur maximale de la corrélation entre toutes les paires de colonnes de  $\Phi$  et  $\Psi$ . Mathématiquement, cela peut être exprimé comme suit [55] :

$$\mu(\Phi, \Psi) = \alpha \max_{1 \leq k, j \leq N} |\langle \Phi_k, \Psi_j \rangle| \quad (\text{III. 9})$$

Où  $\Phi_k$  et  $\Psi_j$  représentent respectivement les  $k^{\text{ième}}$  et  $j^{\text{ième}}$  colonnes de  $\Phi$  et  $\Psi$ . La valeur  $\alpha$  dépend des propriétés algébriques des matrices choisies  $\Phi$  et  $\Psi$ . Divers auteurs ont proposé différentes valeurs pour  $\alpha$ , telles que  $1/N$  ( $N$  représente la taille des vecteurs dans les bases  $\Phi$  et  $\Psi$ ),  $1/\sqrt{N}$  ou une valeur normalisée à un.

La cohérence est une mesure grossière du degré de similarité ou de corrélation entre la parcimonie du signal et les systèmes de mesure. Une valeur de  $\mu = 1$  correspond à une incohérence minimale, ce qui signifie que les lignes de la matrice de mesure  $\Phi$  sont étalées dans le domaine  $\Psi$ . En d'autres termes, plus la cohérence est faible, plus l'incohérence est grande. Cette idée est fondamentale dans le contexte de l'acquisition comprimée, car le choix de la matrice de mesure doit être incohérent avec la base dans laquelle la parcimonie du signal est exprimée.

Pour illustrer cela, considérons deux exemples spécifiques de matrices  $\Phi$  et  $\Psi$ .

Dans le premier exemple,  $\Phi$  représente une matrice de mesure correspondant à un échantillonnage classique dans le temps ou dans l'espace, tandis que  $\Psi$  est la base de Fourier. Dans ce cas, le couple temps-fréquence satisfait  $\mu(\Phi, \Psi) = 1$ , ce qui implique une incohérence maximale. Cette notion d'incohérence prolonge la dualité entre le temps et la fréquence, où un signal fortement localisé dans le temps sera étalé en fréquence (principe d'incertitude).

Dans le deuxième exemple, la matrice de mesure  $\Phi$  est choisie comme une matrice aléatoire gaussienne dont les éléments sont indépendants les uns des autres, tandis que la matrice de représentation  $\Psi$  est une base d'ondelettes orthogonales telle que les ondelettes de Daubechies. Ce choix est motivé par le fait qu'une matrice de mesure aléatoire est

largement incohérente avec n'importe quelle base fixe  $\Psi$ . Cela permet d'exploiter l'incohérence entre  $\Phi$  et  $\Psi$  pour obtenir des performances élevées en acquisition comprimée [56].

L'incohérence en acquisition comprimée est un concept clé qui étend la notion de dualité entre le temps et la fréquence. Elle exprime l'idée que les objets qui ont une représentation parcimonieuse dans un certain domaine, noté  $\Psi$ , doivent être dispersés dans le domaine dans lequel ils sont acquis. Pour mieux comprendre ce concept, prenons l'exemple d'un signal dans le domaine temporel. Si ce signal est représenté par un Dirac ou un pic, sa représentation dans le domaine fréquentiel sera étalée sur plusieurs fréquences, ce qui signifie qu'il sera dispersé.

L'incohérence est essentielle dans le contexte de l'acquisition comprimée car elle permet de garantir la possibilité de récupérer le signal d'intérêt de manière exacte à partir d'un nombre limité de mesures. En d'autres termes, elle assure que le signal peut être reconstruit de manière précise et efficace à partir d'un nombre réduit de mesures.

### III.3.3 Parcimonie

Dans la plupart des cas, un vecteur de signal dans un espace vectoriel de grande dimension transporte des informations pouvant être exprimées dans des dimensions beaucoup plus basses que l'espace auquel appartient le signal. Surtout dans le cas des signaux naturels, on peut observer que la plupart de la durée du signal contient une faible variance et que seules quelques parties contiennent des informations à variance élevée. C'est précisément la propriété de parcimonie, où quelques coefficients contiennent une grande partie de l'énergie de nombreux coefficients. En conséquence, dans un tel vecteur de signal, il existe  $k$  éléments les plus dominants qui portent le plus d'impact. Un tel signal est appelé un signal  $k$  – *parcimonieux* [57].

Mathématiquement, nous avons un vecteur  $f \in R^n$ , représentant une image de  $n$  pixels comme indiqué dans la Figure III.2, que nous développons dans une base orthonormée telle qu'une base d'ondelettes  $\Psi = [\Psi_1, \Psi_2, \dots, \Psi_n]$  de la manière suivante :

$$f(t) = \sum_{i=1}^n x_i \Psi_i(t) \quad (\text{III. 10})$$

Ici,  $x$  est la suite de coefficients de  $f$ , où  $x_i$  représente le produit scalaire entre  $f$  et  $\Psi_i$ . Nous pouvons donc exprimer  $f$  comme une combinaison linéaire des éléments de la base d'ondelettes  $\Psi$  multipliés par les coefficients  $x$  correspondants.

L'implication de la parcimonie est maintenant évidente : lorsque le signal a une représentation parcimonieuse, nous pouvons négliger les petits coefficients sans grande perte d'information. Formellement, considérons  $f_s(t)$  obtenu en ne conservant que les termes correspondant aux  $S$  plus grandes valeurs de  $(x_i)$ . Par définition,  $f_s := \psi x_s$ , où  $x_s$  est le vecteur de coefficients  $(x_i)$  avec tous les coefficients mis à zéro, sauf les  $S$  plus grandes valeurs. Ce vecteur est parcimonieux au sens strict, car toutes ses entrées, sauf quelques-unes, sont nulles. Nous appellerons  $S$ -parcimonieux de tels objets avec au plus  $S$  entrées non nulles. Puisque  $\psi$  est une base orthonormée, nous avons  $\|f - f_s\|_{l_2} = \|x - x_s\|_{l_2}$ , et si  $x$  est parcimonieux ou compressible, ce qui signifie que les valeurs triées des  $x_i$  décroissent rapidement, alors  $x$  est bien approximé par  $x_s$  et, par conséquent, l'erreur  $\|f - f_s\|_{l_2}$  est petite. En d'autres termes, nous pouvons négliger une grande partie des coefficients sans subir de grandes pertes [58].

### III.3.4 Domaines de transformation

Le signal d'échantillonnage doit être parcimonieux, ou approximativement parcimonieux, pour être traité par des algorithmes d'échantillonnage compressé. En conditions normales, les signaux de parole en général ne sont pas parcimonieux par nature, mais grâce à une certaine transformation, nous pouvons obtenir un signal approximativement parcimonieux. Plusieurs transformations peuvent rendre les signaux de parole parcimonieux.

#### (i). Transformée de Fourier discrète

Il s'agit de la méthode de parcimonie la plus élémentaire. Cependant, la transformée de Fourier discrète (DFT : discrete Fourier transform) n'est pas adaptée à la parole car il y a beaucoup de composantes de fréquence dans la parole, ce qui entraîne des coefficients de la DFT de grandes amplitudes empêchant ainsi de supprimer des composantes spectrales pour augmenter la parcimonie. De plus, les coefficients de la DFT sont généralement des nombres complexes, ce qui rend le calcul plus compliqué.

**(ii). Transformée en ondelettes discrètes**

La transformée en ondelettes discrète ( DWT : discrete wavelet transform) utilise des ondelettes comme fonctions de base. Une ondelette a généralement un support fini, contrairement aux transformées de Fourier. Par conséquent, elle peut être utilisée pour rendre les signaux parcimonieux. Cependant, elle n'est pas totalement adaptée à la parole.

**(iii). Transformée en cosinus discrète**

La transformée en cosinus discrète (DCT : discrete cosine transform) peut éliminer la redondance entre les valeurs voisines conduisant à des coefficients de transformation non corrélés qui peuvent être encodés indépendamment. Une caractéristique importante de la DCT est que l'énergie des signaux se concentre dans les coefficients d'ordre inférieur. Par conséquent, nous pouvons supprimer les coefficients d'ordre supérieur pour rendre le signal parcimonieux dans le domaine fréquentiel. De plus, tous les coefficients de la DCT sont des nombres réels purs.

Parmi ces domaines de transformation, la DCT offre la meilleure représentation parcimonieuse de la parole. Ainsi, nous pouvons même augmenter la parcimonie de la parole en seuillant les composantes de faible valeur sans affecter considérablement la qualité de la parole [59].

**III.3.5 Traitement du signal**

Trouver le dictionnaire idéal et l'algorithme de transformation parcimonieuse correspondant au signal est nécessaire pour transformer le signal de manière parcimonieuse. Généralement, la parcimonie dans divers domaines de transformation est obtenue en utilisant différentes approches pour obtenir un dictionnaire adapté au signal à traiter. Pour ce problème, une équation d'optimisation est utilisée pour trouver l'équation sous-déterminée.

$$\hat{s} = \arg \min \|s\|_0 \text{ s. t. } x = D * s \quad (\text{III. 11})$$

En pratique, pour faciliter les calculs mathématiques, la norme  $l1$  et la norme  $lp$  peuvent également être utilisées pour approximer la norme  $l0$  afin de mesurer la parcimonie du signal [53].



La conception de la matrice de mesure du signal est un autre problème clé à résoudre dans la théorie de la compression par échantillonnage. La limitation fondamentale à prendre en compte lors de la conception d'une matrice de mesure optimale pour le signal est l'incohérence. En pratique, la matrice de mesure du signal doit être utilisée pour réduire la dimension d'un signal original de grande dimension. La conception d'une matrice de mesure appropriée permettra de réduire significativement le nombre de signaux originaux de grande dimension et de préserver autant que possible les signaux originaux lors du processus de réduction. Le degré d'incohérence entre la matrice de mesure et la matrice parcimonieuse est associé aux performances de récupération des signaux originaux.

L'essence de la propriété d'isométrie restreinte (RIP) est que tous les sous-ensembles de colonnes dans la matrice de mesure sont orthogonaux. Comme il est impossible de déterminer la RIP de la plupart des matrices, en réalité, la cohérence est souvent utilisée pour évaluer les performances de récupération de la matrice de mesure [53].

Plusieurs algorithmes ont été développés pour la récupération des signaux dans l'acquisition comprimée. Ces algorithmes visent à résoudre le problème d'optimisation qui consiste à trouver le signal parcimonieux qui correspond le mieux aux mesures compressées. Certains des algorithmes couramment utilisés comprennent :

### **L'algorithmes glouton (greedy algorithm)**

L'algorithme glouton est l'algorithme le plus direct pour résoudre des équations sous-déterminées, qui progresse à travers des itérations successives pour atteindre la solution finale étape par étape. L'algorithme glouton a une complexité faible, mais il est facilement perturbé par le bruit. L'algorithme de poursuite orthogonale (Orthogonal Matching Pursuit) est le plus représentatif, qui utilise l'idée de l'orthogonalisation pour éviter la sélection répétée des atomes et recalcule les vecteurs de coefficients non nuls après chaque nouvel ensemble d'indices obtenu, ce qui améliore davantage la vitesse de convergence de l'algorithme [53].

### **L'algorithme d'optimisation convexe (Convex optimization algorithm)**

L'algorithme d'optimisation convexe est une méthode courante dans le domaine de la récupération des signaux. L'idée fondamentale est de transformer un problème non convexe

en un problème convexe accessible grâce à une optimisation sans contrainte, afin d'obtenir le signal approximatif.

Les algorithmes classiques d'optimisation convexe comprennent l'algorithme de poursuite de base (Basis Pursuit), l'algorithme de poursuite de base avec débruitage (Basis Pursuit De-Noiseing) et l'algorithme de projection du gradient (Gradient Projection). L'algorithme Basis Pursuit propose de résoudre le problème d'optimisation en remplaçant la norme  $l_0$  par la norme  $l_1$  [53].

### **Algorithme d'optimisation non convexe (Non-convex optimization algorithm)**

Bien qu'il existe de nombreux algorithmes bien établis pour résoudre des problèmes convexes, leur modèle doit utiliser une optimisation convexe plutôt qu'une optimisation non convexe, ce qui aura un impact sur la récupération des signaux. L'algorithme d'optimisation non convexe diffère de l'algorithme d'optimisation convexe en ce qu'il utilise une fonction d'optimisation parcimonieuse continue non convexe pour obtenir une représentation parcimonieuse et ne prend pas en compte la solution optimale existante dans l'itération. Par conséquent, le résultat global de chaque itération est meilleur que celui de l'algorithme glouton. De plus, l'algorithme d'optimisation non convexe n'utilise pas la norme  $l_1$  à la place de la norme  $l_0$ , il n'a donc pas d'exigences strictes sur la propriété d'isométrie restreinte (RIP : restricted isometry property) de la matrice de mesure. Les algorithmes d'optimisation non convexe couramment utilisés comprennent l'algorithme de la norme  $l_0$  lissée (Smoothed  $l_0$  norm) et l'algorithme de pénalité de somme logarithmique (Log Sum Penalty). L'algorithme de la norme  $l_0$  lissée consiste à transformer le problème de minimisation d'une fonction discrète en un problème de valeur extrême d'une fonction continue et à le résoudre en utilisant la méthode de la descente la plus raide (Steepest-descent) et le principe de projection du gradient (gradient projection) [53].

### **III.3.6 Exemple d'acquisition et de reconstruction**

L'exemple met en œuvre la technique d'acquisition comprimée en exploitant la structure parcimonieuse des signaux dans des bases appropriées.

Tout d'abord, un signal original  $f$  représentant une combinaison linéaire de deux sinusoides de fréquences est généré.

$$f(t) = \sin(1394\pi t) + \sin(3266\pi t) \quad (\text{III. 12})$$

Ce signal est ensuite sous-échantillonné pendant 1/8 de seconde à une fréquence de 40000 Hz, le résultat est un vecteur  $f$  de longueur  $N = 5000$ , Le graphique supérieur de la Figure III.3 montre une partie de ce signal, ainsi que certains des  $M = 500$  échantillons prélevés aléatoirement.

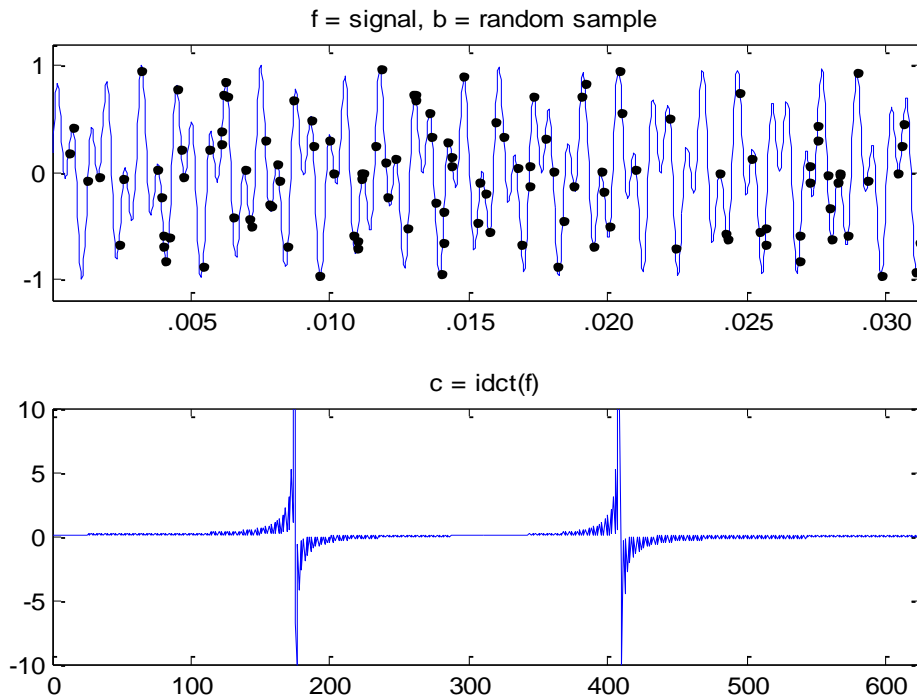
Le graphique inférieur montre les coefficients  $c$ , obtenus en prenant la transformée en cosinus discrète inverse de  $f$ , avec deux pics aux fréquences appropriées. Étant donné que les deux fréquences sont incommensurables, ce signal ne se trouve pas exactement dans l'espace engendré par les fonctions de base de la DCT, et il y a donc quelques dizaines de coefficients significatifs non nuls, le signal compressé est un vecteur  $b$  de  $M$  échantillons aléatoires du signal original.

Le système linéaire résultant,  $Ax = b$ , a une dimension de  $M * N$ , soit  $500*5000$ . Il y a dix fois plus d'inconnues que d'équations.

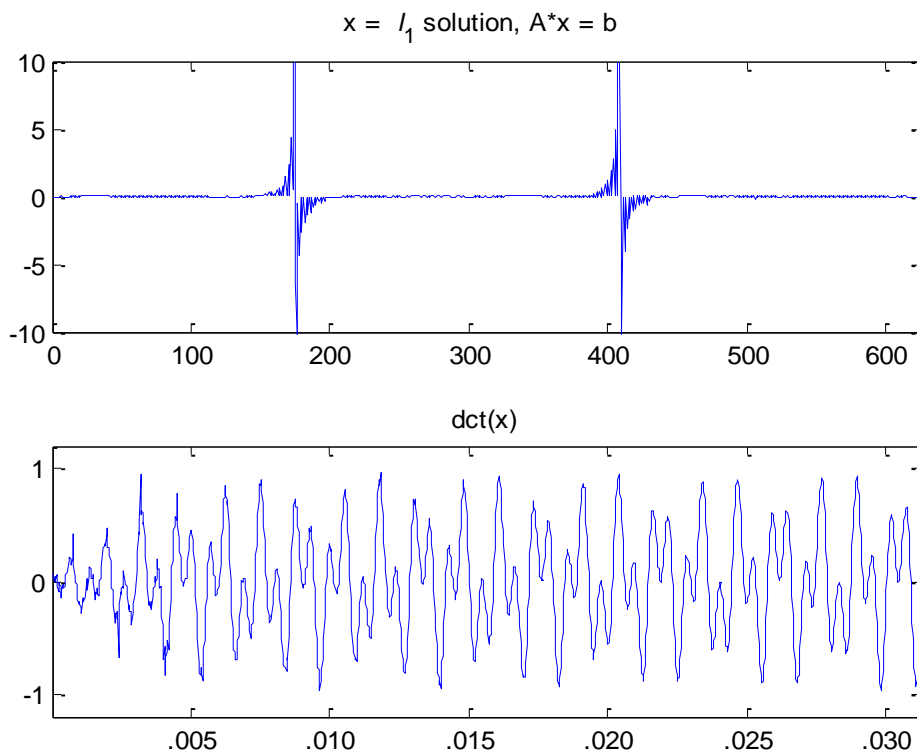
Pour reconstruire le signal, nous devons trouver la solution à  $Ax = b$  qui minimise la norme  $l_1$  de  $x$ . Il s'agit d'un problème d'optimisation non linéaire, et il existe plusieurs programmes basés sur MATLAB disponibles pour le résoudre. Le programme  $l_1$ -magic, écrit par Justin Romberg et Emmanuel Candès [60] a été utilisé dans cet exemple. Le graphique supérieur de la Figure III.4 montre la solution obtenue  $x$ . Le signal  $\text{DCT}(x)$  est presque identique au signal original.

Nous remarquons qu'elle comporte relativement peu de composantes importantes et qu'elle ressemble étroitement à la DCT du signal original. De plus, la transformée en cosinus discrète de  $x$ , montrée dans le graphique inférieur, ressemble de près au signal original [61].

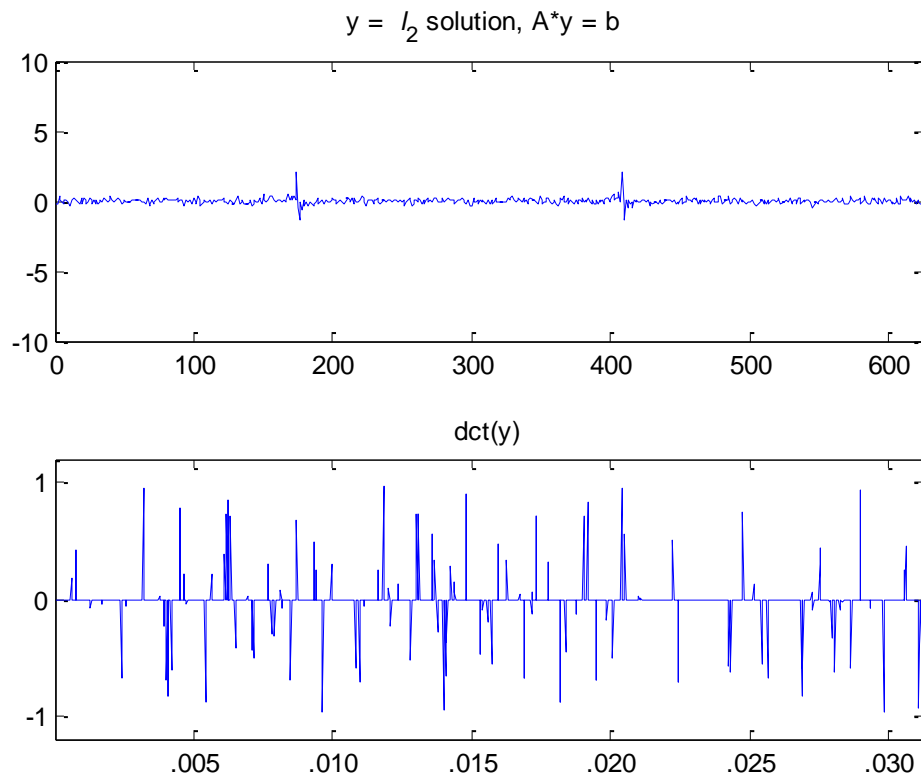
A titre de comparaison, la Figure III.5 montre la solution traditionnelle de moindres carrés ( $l_2$ ) à  $Ay = b$ . Le signal  $\text{DCT}(y)$  ne ressemble guère au signal original.



**Figure III-3** Échantillons aléatoires du signal original (en haut) et transformée en cosinus discrète inverse du signal (en bas).



**Figure III-4** Solution  $l_1$  de l'équation  $Ax = b$ .



**Figure III-5** Solution *l<sub>2</sub>* de l'équation  $Ay = b$ .

### **III.4 Conclusion**

Dans ce chapitre nous avons exploré les principes théoriques, les concepts clés et les aspects pratiques liés à l'acquisition comprimée. Nous avons pu constater son potentiel dans divers domaines. De plus, l'exemple de l'algorithme d'optimisation  $l_1$ -Magic utilisé a illustré de manière concrète l'efficacité de cette approche dans la récupération de signaux parcimonieux à partir d'échantillons compressés.

---

# Chapitre IV

Résultats et discussion

---

## IV.1 Introduction

Les applications de la télémédecine pour l'évaluation clinique à distance de la qualité de la voix dépendent de la technologie existante pour la transmission du signal vocal créé pour les communications cellulaires et Internet. La technique de compression du signal est un élément crucial de ces systèmes [62]. Le format MP3 bien connu développé pour la compression audio est largement utilisé dans ce contexte. Cependant, la compression du signal vocal au format MP3 présente de nombreuses limitations, en particulier pour la télémédecine, comme la transmission sécurisée [63,12] et l'authenticité des données [13].

L'acquisition compressée (CS : compressed sensing) est bien adaptée pour l'acquisition et la compression simultanée des signaux parcimonieux. Le signal de parole voisée est parcimonieux dans le domaine fréquentiel. Dans [19], la propriété de parcimonie a été exploitée pour compresser le signal de parole pour les applications de la télémédecine à l'évaluation des troubles de la voix. Le signal de parole produit par le patient est compressé en utilisant la technique de compression puis transmis à travers un moyen de communication pour être exploité par le médecin. Le signal compressé reçu par le médecin est reconstruit pour calculer les indices acoustiques servant de descripteurs des troubles de la voix. Il a été montré que l'acquisition compressée présente des performances de compression élevées. De plus, plusieurs travaux ont rapporté la capacité de l'acquisition compressée à résoudre des problèmes de transmission sécurisée et d'authenticité des données [17,64].

Dans ce chapitre, on utilise l'acquisition compressée pour la compression de la parole dysarthrique dans le contexte de l'analyse à distance de la parole pour la détection de la dysarthrie. L'effet de compression par la technique d'acquisition compressée est étudié en calculant les indices acoustiques couramment utilisés pour la détection et l'évaluation de la dysarthrie à partir du signal reconstruit et en les comparant aux indices acoustiques obtenus à partir du signal original sans compression ainsi qu'aux indices acoustiques calculés à partir du signal MP3 décompressé. Les indices acoustiques sont la période fondamentale (F0), les perturbations rapides (jitter) et lentes (tremor) des durées des cycles, les perturbations de l'amplitude (shimmer), le rapport harmonique sur bruit (HNR : harmonics-to-noise ratio) et la proéminence du pic cepstral (CPP : cepstral peak proeminence). Pour le calcul des indices acoustiques, on utilise le logiciel Praat à l'exception du tremor qui est implémenté avec Matlab à cause de sa non disponibilité dans Praat.



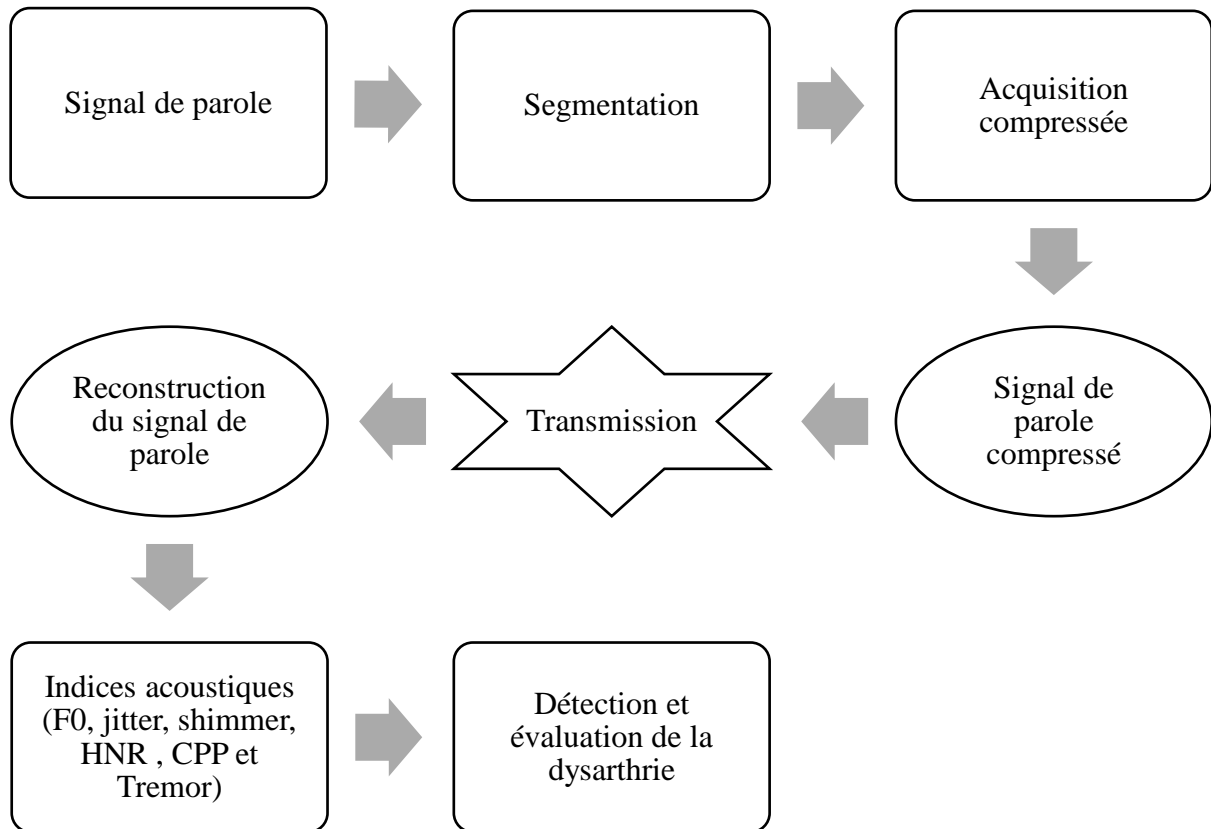
## IV.2 Base de données

La base de données utilisée contient des enregistrements de voyelles espagnoles prononcées par 100 locuteurs, dont 50 patients atteints de la maladie de Parkinson et 50 sujets sains (témoins ou contrôles), tous équilibrés en termes d'âge et de sexe [65]. Pour le groupe de Parkinsoniens, les hommes sont âgés de 33 à 77 ans (moyenne de  $62,2 \pm 11,2$ ), tandis que les femmes sont âgées de 44 à 75 ans (moyenne de  $60,1 \pm 7,8$ ). En ce qui concerne le groupe témoin, les hommes ont un âge compris entre 31 et 86 ans (moyenne de  $61,2 \pm 11,3$ ), et les femmes ont un âge compris entre 43 et 76 ans (moyenne de  $60,7 \pm 7,7$ ). Tous les participants ont été enregistrés dans une cabine insonorisée à l'aide d'un microphone dynamique omnidirectionnel et d'une carte audio professionnelle. La fréquence d'échantillonnage est de 44,1 kHz avec une résolution de 16 bits. Tous les patients ont été enregistrés en état d'activation, c'est-à-dire pas plus de 3 heures après leur médication matinale. Les patients ont été diagnostiqués par un neurologue expert.

## IV.3 Compression et reconstruction du signal de parole par acquisition comprimée

### Méthode

L'objectif de l'étude est d'étudier les effets de la compression de la parole avec la méthode d'acquisition comprimée dans le cadre de l'évaluation par télésanté de la parole dysarthrique. Les différentes étapes de la procédure sont illustrées à la figure IV.I. Tout d'abord, le signal vocal enregistré et échantillonné à une fréquence d'échantillonnage appropriée est segmenté en trames sans chevauchement. Le signal vocal est compressé en appliquant la méthode d'acquisition comprimée à chaque trame pour réduire la taille du signal. Ensuite, le signal vocal compressé est transmis à travers un canal de transmission. Au niveau du récepteur, le signal de parole est récupéré à partir du signal compressé en utilisant un algorithme basé sur la minimisation de la norme  $l1$ . Enfin, des indices acoustiques ( $F0$ , jitter, tremor, shimmer, HNR et CPP) sont calculés pour l'analyse de la parole et la détection de la dysarthrie.



**Figure IV-1** Illustration du système de télésanté pour la détection et l'évaluation de la dysarthrie.

### IV.3.1 Exemple d'illustration :

Comme exemple d'illustration, nous appliquons l'acquisition compressée aux voyelles [a] produites par un locuteur normophonique et un locuteur Parkinsonien. Les signaux sont reconstruits en utilisant le programme d'optimisation Matlab *l1-magic* basé sur la minimisation de la norme *l1*. La figure IV-2 montre des segments de 120 ms extraits des signaux reconstruits superposés aux segments correspondants du signal original pour une longueur de trame  $N=1323$  (une durée de 30 ms) et un nombre de mesures  $m=396$  ( $m=\lfloor 0.3*N \rfloor$ ). Les indices acoustiques sont calculés à partir du signal reconstruit par minimisation de la norme *l1* à l'aide du logiciel Praat à l'exception du tremor qui est implémenté avec Matlab.

Le tableau IV.I montre des valeurs des indices acoustiques calculés à partir du signal reconstruit par minimisation de la norme *l1*. Pour comparaison, nous avons également reporté les valeurs des indices acoustiques calculés à partir du signal original ainsi qu'à partir du signal MP3 décompressé.

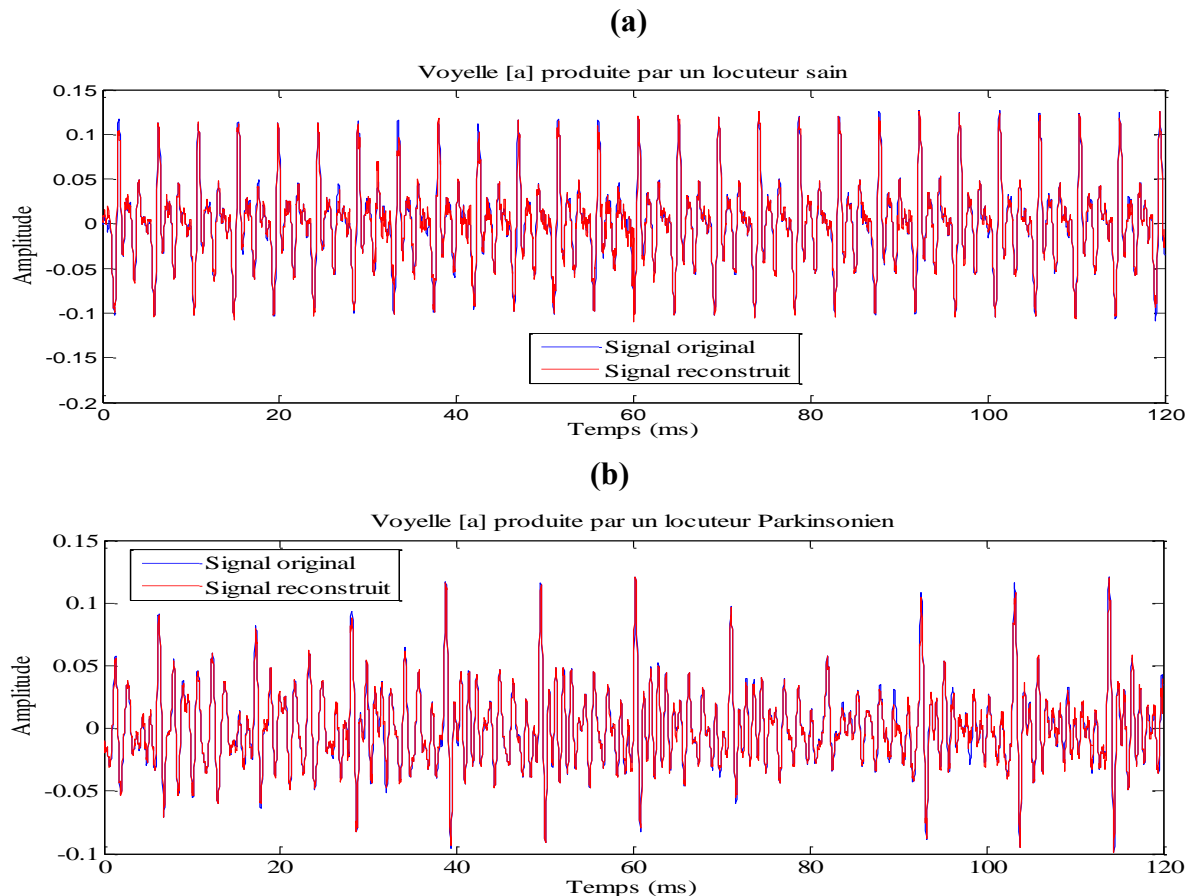
**Tableau IV.1** Valeurs des indices acoustiques calculés à partir du signal original, du signal reconstruit par minimisation de la norme  $l1$  et du signal MP3 décompressé.

	Locuteur sain			Locuteur parkinsonien		
	Original	CS	MP3	Original	CS	MP3
F0 (Hz)	188.4	188.5	188.87	91.5	91.5	91.6
Jitter (%)	0.20	0.223	0.2	0.86	0.84	0.858
Shimmer (%)	1.86	7.36	1.82	7.95	9.48	7.90
Tremor (%)	1.03	1.04	1.08	1.13	1.16	1.05
HNR (dB)	25.9	15.14	25.82	12.62	11.60	12.67
CPP (dB)	20.15	12.20	19.96	14.5900	11.39	14.36

#### IV.4 Résultats et discussion

Pour montrer l'efficacité de la méthode d'acquisition compressée pour la compression du signal vocal dans le cadre de l'évaluation par télésanté de la dysarthrie, la fréquence fondamentale et les indices acoustiques jitter, tremor, shimmer, HNR et CPP ont été calculés à partir du signal original et du signal reconstruit à partir de la version compressée par acquisition compressée. Les quartiles des valeurs des indices acoustiques calculés sont présentés sous forme de boîtes à moustaches (boxplots) dans les figures IV.3 à IV.8.

La Figure IV.2 illustre la superposition du signal original (en bleu) et du signal reconstruit par acquisition compressée (en rouge) pour un locuteur sain (a) et un l'autre parkinsonien (b).



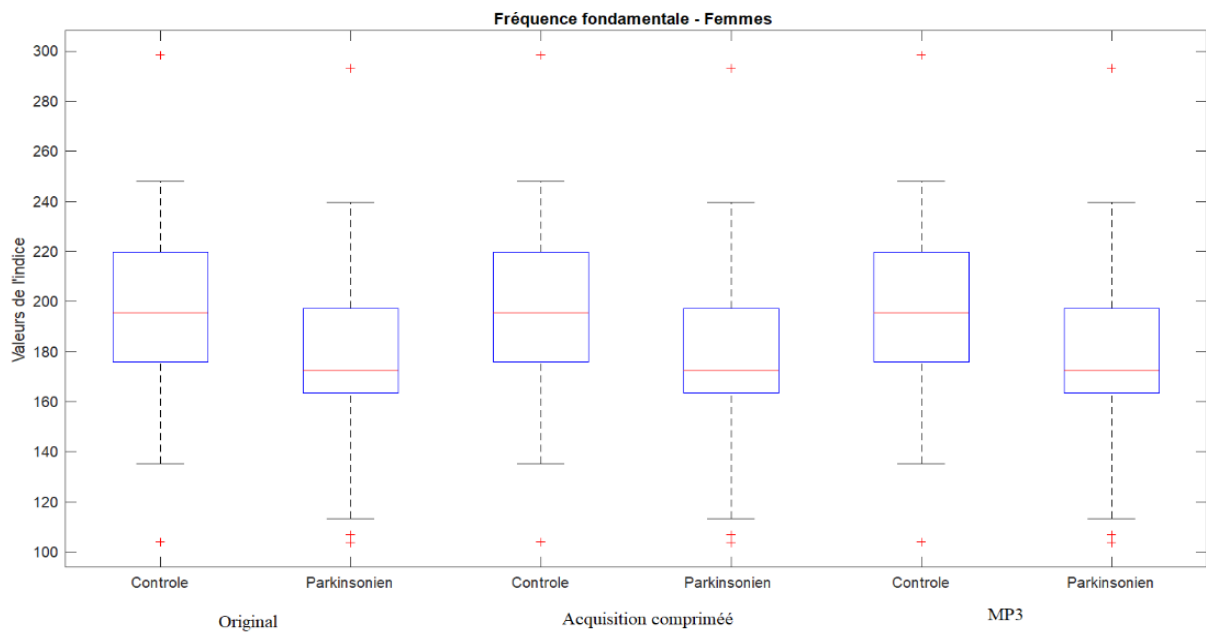
**Figure IV-2** Signal reconstruit par acquisition compressée (en rouge) superposé au signal original (en bleu). (a) Locuteur sain. (b) Locuteur Parkinsonien.

#### IV.4.1 Fréquence fondamentale

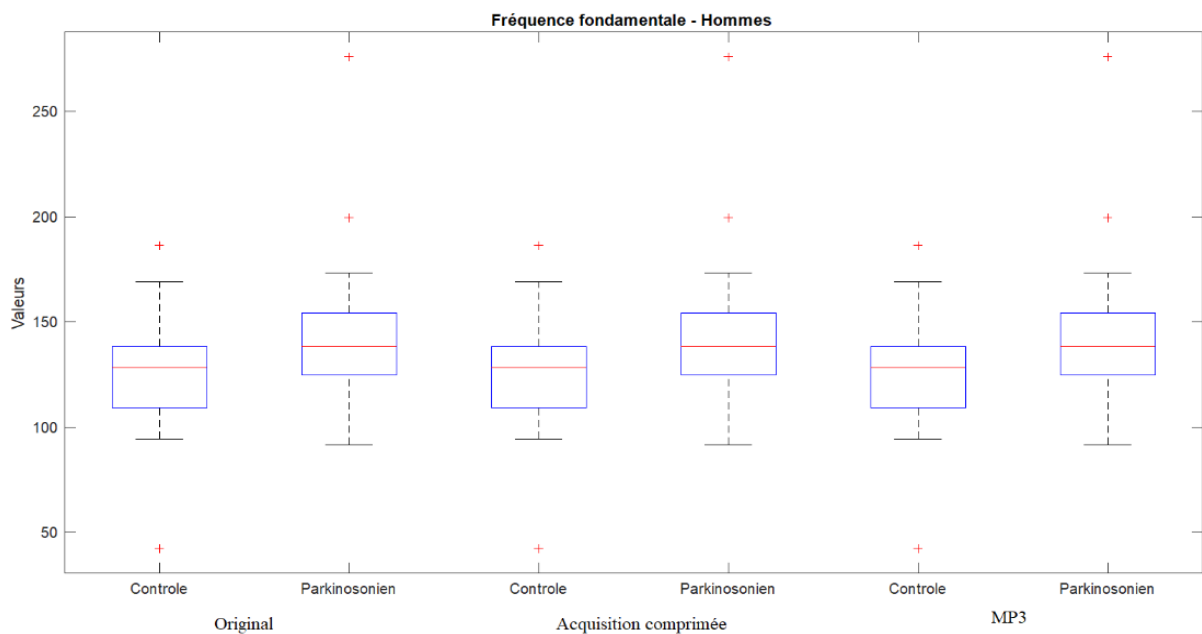
La figure IV.3 montre les boîtes à moustaches des fréquences vocales représentées séparément hommes et femmes. On constate que les fréquences vocales des hommes augmentent et les fréquences vocales des femmes diminuent en passant des locuteurs sains aux locuteurs Parkinsonien quel que soit le signal utilisé pour le calcul de la fréquence fondamentale. Les valeurs de la fréquence fondamentale obtenues à partir du signal de parole reconstruit par minimisation de la norme  $l_1$  ne diffèrent pas significativement des valeurs de la fréquence fondamentale obtenues à partir du signal de parole original ou MP3 décompressé. Cette similarité montre que la reconstruction du signal par minimisation  $l_1$  a préservé les caractéristiques du signal qui contribuent à la qualité et à la perception de la voix. En effet, l'algorithme MP3 utilise des techniques de compression perceptuelle qui exploitent les caractéristiques de l'audition humaine. Il élimine les informations auditivement moins perceptibles ou moins importantes pour réduire la taille du fichier audio.

Cependant, il vise à préserver les caractéristiques auditivement essentielles, telles que les variations de fréquence et de pitch qui contribuent à la qualité et à la perception de la voix.

(a)



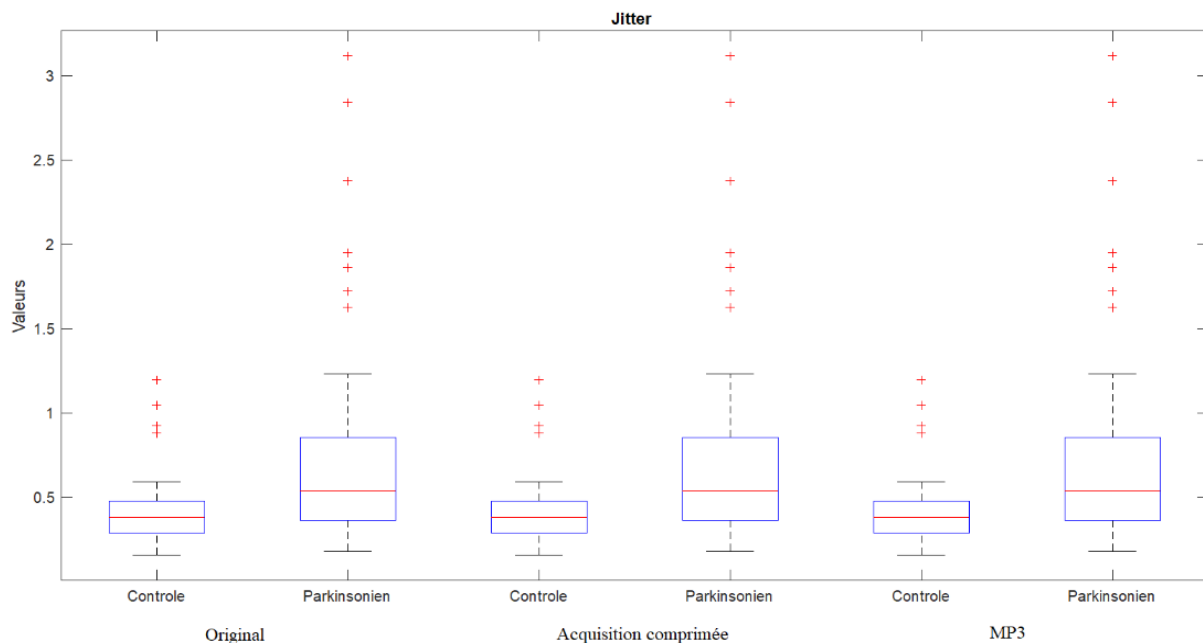
(b)



**Figure IV-3** Boîte à moustaches de la fréquence vocale en Hz calculée à partir du signal original, et du signal reconstruit par minimisation  $l_1$  et du signal MP3 décompressé pour les locuteurs contrôles et Parkinsoniens. (a) Fréquence vocale des Femmes. (b) Fréquence vocale des Hommes

### IV.4.2 Jitter

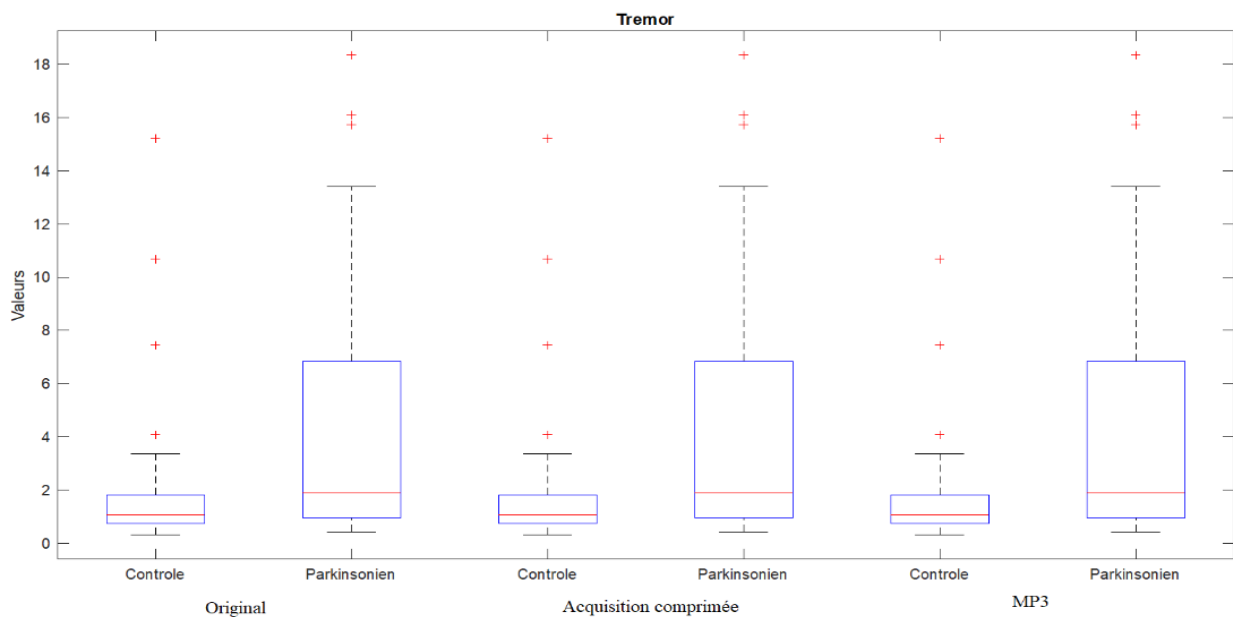
La figure IV.4 montre la boîte à moustaches des perturbations des durées des cycles (jitter). Une inspection visuelle montre que les valeurs du jitter sont plus élevées pour les locuteurs Parkinsoniens que pour les locuteurs contrôles. De même, on observe que les valeurs du jitter obtenues à partir du signal de parole reconstruit par minimisation de la norme  $l1$  diffèrent peu des valeurs du jitter obtenues à partir du signal de parole original ou MP3 décompressé. Ces résultats sont cohérents avec ce qui rapporté dans la littérature sur la maladie de Parkinson, qui est caractérisée par des troubles de la voix et de la parole, tels que l'augmentation du jitter, qui correspond à des fluctuations irrégulières de la fréquence fondamentale.



**Figure IV-4** Boîte à moustaches du jitter calculé à partir du signal original, et du signal reconstruit par minimisation  $l1$  et du signal MP3 décompressé pour les locuteurs contrôles et Parkinsoniens.

### IV.4.3 Tremor

Les percentiles des valeurs du tremor vocal calculées à partir du signal original ainsi qu'à partir du signal reconstruit par minimisation de la norme  $l_1$  et du signal MP3 décompressé sont représentées sur ma figure IV.5. On observe que les valeurs du tremor sont plus élevées pour les locuteurs Parkinsoniens que pour les locuteurs contrôles quel que soit le signal utilisé pour le calcul de l'indice acoustique. Cela suggère que les locuteurs Parkinsoniens présentent généralement un niveau de tremor plus élevé que les locuteurs contrôles. Cela confirme donc l'importance du tremor comme un descripteur de la dysarthrie. De même, on constate un chevauchement entre les intervalles des valeurs des deux groupes de locuteurs. Les valeurs du tremor obtenues à partir du signal reconstruit par minimisation  $l_1$  sont plus grandes que celles obtenues à partir du signal original et signal MP3 décompressé, cependant, cette augmentation des valeurs n'a pas d'effet sur l'efficacité de l'indice acoustique en termes de discrimination contrôle/Parkinsonien.

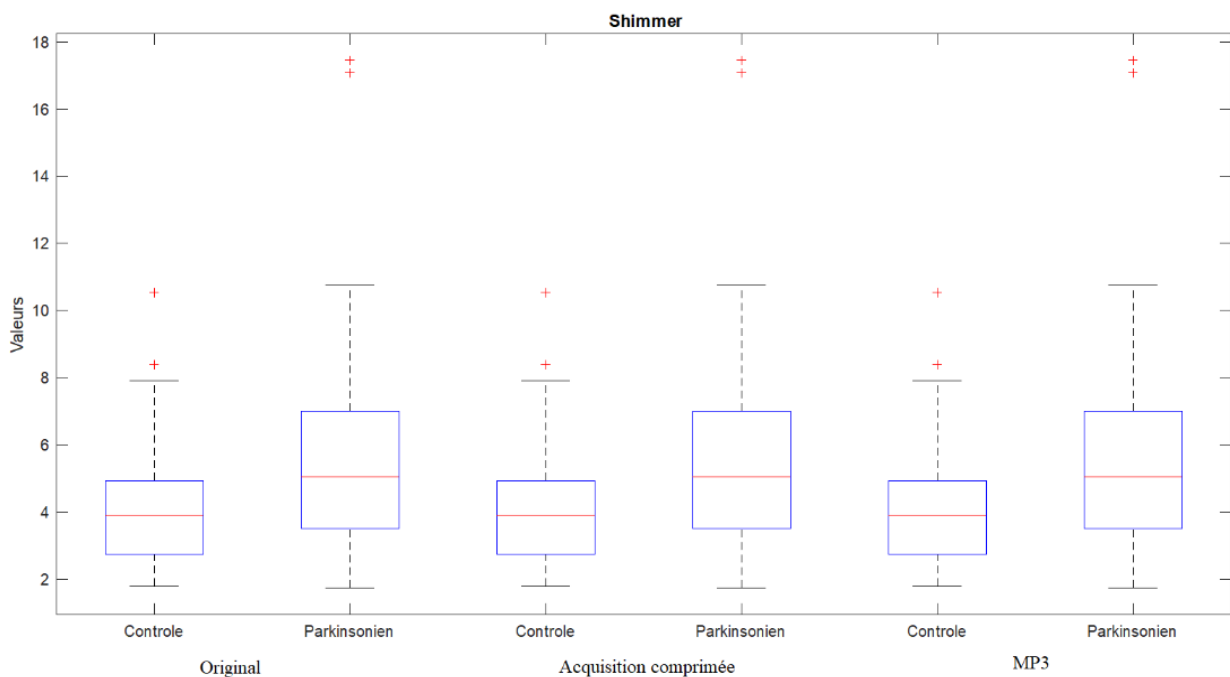


**Figure IV-5** Boîte à moustaches du tremor calculé à partir du signal original, et du signal reconstruit par minimisation  $l_1$  et du signal MP3 décompressé pour les locuteurs contrôles et Parkinsoniens.

### IV.4.4 Shimmer

Le shimmer est un paramètre acoustique utilisé pour évaluer la variation de l'amplitude de la voix. En examinant la boîte à moustaches de la figure IV.6, On remarque que les valeurs

du shimmer sont plus élevées pour les locuteurs Parkinsoniens que pour les locuteurs contrôles quel que soit le signal utilisé pour le calcul de l'indice acoustique. Cela suggère une plus grande variabilité et une irrégularité accrue dans la voix des locuteurs Parkinsoniens par rapport aux locuteurs contrôles. Les valeurs du shimmer obtenues à partir du signal de parole reconstruit par minimisation de la norme  $l_1$  ont une tendance d'être plus élevées que des valeurs du shimmer obtenues à partir du signal de parole original, cependant cette augmentation des valeurs n'affecte pas la discrimination contrôle/parkinsonien du shimmer. La compression MP3 a pour effet d'augmenter des valeurs du shimmer des locuteurs contrôles par rapport aux valeurs obtenues à partir du signal original. Contrairement, pour les locuteurs Parkinsoniens, la compression MP3 a conduit à une diminution des valeurs du shimmer par rapport aux signaux originaux. Cela indique que la compression MP3 a atténué certaines des variations naturelles de l'amplitude de la voix présentes dans les signaux vocaux des locuteurs Parkinsoniens. Cette diminution des valeurs peut également être attribuée aux pertes de données et aux distorsions induites par la compression MP3.



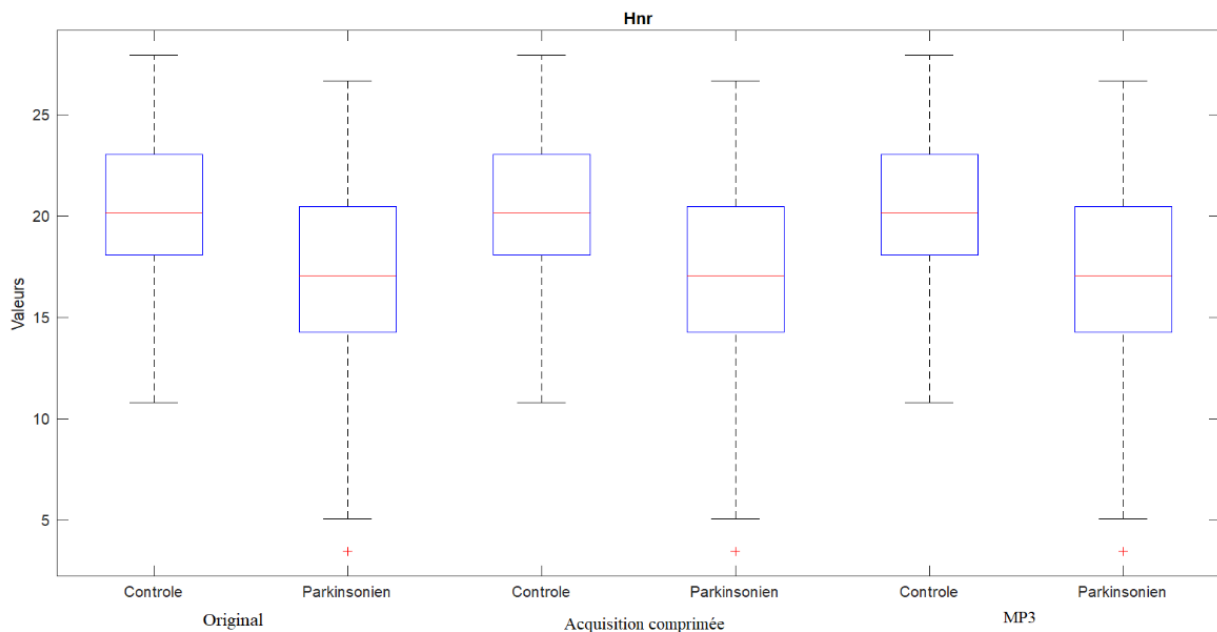
**Figure IV-6** Boîte à moustaches du shimmer calculé à partir du signal original, et du signal reconstruit par minimisation  $l_1$  et du signal MP3 décompressé pour les locuteurs contrôles et Parkinsoniens.



#### IV.4.5 Rapport harmonique sur bruit (HNR)

L'indice acoustique HNR mesure la qualité vocale en évaluant le rapport entre les harmoniques (sons purs) et le bruit dans un signal vocal. Une valeur élevée de l'indice HNR indique une meilleure qualité vocale avec moins de bruit, tandis qu'une valeur plus faible indique une détérioration de la qualité vocale et une augmentation du bruit.

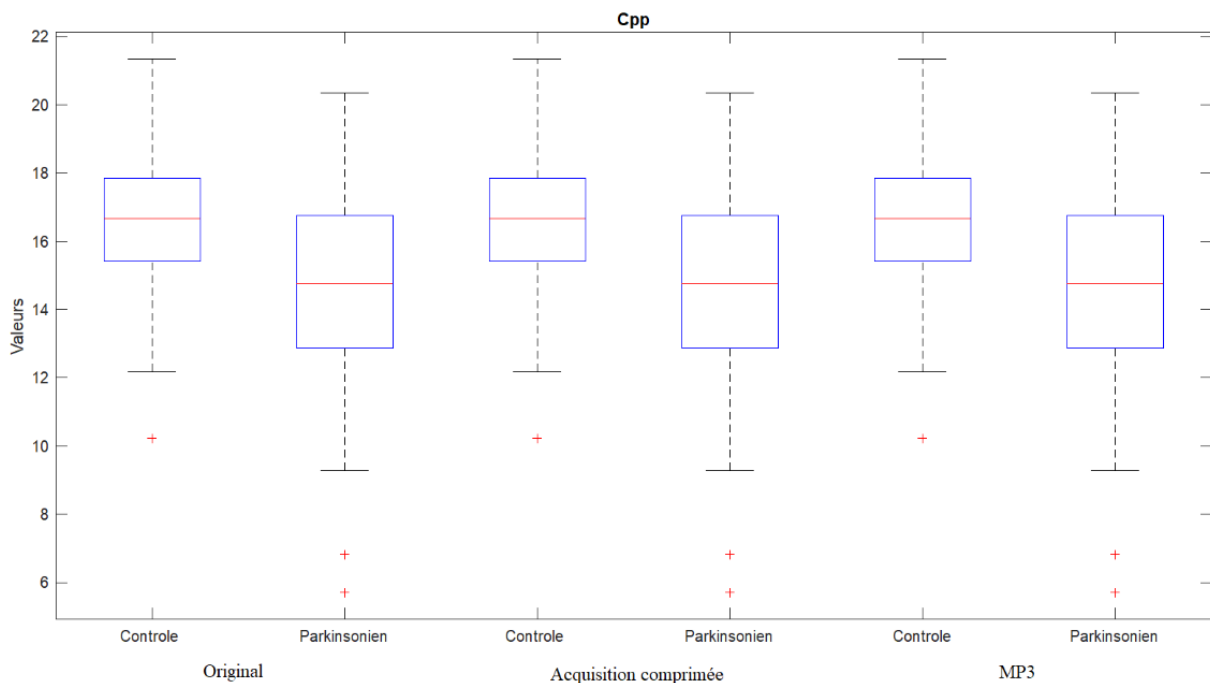
La figure IV.7 montre les percentiles des valeurs du HNR. On observe que les valeurs du HNR sont plus élevées pour les locuteurs contrôles que pour les locuteurs Parkinsoniens quel que soit le signal utilisé pour le calcul de l'indice acoustique. On observe une diminution des valeurs de l'indice HNR pour les signaux reconstruits par minimisation de la norme  $l_1$  par rapport aux signaux originaux aussi bien pour les locuteurs contrôles que pour les locuteurs Parkinsoniens. Cependant, l'efficacité du HNR en termes de discrimination contrôle/Parkinsonien n'est pas affectée. Les valeurs du HNR obtenues à partir des signaux MP3 décompressés sont quasi identiques aux valeurs obtenues à partir des signaux originaux. Le format MP3 utilise une compression audio avec perte, ce qui entraîne une diminution de la qualité audio. Cependant, les pertes sont souvent perceptuellement acceptables, ce qui signifie que les informations vocales importantes sont préservées.



**Figure IV-7** Boîte à moustaches du rapport harmonique sur bruit calculé à partir du signal original, et du signal reconstruit par minimisation  $l_1$  et du signal MP3 décompressé pour les locuteurs contrôles et Parkinsoniens.

#### IV.4.6 Proéminence du pic cepstral (CPP)

La figure IV.8 montre les percentiles des valeurs du CPP. On observe que les valeurs du CPP sont plus élevées pour les locuteurs contrôles que pour les locuteurs Parkinsoniens quel que soit le signal utilisé pour le calcul de l'indice acoustique. De plus, la dispersion des valeurs du CPP est plus faible chez les locuteurs contrôles, ce qui signifie que les valeurs du CPP se situent dans une plage plus restreinte pour les locuteurs contrôles comparés aux locuteurs Parkinsoniens. Les locuteurs Parkinsoniens montrent une plus grande variabilité et une plus grande étendue des valeurs du CPP. Par inspection visuelle, on constate que la discrimination contrôle/Parkinsonien est meilleure lorsque le CPP est calculé à partir du signal reconstruit par minimisation de la norme  $l_1$  comparé au signal MP3 décompressé.



**Figure IV-8** Boîte à moustaches de la proéminence du pic cepstral (CPP) calculée à partir du signal original, et du signal reconstruit par minimisation  $l_1$  et du signal MP3 décompressé pour les locuteurs contrôles et Parkinsoniens.

## **IV.5 Conclusion**

Dans ce chapitre, nous avons investigué les effets de la compression audio sur les paramètres acoustiques de la parole chez les contrôles et Parkinsoniens. À travers cette analyse, nous avons examiné les différences entre les locuteurs contrôles et Parkinsoniens en termes des indices de la fréquence fondamentale, du jitter, du tremor, du shimmer, du HNR et du CPP. Nous avons également évalué les effets de la compression audio sur ces paramètres, cherchant ainsi à comprendre son influence sur la qualité de la parole des patients.

Les résultats de cette étude ont révélé des variations significatives dans les paramètres acoustiques étudiés chez les locuteurs Parkinsoniens par rapport aux locuteurs contrôles. Ces différences mettent en évidence l'impact de la maladie de Parkinson sur la production de la parole et soulignent l'importance de l'évaluation de la dysarthrie dans cette population.

De plus, nous avons observé que la compression audio avait des effets variables sur les paramètres acoustiques de la parole chez les locuteurs Parkinsoniens. Cela soulève des questions importantes concernant l'adaptation des techniques de compression audio pour répondre aux besoins spécifiques de l'analyse de la parole dysarthrique.

Le travail de recherche réalisé met en évidence l'importance de prendre en compte la compression audio dans l'évaluation et la réhabilitation de la dysarthrie chez les locuteurs Parkinsoniens. Ces résultats fournissent des indications précieuses pour la détection, l'évaluation et la réhabilitation des locuteurs dysarthriques dans le contexte de la télésanté.

---

## *Conclusion générale*

---

## *Conclusion générale*

La télémédecine offre une solution prometteuse pour suivre l'évolution des patients souffrant de troubles de la voix et de la parole, en permettant des examens et des traitements à distance, malgré la distance géographique entre le fournisseur de soins de santé et le patient. L'avantage majeur de la télémédecine est la capacité de fournir des soins spécialisés à distance et d'offrir des services de santé aux patients qui ne pourraient pas autrement y accéder facilement. Cela ouvre de nouvelles opportunités pour un dépistage précoce des troubles de la voix, un suivi régulier et un traitement personnalisé, améliorant ainsi la qualité des soins et la satisfaction des patients.

L'acquisition comprimée, une méthode d'acquisition et de compression simultanées des signaux vocaux, est une approche prometteuse pour la compression des signaux de parole dans le cadre de l'évaluation des troubles de la voix et de la parole en télémédecine. Cette méthode utilise des algorithmes avancés pour reconstruire les signaux vocaux à partir d'un nombre réduit d'échantillons, tout en préservant les caractéristiques essentielles nécessaires à l'évaluation clinique.

L'effet de compression par la technique d'acquisition compressée a été étudié sur une base de données incluant des locuteurs contrôles et parkinsoniens. Nous avons calculé les indices acoustiques couramment utilisés pour la détection et l'évaluation de la dysarthrie à partir du signal reconstruit et en le comparant aux indices acoustiques obtenus à partir du signal original sans compression ainsi qu'aux indices acoustiques calculés à partir du signal MP3 décompressé. Les indices acoustiques sont la période fondamentale (F0), les perturbations rapides (jitter) et lentes (tremor) des durées des cycles, les perturbations de l'amplitude (shimmer), le rapport harmonique sur bruit (HNR : harmonics-to-noise ratio) et la prééminence du pic cepstral (CPP : cepstral peak proeminence). Les résultats de notre étude montrent que l'acquisition comprimée offre des performances meilleures ou comparables à celle de la méthode de compression MP3.

Il a été observé que la compression audio avait des effets variables sur les paramètres acoustiques de la parole chez les locuteurs Parkinsoniens. Cela soulève des questions importantes concernant l'adaptation des techniques de compression audio pour répondre aux besoins spécifiques de l'analyse de la parole dysarthrique.

Le travail de recherche réalisé met en évidence l'importance de prendre en compte la compression audio dans l'évaluation et la réhabilitation de la dysarthrie chez les locuteurs Parkinsoniens.

Des recherches futures pourraient étendre l'application de cette méthode à d'autres types de troubles de la parole et de la voix et à des bases de données de plus grande taille, ce qui renforcerait davantage sa validité et sa généralisation dans le domaine de la télémédecine.

---

*Références bibliographiques*

---

- [1] A. Castillo-Allendes, F. Contreras-Ruston, L. C. Cantor-Cutiva, et al., "Voice therapy in the context of the COVID-19 pandemic: guidelines for clinical practice," *J. Voice*, vol. 35, no. 5, pp. 717-727, 2021.
- [2] J. S. Sevitz, B. R. Kiefer, J. E. Huber, et al., "Obtaining objective clinical measures during telehealth evaluations of dysarthria," *Am. J. Speech Lang. Pathol.*, vol. 30, no. 2, pp. 503-516, 2021.
- [3] K. Weidner and J. Lowman, "Telepractice for adult speech-language pathology services: a systematic review," *Perspect. ASHA Spec. Interest Groups*, vol. 5, no. 1, pp. 326-338, 2020.
- [4] A. J. Hill, D. G. Theodoros, T. G. Russell, L. M. Cahill, E. C. Ward, and K. M. Clark, "An Internet-Based Telerehabilitation System for the Assessment of Motor Speech Disorders: A Pilot Study," *Am. J. Speech Lang. Pathol.*, vol. 15, no. 1, pp. 45, 2006.
- [5] A. J. Hill, D. G. Theodoros, T. G. Russell, et al., "The redesign and re-evaluation of an internet-based telerehabilitation system for the assessment of dysarthria in adults," *Telemed. e-Health*, vol. 15, no. 9, pp. 840-850, 2009.
- [6] G. Constantinescu, D. Theodoros, T. Russell, et al., "Assessing disordered speech and voice in Parkinson's disease: a telerehabilitation application," *Int. J. Lang. Commun. Disord.*, vol. 45, no. 6, pp. 630-644, 2010.
- [7] S. Fu, D. G. Theodoros, and E. C. Ward, "Delivery of intensive voice therapy for vocal fold nodules via telepractice: a pilot feasibility and efficacy study," *J. Voice*, vol. 29, pp. 696-706, 2015. [Online]. Available: <https://doi.org/10.1016/j.jvoice.2014.12.003>.
- [8] K. L. Dahl, H. R. Weerathunge, D. P. Buckley, et al., "Reliability and accuracy of expert auditory-perceptual evaluation of voice via telepractice platforms," *Am. J. Speech Lang. Pathol.*, vol. 30, no. 6, pp. 2446-2455, 2021.
- [9] H. R. Weerathunge, R. K. Segina, L. Tracy, et al., "Accuracy of acoustic measures of voice via telepractice videoconferencing platforms," *J. Speech Lang. Hear. Res.*, vol. 64, no. 7, pp. 2586-2599, 2021.
- [10] C. L. Payten, D. D. Nguyen, D. Novakovic, et al., "Telehealth voice assessment by speech language pathologists during a global pandemic using principles of a primary contact model: An observational cohort study protocol," *BMJ Open*, vol. 12, no. 1, p. e052518, 2022.



- [11] R. L. Utianski, S. Sandoval, V. Berisha, et al., "The effects of speech compression algorithms on the intelligibility of two individuals with dysarthric speech," *Am. J. Speech Lang. Pathol.*, vol. 28, no. 1, pp. 195-203, 2019.
- [12] Y. Yang, X. Yi, X. Zhao, et al., "A fast and secure MP3 steganographic scheme with multi-domain," *Signal Process.*, vol. 190, p. 108332, 2022.
- [13] K. Wang, C. Li, and L. Tian, "Audio zero watermarking for MP3 based on low frequency energy," in *2017 6th International Conference on Informatics, Electronics and Vision & 2017 7th International Symposium in Computational Medical and Health*, 2017.
- [14] R. Kumar and A. Kumar, "Compressed sensing based ECG signal compression for telemedicine," in *2015 IEEE Bombay Section Symposium (IBSS)*, Mumbai, India, pp. 1-4, 2015, doi: 10.1109/IBSS.2015.7456628.
- [15] D. L. Donoho, "Compressed sensing," in *IEEE Transactions on Information Theory*, vol. 52, no. 4, pp. 1289-1306, April 2006.
- [16] G. Peyré, "Best basis compressed sensing," in *IEEE Transactions on Signal Processing*, vol. 58, no. 5, pp. 2613-2622, May 2010.
- [17] X. Zhang, Z. Qian, Y. Ren, and G. Feng, "Watermarking with flexible self-recovery quality based on compressive sensing and compositive reconstruction," *IEEE Trans. Inf. Forensics Security*, vol. 6, no. 4, pp. 1223-1232, Dec. 2011.
- [18] W. Lu, Z. Chen, L. Li, et al., "Watermarking based on compressive sensing for digital speech detection and recovery," *Sensors*, vol. 18, no. 7, p. 2390, 2018.
- [19] B. Menia et L. Samer, *Compression des signaux de parole par acquisition comprimée pour les systèmes de télémedecine, mémoire de Master, Université de Jijel*, 2022.
- [20] C. Debilou et S. Boudaoud, *Amélioration d'un synthétiseur de la parole par concaténation, Mémoire de Fin d'Étude, Université Echahid Hamma Lakhdar El-Oued*, juin 2019.
- [21] A. Feniza et H. Laissaoui, *Analyse multivariée de la parole pour l'évaluation objective de la qualité vocale, Mémoire de Fin d'Étude, Université Mohammed Seddik BENYAHIA -JIJEL*, juillet 2019.
- [22] <http://club.cso.free.fr/sectiontechnique/cours/n4/n4anatomie1.htm>
- [23] G. Léothaud, "Théorie de la Phonation", *Cours de DEUG 2e année DMU3D1B*, 2004-2005.
- [24] <http://infirmier-freedom.blogspot.com/2015/05/le-larynx-et-cordes-vocales.html>

- [25] [http://staff.univ-batna2.dz/sites/default/files/haddadi\\_radhia/files/cours\\_ndeg\\_03\\_pca.pdf](http://staff.univ-batna2.dz/sites/default/files/haddadi_radhia/files/cours_ndeg_03_pca.pdf)
- [26] [https://www.ling.upenn.edu/courses/Fall\\_2013/ling520/LectureNotes2.html](https://www.ling.upenn.edu/courses/Fall_2013/ling520/LectureNotes2.html)
- [27] I. Jemâa, Suivi de formants par analyse en multirésolution, Thèse de doctorat, Université de Lorraine, Nancy, France, février 2013, ffNNT : 2013LORR0026f. [cs.OH].
- [28] E. A. Boualam, Synthèse de la Parole par méthode TD-PSOLA, thèse de doctorat, École Nationale Polytechnique, Algerie, 2013.
- [29] L. Cnockaert, Analyse du tremblement vocal et application à des locuteurs parkinsoniens, thèse de doctorat, Université libre de Bruxelles, Bruxelles, Belgique, 2008
- [30] S. Amiar, N. Ferhoun, Débruitage du signal de parole, mémoire de fin d'études, Université Larbi Ben M'Hidi de Oum El Bouaghi, 2012.
- [31] I. Larak, S. Affoun, Systèmes de Reconnaissance Automatique de la Parole, mémoire de Master, Université Akli Mohand Oulhadj – Bouira, 2017.
- [32] C. Fredouille, "Traitement Automatique et Troubles de la Voix et de la Parole : champs d'application, contraintes et limites," Informatique et langage, [cs.CL], Université d'Avignon et du Pays de Vaucluse, France, 2016.
- [33] M. Ghenay et R. Khelifi, Développement d'un système de détection de la voix pathologique, Mémoire de Master, Université Badji Mokhtar - Annaba, Annaba, Algérie, 2021.
- [34] M. Boudjerda, Analyse du signal de parole pour l'évaluation automatique des voix pathologiques, thèse, Université Mohammed Seddik BENYAHIA - Jijel, 2018.
- [35] I. Laaridh, Évaluation de la parole dysarthrique : Apport du traitement automatique de la parole face à l'expertise humaine, thèse de doctorat, Université d'Avignon, France, 2017, ffNNT : 2017AVIG0218ff.
- [36] F. Viallet and B. Teston, La dysarthrie dans la maladie de Parkinson, mémoire d'orthophonie, Université de Nantes, France, pp. 169-174, 2007. [En ligne]. Disponible sur: hal-00242418.
- [37] J. R. Orozco-Aroyave, E. A. Belalcazar-Bolaños, J. D. Arias-Londoño, J. F. Vargas- Bonilla, S. Skodda, J. Ruz, et al., "Characterization methods for the detection of multiple voice disorders: Neurological functional and laryngeal diseases", IEEE Journal of Biomedical and Health Informatics, vol. 19, no. 6, pp. 1820-1828, 2015.

- [38] C. Mertz, Étude des corrélations entre les évaluations subjectives et objectives de la voix chez des patients présentant une paralysie laryngée bilatérale, mémoire, Université de Rouen Normandie, 2018.
- [39] <https://www.fon.hum.uva.nl/praat/>
- [40] A. V. Oppenheim and R. W. Schaffer, "Digital signal Processing", Prentice-Hall, 1975.
- [41] E. S. Heller Murray, A. Chao, and L. Colletti, "A Practical Guide to Calculating Cepstral Peak Prominence in Praat," in press, 2022.
- [42] S. Stankovic', I. Orovic', and E. Sejdic', Multimedia Signals and Systems Basic and Advanced Algorithms for Signal Processing, 2nd edition, 2016.
- [43] A. Haleem, M. Javaid, R.P. Singh, R. Suman, "Telemedicine for healthcare: Capabilities, features, barriers, and applications," Sensors International, vol. 2, Jul. 24, 2021.
- [44] A. Tesnière, J.-P. Blanchère, A. Dompormartin, "Télédermatologie," EMC – Dermatologie, vol. 11, no. 1, pp. 1-7, 2015.
- [45] S. Pierre, A. Dominique, "La place de la télémédecine dans l'organisation des soins", rapport de la Direction Générale de l'Hospitalisation et de l'Organisation des Soins pour le Ministère de la santé et des Sports, France, 2008.
- [46] F. Kohler, Télémédecine, télésanté, eSanté, université Nancy, sante, France, 2010.
- [47] D. Florence, Fusion de données multicateurs pour un système de télésurveillance médicale de personnes à domicile, Thèse de Doctorat, Université Joseph-Fourier-Grenoble I, 2004.
- [48] D.G. Theodoros, A.J. Hill, Russell, T. Ward, E.C. and Wootton, R., "Assessing acquired language disorders in adults via the Internet," Telemedicine and e-Health, vol. 14, pp. 552-559, 2008.
- [49] S. Bouasli, A. Noumeri, compression et codage de la parole par la transformée KLT automatisée, Mémoire de Master, Systèmes de Télécommunications, Université Djilali Bounaama Khemis Miliana, 2016.
- [50] J. Raverdino, J.M. Guatteri, Techniques multimédia pour le son.
- [51] [https://olczyk.pagesperso-orange.fr/physique-chimie/Ens-Sc1ere/EnSc14.2%20Numérisation fichiers/MP3.pdf](https://olczyk.pagesperso-orange.fr/physique-chimie/Ens-Sc1ere/EnSc14.2%20Numérisation%20fichiers/MP3.pdf)
- [52] A. Ravelomanantsoa, Deterministic Approach for Compressed Sensing and Reconstruction of Signals from Distributed Intelligent Sensors, PhD thesis, Université de Lorraine, Nov. 9, 2015.

- [53] H. Wang, "Compressed Sensing: Theory and Applications," in *Journal of Physics: Conference Series*, vol. 2419, 2023
- [54] A. Sharma, "Compressive Sensing," *International Journal of Engineering Technologies and Management Research*, vol. 5, no. 2, pp. 249-255, 2018
- [55] E. J. Candès and T. Tao, "Near-Optimal Signal Recovery From Random Projections: Universal Encoding Strategies?" *IEEE Transactions on Information Theory*, vol. 52, no. 12, pp. 5406–5425, 2006.
- [56] H. Haneche, *Compressed Sensing in Mobile Systems*, Doctoral Thesis, Telecommunications and Information Processing, Electronics Department, University of Science and Technology Houari Boumediene, Algérie, 2020.
- [57] M. Khosravy, N. Dey, and C. A. Duque, Eds., *Compressive Sensing in Healthcare*, vol. 11, *Advances in ubiquitous sensing applications for healthcare*. Elsevier Science Publishing Co Inc, 2020.
- [58] J. E. Candès and M. B. Wakin, "An Introduction To Compressive Sampling," *IEEE SIGNAL PROCESSING MAGAZINE*, 2008
- [59] N. Tan, P. Tran, and M. Vozňák, "A Novel Compressed Sensing Approach to Speech Signal Compression," in *AETA 2015: Recent Advances in Electrical Engineering and Related Sciences*, pp. 75-85, March 2016.
- [60] E. Candès and J. Romberg, "1-magic: Recovery of Sparse Signals via Convex Programming, October 2005.
- [61] C. Moler, "Magic Reconstruction: Compressed Sensing," *MathWorks News&Notes, Cleve's Corner*, 2010.
- [62] R. L. Utianski, S. Sandoval, V. Berisha, et al., "The effects of speech compression algorithms on the intelligibility of two individuals with dysarthric speech," in *American Journal of Speech-Language Pathology*, vol. 28, no. 1, pp. 195-203, 2019.
- [63] C. Liu and P. Chang, "An efficient audio fingerprint design for MP3 music," in *Proceedings of the 9th International Conference on Advances in Mobile Computing and Multimedia*, pp. 190-193, 2011.
- [64] Y. Zhang, Y. Xiang, and L. Y. Zhang, *Secure Compressive Sensing in Multimedia Data*, *Cloud Computing and IoT*. Springer, 2018.
- [65] J. R. Orozco-Arroyave, E. A. Belalcazar-Bolaños, J. D. Arias-Londoño, J. F. Vargas-Bonilla, S. Skodda, J. Ruzs, et al., "Characterization methods for the detection of multiple voice disorders: Neurological functional and laryngeal

diseases", IEEE Journal of Biomedical and Health Informatics, vol. 19, no. 6, pp. 1820-1828, 2015

## Résumé

La télé-médecine est une pratique médicale qui permet la prestation de soins à distance en utilisant les technologies de l'information et de la communication. Elle offre aux patients la possibilité de consultations et de suivis médicaux à distance, notamment pour des troubles de la parole tels que la dysarthrie.

Ce mémoire explore l'application de l'acquisition comprimée de la parole dans les systèmes de télé-médecine pour la compression, en se concentrant sur la dysarthrie. L'effet de compression est étudié en comparant les indices acoustiques obtenus à partir du signal reconstruit à ceux calculés à partir du signal original et du signal compressé au format MP3. Les résultats de cette étude contribueront à évaluer la performance de cette approche dans l'évaluation de la dysarthrie en télé-médecine et à fournir des informations clés pour son application clinique.

**Mots clés :** acquisition comprimée, télé-médecine, indices acoustiques, compression du signal de parole, évaluation des troubles de la voix, dysarthrie.

## Abstract

Telemedicine is a medical practice that enables remote healthcare delivery using information and communication technologies. It provides patients with the opportunity for remote consultations and medical follow-ups, including for speech disorders such as dysarthria. This Master thesis explores the application of compressed sensing of speech in telemedicine systems for compression, with a specific focus on dysarthria.

The effect of compression is investigated by comparing the acoustic indices derived from the reconstructed signal to those calculated from the original and compressed MP3 signals. The findings from this study will contribute to evaluating the performance of this approach in dysarthria assessment within telemedicine, providing crucial insights for its clinical application.

**Key words:** compressed sensing, telemedicine, acoustic indices, speech signal compression, voice disorder evaluation, dysarthria.

## ملخص

الطب عن بُعد هو ممارسة طبية تتيح تقديم الرعاية عن بُعد باستخدام تكنولوجيا المعلومات والاتصالات. يتيح للمرضى إمكانية الاستشارات ومتابعة الرعاية الطبية عن بُعد، بما في ذلك لاضطرابات النطق مثل صعوبة النطق. تستكشف هذه الأطروحة تطبيق الحصول المضغوط للنطق في أنظمة الطب عن بُعد للضغط، مع التركيز على صعوبة النطق. يتم دراسة تأثير الضغط من خلال مقارنة المؤشرات الصوتية المستخلصة من الإشارة المُعيدة بتلك المُحسوبة من الإشارة الأصلية. سُسِّم نتائج هذه الدراسة في تقييم أداء هذا النهج في تقييم صعوبة النطق في الطب عن MP3 والإشارة المُضغوطة بتنسيق بُعد وتوفير معلومات رئيسية لتطبيقه السريري. الكلمات الرئيسية: الحصول المضغوط، الطب عن بُعد، المؤشرات الصوتية، ضغط إشارة النطق، تقييم اضطرابات الصوت، صعوبة النطق.

**الكلمات المفتاحية:** اكتساب مضغوط، الطب عن بُعد، مؤشرات صوتية، ضغط إشارة الكلام، تقييم اضطرابات الصوت، عسر النطق.