

الجمهورية الجزائرية الديمقراطية الشعبية
وزارة التعليم العالي والبحث العلمي

UNIVERSITÉBADJI MOKHTAR-ANNABA
BADJI MOKHTAR– ANNABAUNIVERSITY



جامعة باجي مختار – Annaba

Faculté : Sciences De L'ingéniorat

Département : Electronique

Domaine : Sciences Et Techniques

Filière : Télécommunications

Spécialité : Systèmes Des Télécommunications

Mémoire

Présenté En Vue De L'obtention Du Diplôme De Master

Thème : Réduction de dimension en vue de la détection de la voix pathologique

Présenté par: *Kadri Ibrahim Aimen&Debabi Abdeljalil Mohamed*

Encadrant: *Amara Fethi* Docteur Université : UBM Annaba

Jury de Soutenance :

F. Amara	Docteur	UBM ANNABA	Encadreur
K. Boukari	Docteur	UBM ANNABA	Examineur
L. Sahraoui	Docteur	UBM ANNABA	Président

Année Universitaire : 2020/2021

Remerciements

Nous remercions Dieu le tout puissant, qui nous a donné la force et la patience pour accomplir ce travail.

Nous adressons nos vifs remerciements à notre encadreur Mr Amara Fethi pour son aide, ses conseils, ses contributions, ses orientations précieuses et ses compétences.

Nous remercions vivement les membres du jury qui nous ont fait l'honneur d'accepter de juger notre travail.

N'oublions pas, bien sûr, nos professeurs de tout le cycle universitaire, de leur présenter nos meilleurs vœux.

Et enfin, nous tenons à remercier tous ceux qui ont contribué de près ou de loin, par leurs encouragements, leurs conseils et leurs soutiens à mener bien ce travail.

Dédicaces

Je dédie ce modeste travail

À mes chers Parents ma mère et mon père Pour leur patience, leur soutien et leur encouragement.

À mes grands-parents que je souhaite une bonne santé.

À mes chers frères.

À tous mes proches, mes amis, mes collègues de classe.

À mon binôme «AbdelJalile Mohamed » et à toute sa famille.

À notre promoteur Mr Amara Fethi.

À tous nos profs de la spécialité.

Et enfin, à tous ceux et celle que j'aime dans ma vie.

Kadri Ibrahim Aimen

Dédicaces

Je dédie ce mémoire :

A mes très chers Parents pour leur amour et encouragements qui n'ont jamais cessé de me motiver.

A mon cher frère Abderaouf pour son soutien moral.

A mon très cher oncle Fawzi pour son aide si précieuse tout au long de mes études.

A mes chers regrettés grands parents que je n'ai jamais cessé d'aimer.

A ma chère grand-mère Youma que j'aime tant.

A mes oncles et tantes paternels (elles) et maternels (elles)

A mon amie proche Manel qui a été toujours à mes côtés pour m'encouragé.

A mes cousins, cousines et tous mes amis (ies).

A mon binôme Ibrahim Aimen

Debabi Abdeljalil Mohamed

Résumé

L'analyse du signal de la parole est à la base de beaucoup d'applications. Ce mémoire s'inscrit dans le cadre des travaux de recherche qui ont pour objectif le développement d'un système de discrimination entre la voix normale et la voix pathologique à partir de ce signal. Ce mémoire est destiné à contribuer à l'amélioration des taux de détection par la combinaison de classificateur et les méthodes de réduction les plus répandus dans le domaine. Pour répondre à cet objectif, nous avons procédé en premier lieu à une caractérisation qualitative des voix pathologiques en utilisant les coefficients MFCC et leurs dérivées premières et deuxièmes. Ensuite, deux étapes d'expériences sont réalisées : dans la première étape nous nous intéressons à tester l'efficacité du SVM reconnu par sa capacité discriminative. Le classificateur a prouvé son efficacité dans ce contexte, le taux de reconnaissance 80%. La deuxième étape d'expériences concerne l'utilisation des méthodes de réduction pour améliorer les taux obtenus qui passent à 97%.

Abstract

Speech signal analysis is the basis of many applications. This thesis is part of research that aims to develop a system of discrimination between normal voice and pathological voice based on this signal. This dissertation is intended to contribute to the improvement of detection rates by the combination of classifier and the most answered reduction methods in the field. This thesis is intended to contribute to the improvement of detection rates by the combination of classifier and the most widely used reduction methods in the field. To meet this objective, we first carried out a qualitative characterization of pathological voices. using the MFCC coefficients and their first and second derivatives. Then, two experimental stages are carried out: in the first stage we are interested in testing the effectiveness of the SVM recognized by

its discriminative capacity. The classifier has proven its effectiveness in this context, the recognition rate 80%. The second stage of experiments concerns the use of reduction methods to improve the rates obtained, which increases to 97%.

ملخص

تحليل إشارة الكلام هو أساس العديد من التطبيقات. هذه الأطروحة جزء من بحث يهدف إلى تطوير نظام تمييز بين الصوت الطبيعي والصوت المرضي بناءً على هذه الإشارة. تهدف هذه الرسالة إلى المساهمة في تحسين معدلات الكشف عن طريق الجمع بين المصنف وأساليب التخفيض الأكثر إجابة في هذا المجال. لتحقيق هذا الهدف ، أجرينا أولاً توصيفاً نوعياً للأصوات المرضية باستخدام معاملات ومشتقات MFCC الأولى والثانية.

بعد ذلك ، يتم تنفيذ مرحلتين تجريبيتين: في المرحلة الأولى نحن مهتمون SVM المعترف بها من خلال قدرتها التمييزية. وقد أثبت المصنف فعاليته في هذا السياق بمعدل التعرف 80%. المرحلة الثانية من التجارب تتعلق باستخدام طرق الاختزال لتحسين المعدلات التي تم الحصول عليها والتي تصل إلى 97%.

Liste Des Abréviations

- TMP Temps maximum de phonation.
- GMM Gaussian Mixture Model.
- SVM Support Vectors machine.
- MFCC Mel Frequency Cepstral Coefficients.
- LPC Linear Predictive Coefficients.
- HNR Harmonic to noise ratio.
- ANN Artificial Neural Network.
- LDA Linear Discriminant Analysis.
- SVDS Saarbrücken Voice Database.
- PCA Principal Component Analysis.
- Drtoolbox Matlab Toolbox for Dimensionality Reduction

Listes Des Figures

Figure 1.1:	Anatomie de l'appareil phonatoire humain.....	4
Figure 2.1:	le cartilage cricoïde.....	6
Figure 3.1:	Vue sagittale du larynx et du Pavillon bucco-pharyngo-nasal.....	6
Figure 4.1:	le boitier laryngé.....	7
Figure 5.1:	Muscles Intrinsèques Du Larynx Coupe Axiale.....	8
Figure 6.1:	Bilan instrumental de la dysphonie (Changement de la voix).....	17
Figure 7.1:	Le système de reconnaissance.....	18
Figure 1.2:	Exemple de classification par SVM.....	25
Figure 2.2:	Exemple graphique de données linéairement séparables.....	26
Figure 3.2:	Exemple de transformation des données non linéaires dans un espace de dimension supérieure.....	27
Figure 4.2:	Organigramme Du Système D'aide Au Diagnostique.....	27

Listes des tableaux

Tableau 3.1 - Corpus de données « Homme ».....	38
Tableau 3.2 - Corpus de données « Femme ».....	38
Tableau 3.3 Corpus de données – Base de données SVD.....	38
Tableau3.4 : résultat obtenu sans réduction en utilisant SVM pour la base SVD « 13 Coefficients MFCC ».....	40
Tableau3.5 : résultat obtenu sans réduction en utilisant SVM pour la base SVD « 26 Coefficients MFCC ».....	40
Tableau3.6 : résultat obtenu sans réduction en utilisant SVM pour la base SVD « 39 Coefficients MFCC ».....	41
Tableau3.7: résultat obtenu avec réduction « LDA » pour la base SVD « 13 Coefficients MFCC ».....	41
Tableau3.8: résultat obtenu avec réduction « LDA » pour la base SVD « 26 Coefficients MFCC ».....	42
Tableau3.9: résultat obtenu avec réduction « LDA » pour la base SVD « 39 Coefficients MFCC ».....	42
Tableau3.10: résultat obtenu avec réduction « PCA » pour la base SVD « 13 Coefficients MFCC ».....	43

Tableau3.11: résultat obtenu avec réduction « PCA » pour la base SVD	
« 26 Coefficients MFCC ».....	43
Tableau3.12: résultat obtenu avec réduction « PCA » pour la base SVD	
« 39 Coefficients MFCC ».....	44
Tableau3.13 : résultat obtenu sans réduction en utilisant SVM pour la base HUPA	
« 13 Coefficients MFCC ».....	44
Tableau3.14 : résultat obtenu sans réduction en utilisant SVM pour la base HUPA	
« 26 Coefficients MFCC ».....	45
Tableau3.15 : résultat obtenu sans réduction en utilisant SVM pour la base HUPA	
« 39 Coefficients MFCC ».....	45
Tableau3.16 : résultat obtenu avec réduction « LDA » pour la base HUPA	
« 13 Coefficients MFCC ».....	46
Tableau 3.17 : résultat obtenu avec réduction « LDA » pour la base HUPA	
« 26 Coefficients MFCC ».....	46
Tableau 3.18 : résultat obtenu avec réduction « LDA » pour la base HUPA	
« 39 Coefficients MFCC ».....	47
Tableau 3.19 : résultat obtenu avec réduction « PCA » pour la base HUPA	
« 13 Coefficients MFCC ».....	47
Tableau 3.20 : résultat obtenu avec réduction « PCA » pour la base HUPA	
« 26 Coefficients MFCC ».....	48
Tableau 3.21 : résultat obtenu avec réduction « PCA » pour la base HUPA	
« 39 Coefficients MFCC ».....	48

Table Des Matières

Introduction générale

0.1 Contexte et motivation.....	1
0.2 Contributions.....	3
0.3 Organisation du manuscrit.....	5

Contents

1.1 Introduction.....	4
• 1.2 Appareil phonatoire	4
• 1.2.1 Poumons.....	5
• 1.2.2 Le larynx, organe vibrant.....	5
• 1.2.3 Cartilage du larynx.....	5
• 1.2.4 Structure des cordes vocales.....	7
• 1.2.5 Muscles Du Larynx.....	9
• 1.2.6 Rôle des cordes vocales.....	9
• 1.2.7 Les Cavités Résonantes.....	10

1.2 Production	de	la	parole	11
.....				
1.4	Troubles	de	la	voix
.....				
• 1.4.1	Définitions			12
• 1.4.2		Types	de	
	troubles			13
• 1.4.3	Tremblement vocal			13
• 1.4.2.1	Troubles organiques			13
• 1.4.2.2	Troubles fonctionnels			13
• 1.4.3			Symptômes	
	généraux			14
.....				
1.5 Méthodes	d'identification	des troubles		14
.....				
• 1.7.1	Évaluation	Perceptive		14
• 1.7.2	Evaluation	Instrumentale		16
• 1.7.2.1			Mesures	
	Acoustiques			16
.....				
1.6	Les		Pathologies	
Fonctionnelles				17
.....				
• 1.6.1		Vibrations	Non	
	Modales			17
• 1.6.2		Jitter	Et	
	Shimmer			18
.....				
1.8 Conclusion				19

2.1					
Introduction.....					20
2.2	Bases	de	données	des	voix
pathologiques.....					21
• 2.2.1	SVD « Saarbrucken Voice Database ».....				21
2.3	Travaux		selon		les
paramètres.....					21
• 2.3.1	Cepstre.....				22
• 2.3.1.1	LPC.....			Coefficients	23
• 2.3.1.2	MFCC.....			Coefficients	24
2.4		Mesures			de
bruit.....					24
2.5	travaux		selon	l’approche	de
classification.....					25
• 2.5.1	Séparateurs à vaste marge (SVM).....				25
• 2.5.1.1	Principe de fonctionnement.....				26
• 2.5.1.2	Méthode proposée.....				29
2.6	Méthodes	De	Réduction		De
Dimension.....					29
• 2.6.1	L’Analyse en Composantes Principales « PCA ».....				30

• 2.6.2	Analyse Discriminante Linéaire	
(LDA).....		32
2.7	Les systèmes	
hybrides.....		35
2.8 Conclusion.....		36
3.1 Introduction.....		37
3.2 Logiciel de développement.....		37
• 3.2.1 Conditions d'implémentation.....		37
• 3.2.2 Les Toolboxes.....		37
3.3 Méthodologie.....		38
3.4 Corpus de données.....		38
3.5 Extraction Des Paramètre MFCC.....		39
3.6 Résultats et discussion.....		40-48
3.7 Conclusion.....		49
3.8 Conclusion Générale		50

Introduction Générale

Contenu

• 0.1 Contexte et motivation	1
• 0.2 Contributions	3
• 0.3 Organisation du manuscrit	3

0.1 Contexte Et Motivation

La voix est un phénomène multidimensionnel composé d'un certain nombre d'éléments qui contribuent à sa qualité globale et à son intelligibilité. Une altération de la production de la parole est fréquemment représentée par une dysphonie ou une voix dysphonique. La dysphonie est une qualité perceptive de la voix qui indique que certains changements négatifs se sont produits dans les organes de phonation. Le terme dysphonie signifie littéralement voix anormale / difficile / altérée / voix pathologique. De nos jours, les troubles de la voix augmentent considérablement en raison du mode de vie moderne.

La relation entre la pathologie de la voix et les caractéristiques de la voix acoustique a été cliniquement établie et confirmée à la fois quantitativement et subjectivement par des experts de la parole. Les principales méthodes utilisées par la communauté médicale pour évaluer le système de production de la parole et diagnostiquer les pathologies sont soit des méthodes directes qui nécessitent une inspection directe des cordes vocales et provoquent une gêne pour le patient, soit des méthodes subjectives dans lesquelles la qualité de la voix est évaluée directement par un médecin expert. La détection et l'évaluation de l'état d'un patient est l'étape la plus cruciale et importante dans le diagnostic d'une pathologie. Pendant la consultation de nombreux problèmes peuvent être rencontrés par le médecin ou par le malade qui compliquent l'évaluation et conduisent à des mauvais résultats de diagnostic tel que : la difficulté de la prise de décision par la méthode de l'évaluation subjective, absences ou

inefficacité des moyens de diagnostic, complexité des cas (maladies rares chroniques et urgentes) et problème de suivi du patient (le coût des soins, incapacité de déplacement, mauvaise qualité de traitement, ...). Pour réduire le coût du diagnostic et aider les médecins à diagnostiquer avec précision les troubles vocaux, il y a eu une croissance récente de l'utilisation des techniques de traitement du signal vocal et de l'analyse des données pour détecter et diagnostiquer avec précision les individus. En raison de sa nature non invasive, l'évaluation automatique de la pathologie vocale est fortement considérée comme un outil de dépistage primaire ou un outil d'aide pour le clinicien. Un système d'évaluation automatique peut discriminer entre les échantillons normaux et pathologiques et de classer les pathologies de la voix. Le processus de différenciation entre les sujets normaux et pathologiques est un problème à deux classes appelé détection de pathologie. En revanche, la discrimination entre les différents types de pathologies est un problème multi-classes appelé classification des pathologies. [31]

La détection automatique et la classification des pathologies est un domaine d'actualité et toujours exploré par la communauté des chercheurs. Une large gamme de paramètres acoustiques a été utilisée pour la détection de pathologie à savoir le jitter, le shimmer, le rapport harmonique sur bruit (HNR : Harmonics to Noise Ratio), les coefficients cepstraux (MFCC : Mel-Frequency Cepstral Coefficients), etc. Dans le domaine de la détection automatique de la pathologie vocale, divers classificateurs ont été proposés tels que le perceptron Multi-Couches, le modèle de mélange gaussien, l'analyse discriminante linéaire, les machines à support de vecteurs (SVM : Support Vector Machine), etc.

0.2 Contributions

Le travail que nous présentons dans ce mémoire consiste à utiliser un système de détection automatique qui permet la discrimination entre la voix pathologique et la voix normale. Ce système classe les sons de la voix en utilisant les machines à vecteurs de supports (SVM). La décision finale consiste à indiquer s'il s'agit d'une voix saine ou pathologique. Les bases de données « Saarbrücken Voice Database - SVD- » et Hupa ont été utilisées dans la partie expérimentale. Les mesures utilisées dans cette étude sont : le taux de classification. Pour améliorer le taux obtenu deux méthodes de réduction ont été exploitées qui sont : la LDA et PCA.

Les résultats obtenus nous ont permis de donner des conclusions importantes sur les bases de données des voix pathologiques.

0.3 Organisation du manuscrit

Le manuscrit est composé, essentiellement, de cinq principaux chapitres :

Chapitre 1 : Ce chapitre rappelle les éléments essentiels à la compréhension des mécanismes de la production de la parole, les pathologies de l'appareil phonatoire et leurs influences sur la qualité de la voix. Nous discutons aussi les différentes méthodes d'évaluation.

Chapitre 2 : La première partie de ce chapitre concerne les travaux qui comportent des contributions sur la classification et l'autre partie traite les travaux qui s'appuient sur l'importance de la réduction de la dimension. Présenter l'architecture typique d'un détecteur de la voix pathologique. Nous décrivons en détails les différentes étapes qui composent ce système.

Chapitre 3 : Dans le troisième chapitre nous examinons l'efficacité du classificateur SVM et des méthodes de réductions de dimension PCA et LDA. Deux groupes d'expérimentation sont réalisés dont l'objectif est de trouver le meilleur modèle qui représente les données.

Contents

1.3 Introduction.....	4
• 1.2 Appareil phonatoire	4
• 1.2.1 Poumons.....	5
• 1.2.2 Le larynx, organe vibrant.....	5
• 1.2.3 Cartilage du larynx.....	5
• 1.2.4 Structure des cordes vocales.....	7
• 1.2.5 Muscles Du Larynx.....	9
• 1.2.6 Rôle des cordes vocales.....	9
• 1.2.7 Les Cavités Résonantes.....	10
1.4 Production de la parole.....	11
1.4 Troubles de la voix.....	12
• 1.4.1 Définitions.....	12
• 1.4.2 Types de troubles.....	13
• 1.4.3 Tremblement vocal.....	13

• 1.4.2.1 Troubles organiques.....	13
• 1.4.2.2 Troubles fonctionnels.....	13
• 1.4.3	Symptômes
généraux.....	14
1.5 Méthodes d'identification des troubles.....	14
• 1.7.1 Évaluation Perceptive.....	14
• 1.7.2 Evaluation Instrumentale.....	16
• 1.7.2.1	Mesures
Acoustiques.....	16
1.6	Les Pathologies
Fonctionnelles.....	17
• 1.6.1	Vibrations Non
Modales.....	17
• 1.6.2	Jitter Et
Shimmer.....	18
1.8 Conclusion.....	19

1.1 Introduction

Dans ce chapitre, nous présentons des généralités sur le signal de la parole, y compris l'anatomie de l'appareil phonatoire et le mécanisme de sa production. Ensuite nous discutons les principales pathologies qui atteignent cet appareil et leur influence sur l'aspect acoustique du signal de la parole. La dernière section est réservée à présenter les différentes méthodes d'évaluation de ces changements acoustiques.

1.2 Appareil Phonatoire

La parole est le langage articulé humain destiné à communiquer la pensée, et est à distinguer des communications orales diverses. La parole est singulière et opère un acte de langage qui s'adresse à un interlocuteur, éventuellement soi-même, mentalement, ou à un support par l'écrit par exemple. Les analyseurs de paroles transforment le signal porteur de la parole en une description de ses caractéristiques, sans se préoccuper de sa signification. L'appareil phonatoire ou appareil vocalique est l'ensemble des organes de la parole et des muscles qui les actionnent. Ils permettent la production des phones, ou sons propres à la langue parlée. Elle se compose de trois parties essentielles : les poumons, le larynx et les trois cavités résonantes. [1]

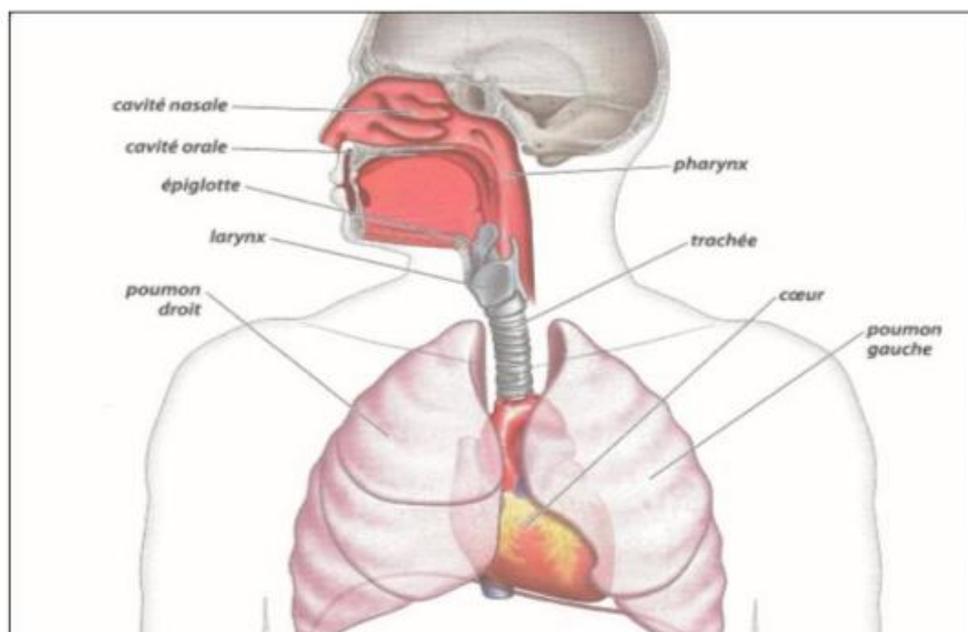


Figure 1.1 : Anatomie de l'appareil phonatoire humain.

1.2.1 Poumons

Les poumons sont des structures du système respiratoire, situés au sein de la cage thoracique. Au nombre de deux, les poumons sont situés dans le thorax, plus particulièrement au sein de la cage thoracique où ils occupent la majeure partie. Les deux poumons, droit et gauche, sont séparés par le médiastin, situé au centre et composé notamment du cœur.

Cavité pleurale. Chaque poumon est entouré par la cavité pleurale, qui est formée à partir de deux membranes : un feuillet interne, en contact avec le poumon, nommée plèvre pulmonaire ; un feuillet externe, en contact avec la paroi thoracique, nommée plèvre pariétale.

Cette cavité est composée d'un liquide séreux, le transsudat, permettant au poumon de glisser. L'ensemble permet également de maintenir le poumon et empêcher qu'il ne s'affaisse.

1.2.2 Le Larynx, Organe Vibrant

Organe de la phonation, puisqu'il joue un rôle très important dans l'émission des sons vocaux, le larynx est placé dans le cou à l'extrémité supérieure de l'arbre respiratoire. Cet instrument vibrant est placé dans le cou, sur le trajet de l'air respiratoire, entre la soufflerie qui commande l'expiration à soufflerie qui commande l'expiration, à savoir les poumons, et les cavités de résonance qui moduleront le son laryngé primaire.

1.2.3 Cartilage Du Larynx

Le larynx est formé d'une charpente cartilagineuse maintenue en place par une musculature interne et externe, et revêtue d'une muqueuse aux replis caractéristiques. Les cartilages qui le constituent s'articulent en bas à la trachée et se rattachent en haut au plancher buccal, par l'intermédiaire de l'os hyoïde. Directement en rapport avec les téguments du cou (pomme d'Adam), le larynx s'ouvre en arrière dans le pharynx, juste avant l'œsophage. Le larynx est constitué d'une série de cartilages

superposés. Le cartilage inférieur (le cartilage cricoïde) apparaît comme un anneau différencié de la trachée, fermée en arrière par une plaque verticale qui lui donne l'aspect d'une chevalière. Au bord du chaton de la chevalière, deux petites pyramides (cartilages aryténoïdes) glissent et pivotent ; elles servent de point d'attache postérieur aux cordes vocales.

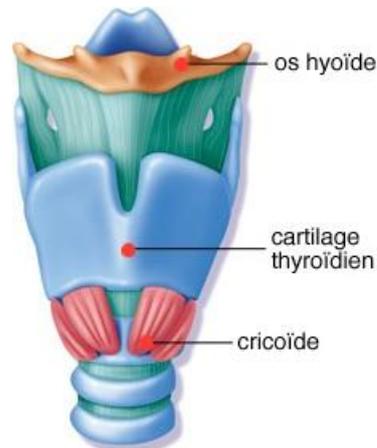


Figure 2.1 : Le cartilage cricoïde

Les cordes vocales s'insèrent en avant dans l'angle d'un volumineux cartilage (cartilage « thyroïde » en forme de bouclier ou de proue de navire, dont la saillie antérieure prend le nom de « pomme d'Adam ». En basculant sous ce bouclier, l'anneau cricoïdien entraîne avec lui les pyramides aryténoïdiennes ; ainsi, les deux points d'attache des cordes vocales s'éloignent, assurant leur mise en tension, les muscles tenseurs sont les muscles cricothyroïdiens.

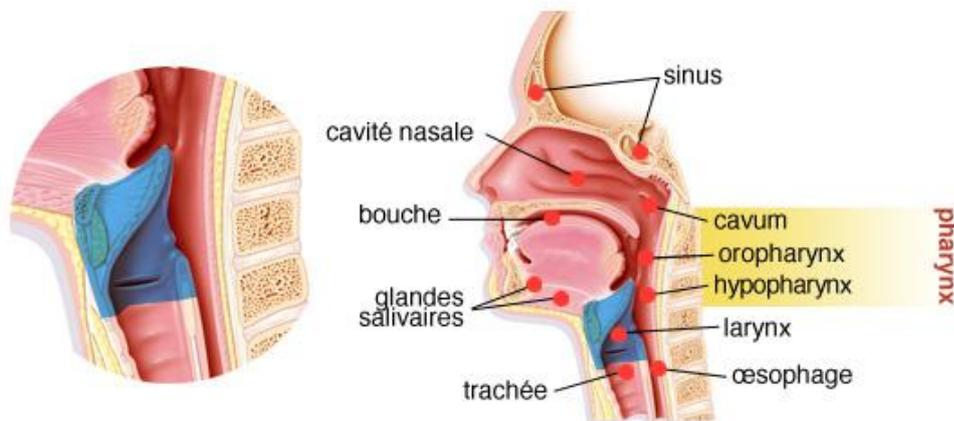


Figure 3.1: Vue sagittale du larynx et du pavillon bucco-pharyngo-nasal

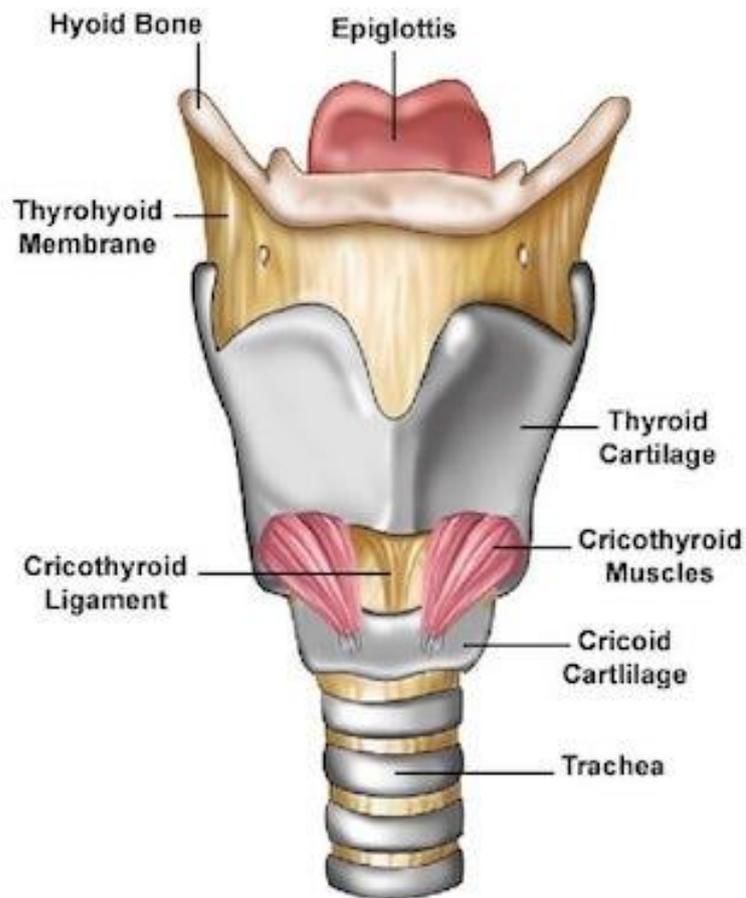


Figure 4.1 : Le boîtier laryngé

1.2.4 Structure des cordes vocales

Une corde vocale est la super position de deux muscles et d'un ligament Il ya tout d'abord, pour chaque corde vocale, un ligament vocal qui va du cartilage thyroïde à un cartilage aryténoïde. Le faisceau Ary-vocal est un muscle qui part du cartilage aryténoïde et va jusqu'au ligament vocal Le faisceau thyro-vocal est un autre muscle qui part du cartilage thyroïde pour se superposer au précédent. La superposition de ces deux faisceaux et du ligament vocal forme une corde vocale.

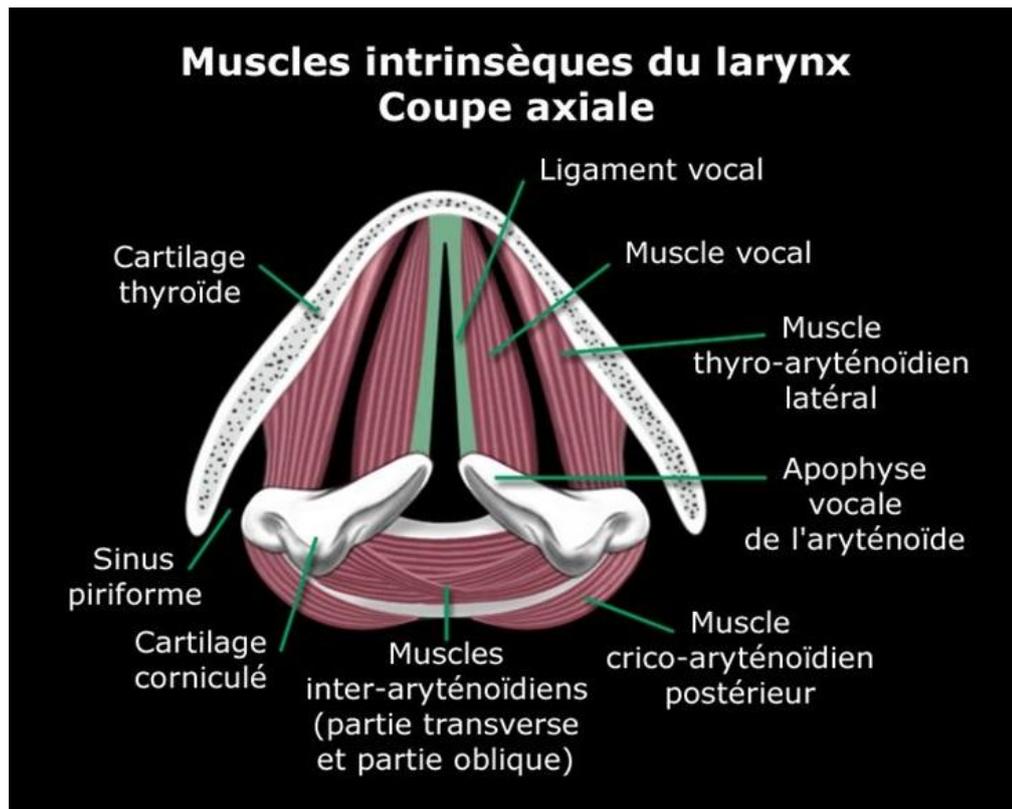


Figure 5.1 : Muscles Intrinsèques Du Larynx Coupe Axiale

Les cordes vocales sont donc formées par des ligaments vocaux, longés par des muscles sur lesquels on peut agir, le tout est recouvert d'une muqueuse qui constitue la partie vibrante qui produit le son. Elles forment un clapet qui peut être ouvert ou fermé.

On parle d'abduction lorsque les cordes vocales sont écartées et d'adduction lorsque les cordes vocales sont rapprochées. L'écart entre les cordes vocales porte le nom de fente glottique ou glotte. A l'avant les ligaments vocaux s'insèrent dans l'avant intérieur du cartilage thyroïde, à l'arrière elles s'insèrent sur les cartilages aryténoïdes.

Le Pivotement des cartilages aryténoïdes permet de rapprocher ou éloigner les cordes vocales, et donc de fermer ou d'ouvrir la glotte.

1.2.5 Muscles Du Larynx

Il y a de nombreux muscles qui permettent des mouvements entre les différents cartilages formant le larynx. Nous n'en distinguons que quelques-uns, indispensables à la compréhension de la phonation. Les muscles crico-aryténoïdes postérieurs (ou CAP), qui relient le cartilage cricoïde aux cartilages aryténoïdes par l'arrière : les contractions de ces muscles crico-aryténoïdes postérieurs rapprochent les apophyses arrière des aryténoïdes et déportent vers l'extérieur les apophyses antérieures (ou vocales), ce qui a pour effet d'écarter les cordes vocales et d'ouvrir complètement la glotte ; ces muscles sont donc responsables du mouvement d'abduction (écartement) des cordes vocales.

Les muscles crico-aryténoïdes latéraux (ou CAL) qui relient les mêmes cartilages par les côtés. Les contractions de ces muscles crico-aryténoïdes latéraux tirent vers l'avant les apophyses arrière des aryténoïdes et ces muscles permettent donc l'adduction (rapprochement) des cordes vocales et donc la phonation.

1.2.6 Rôle des cordes vocales

Les cordes vocales sont donc tendues de l'angle rentrant du cartilage thyroïde à l'apophyse vocale des cartilages aryténoïdes. La glotte est l'espace plus ou moins grand entre les deux cordes vocales. Les cordes vocales ont trois positions fondamentales :

- Soit, elles sont écartées : la glotte est ouverte et l'air circule librement. C'est la respiration. Lors d'une inspiration profonde, l'écartement est maximal, lors de la respiration normale, l'écartement est moyen.
- Soit, elles sont accolées : la glotte est alors fermée et l'air ne passe pas. C'est l'apnée.

- Soit, les cordes sont rapprochées : la glotte est variable. C'est la phonation ou le voisement.

L'adduction est la position requise pour la vibration des cordes vocales (phonation).

Quatre conditions doivent être réunies pendant la phonation :

- Adduction (rapprochement)des cordes vocales
- Tension longitudinale des cordes vocales (réalisée par un mouvement de bascule du thyroïde sur le cricoïde)
- Compression latérale des cordes
- Pression sous glottique (elle doit être équivalente à 3cm d'eau au minimum)

La vibration des cordes vocales n'est pas une succession d'abductions et d'adductions : ce qui s'ouvre et se ferme c'est uniquement la muqueuse, les ligaments sont en contact pendant toute la phonation. Les cordes vocales se comportent comme une valve bloquant le passage de l'air.

1.2.7 Les Cavités Résonantes

La majorité des sons du langage sont le fait du passage d'une colonne d'air venant des poumons, qui traverse un ou plusieurs résonateurs de l'appareil phonatoire

Le pharynx (ou cavité pharyngale) est un conduit musculo- membraneux situé entre la bouche et l'œsophage d'une part et entre les fosses nasales et le larynx d'autre part. La paroi du pharynx est constituée de muscles constricteurs. Effet d'une constriction : modification du diamètre du pharynx. La racine de la langue peut également reculer ou avancer et donc agir sur le volume de cette première cavité supra- glottique.

Les fosses nasales (ou cavités nasales) sont deux cavités cunéiformes séparées par une cloison verticale médiane et sont recouvertes de muqueuses. Une résonance nasale est très caractéristique (nasallement). L'air passe par le nez lorsque le voile du

palais (prolongement) musculaire du palais osseux) est abaissé : passage oro- nasal ouvert.

La bouche (ou cavité buccale) est séparée des fosses nasales par une cloison appelée le palais. Dans cette cavité se situent des articulateurs, certains fixes (passifs), d'autres mobiles (actifs).

La cavité labiale est une cavité que l'on crée lorsqu'on projette en avant les lèvres (protrusion labiale).

1.3 La Production De La Parole

La production de la parole est un processus qui transforme les pensées en parole. Cette activité comprend la sélection des mots, l'organisation des formes grammaticales pertinentes et l'articulation des sons par le système moteur via l'appareil vocal. Cette production peut être spontanée (par exemple, quand une personne crée des mots lors d'une conversation), réactive (par exemple, quand elle identifie une illustration ou qu'elle fait une lecture à voix haute) ou imitative (quand elle répète les mots qu'une autre personne a dit). La production de la parole se distingue de la production langagière puisque le langage peut être produit à l'aide de signes manuscrits.

Habituellement, la parole est produite grâce à une pression appliquée sur les poumons puis par la modulation du courant d'air qui entre dans le canal vocal, ensemble de mouvements coordonnés qui permettent la phonation. Toutefois, la production de la parole est possible sans utiliser les poumons et la glotte ; le discours alaryngé utilise les parties hautes du canal vocal. La production de la parole est régulièrement accompagnée de la production de gestes qui a pour fonction de faciliter la compréhension du discours.

1.4 Troubles de la voix

- **1.4.1 Définitions**

Un trouble de la voix est un changement dans la qualité de la voix normale pendant une longue période de temps. Si la qualité de la voix d'un locuteur X a changé et est resté ainsi pendant plus de deux semaines, c'est très probable que cette personne est atteinte d'un trouble de la voix. Le nom technique de cette maladie est la dysphonie (dis-foh-nee-a). [2]

La plupart des personnes atteintes de troubles de la voix décrivent leur voix comme étant:«rauque», « fragile » ou « faible ». Beaucoup de gens disent aussi qu'ils ont un sentiment d'oppression / tension, la sécheresse ou une boule dans la gorge.

Les troubles de la voix peuvent se produire à tout le monde, mais sont plus fréquents chez les personnes qui utilisent leurs voix lourdement. Cela inclut des groupes comme les enseignants, les agents de centre d'appel, les chanteurs et surtout aussi chez les fumeurs.

Parfois, les gens disent que leur voix a empiré après une infection à la poitrine ou d'un stress événement et est resté mauvais. Un trouble de la voix peut être causé par des conditions à long terme qui affectent la respiration (par exemple asthme ou maladie pulmonaire chronique). IL peut également se produire à la suite d'un accident

vasculaire cérébral ou une maladie qui affecte le cerveau, comme la maladie de Parkinson.

Parfois, les enfants auront des troubles de la voix, soit à cause de quelque chose qu'ils sont nés avec, ou à cause de la façon dont ils utilisent leur voix.[3]

1.4.2 Types de troubles

Les troubles de la voix se répartissent en trois grandes catégories : organiques, fonctionnels, ou une combinaison des deux.[4]

- **1.4.2.1 Troubles organiques**

Les troubles organiques de la voix se répartissent en deux groupes : structurels et neurologiques.

1. Structurels

Les troubles structurels sont causés par une lésion du larynx (anomalie physique).

2. Neurologiques

Les troubles neurologiques vocaux sont causés par un problème dans le système nerveux qui interagit avec le larynx. Lorsque le nerf est endommagé, il provoque une parésie (faiblesse) ou de paralysie (absence totale de mouvement) dans le pli vocal du côté affecté. D'autres troubles de la voix neurologiques sont liés à d'autres types de problèmes dans le 16 système nerveux central (Par exemple, la maladie de Parkinson, Dysphonie spasmodique, la myasthénie, la maladie de la SLA / Lou Gehrig).

- **1.4.2.2 Troubles fonctionnels**

Les troubles fonctionnels sont causés par le mauvais fonctionnement des muscles. Tous les troubles fonctionnels relèvent de la catégorie des dysphonies de la tension musculaire. (Par exemple la dysphonie de tension musculaire (général), constriction antéropostérieure, hyperabduction).

1.4.3 Symptômes généraux

L'adulte ou l'enfant présente un ou plusieurs de ces symptômes :

- Sensation de fatigue vocale
- Besoin de faire des efforts pour parler
- Manque de souffle
- Sensation de boule dans la gorge
- Douleurs
- Présence de sécrétions
- Raclement fréquent de la gorge
- Toute autre sensation inconfortable

1.5 Méthodes d'identification des troubles

Il existe deux méthodes d'identification des troubles de la voix : méthode subjective et méthode objective.

• 1.5.1 Évaluation perceptive

L'évaluation perceptive repose sur le jugement humain et sur les capacités de l'auditeur à évaluer la qualité d'une voix. Sa fiabilité est souvent controversée.

Kreiman et coll montrent un niveau de reproductibilité entre 18 et 100%, avec des protocoles et méthodes d'analyses statistiques si différents que la comparaison des résultats est presque impossible.[5]

Une méthodologie stricte et reproductible doit être mise en place pour améliorer la fiabilité de l'évaluation perceptive et pallier à son caractère subjectif. Elle passe par le choix des variables qualitatives, puis par la mise en place d'une échelle quantitative. Dans la pratique clinique, la nature d'une voix est décrite en termes d'impression auditive : érailement, raucité, souffle, craquements, etc.

Le choix de ces termes a été largement discuté et débattu et différents essais de standardisation des méthodes d'évaluation ont été rapportés dans la littérature. Bien qu'il n'existe actuellement pas de protocole internationalement reconnu et accepté, l'échelle GRBAS proposée par la « Japanese society of logopedics and phoniatrics » et développée par Hirano 2 est la plus largement utilisée au monde. Il s'agit d'une échelle à cinq paramètres où le G (Grade) définit le grade global de la dysphonie, le R (Rough), la raucité, le B (Breathy), le caractère soufflé, le A (Asthenic), l'asthénie et le S (Strained), le forçage.

Dans la mise en place d'un protocole d'évaluation perceptive, le choix du jury est également essentiel. Sa qualité est évaluée en termes de fiabilité qui correspond à la reproductibilité du jugement entre les auditeurs (agrément inter auditeur) et par l'auditeur lui-même, lors de plusieurs sessions d'écoute (agrément intra-auditeur). De nombreux auteurs soulignent l'importance du niveau d'expérience des auditeurs et du recrutement de plusieurs membres afin d'obtenir un groupe homogène. Afin d'améliorer la fiabilité de l'évaluation, plusieurs sessions d'écoute sont organisées pendant lesquelles les voix, un texte lu standardisé, sont passées « à l'aveugle » dans un ordre aléatoire. L'effet de contexte est un facteur d'influence du jury : une voix moyennement dysphonique paraît plus altérée si elle est présentée après une voix normale qu'après une dysphonie sévère.

L'analyse perceptive est la seule capable de rendre les caractéristiques perceptives de la voix et est donc irremplaçable. De plus, elle est facile à mettre en œuvre, accessible à tout clinicien et peu coûteuse. Par contre, elle est chargée de plusieurs biais intrinsèques qui la rendent imparfaite et insuffisante : de nombreux facteurs influençant le jugement perceptif ne peuvent être contrôlés, tels que l'état émotionnel de l'auditeur au moment de l'évaluation, ses valeurs esthétiques, sa langue maternelle, la manière dont il conçoit l'échelle de mesure, etc. C'est pourquoi, plusieurs équipes dirigent leurs travaux dans le but de développer une évaluation instrumentale objective des dysphonies permettant de comparer les résultats vocaux de façon fiable et reproductible.

- **1.5.2 Evaluation instrumentale**

L'analyse instrumentale est conçue pour qualifier et surtout quantifier les dysphonies à partir de mesures physiques du son, acoustiques et aérodynamiques. Elles sont réalisées sur une voyelle tenue, en général le /a/, à l'aide de différents capteurs. Il est nécessaire de combiner différentes mesures complémentaires afin de tenir compte de l'aspect multidimensionnel de la production vocale.

- **1.5.2.1 Mesures acoustiques**

Elles sont effectuées à l'aide d'un microphone et révèlent les caractéristiques audibles de la dysphonie. Il s'agit principalement des mesures de la fréquence et de l'intensité, de leur stabilité, ainsi que de l'analyse du spectre du son.

La fréquence fondamentale (F0) correspond au nombre de vibrations par seconde de la masse vibrante des cordes vocales, exprimée en Hertz (Hz). Elle donne la hauteur tonale de la voix. La F0 dépend de la masse vibrante, de la tension des cordes vocales et de la pression d'air sous-glottique.[15]

L'intensité est dépendante de la PSG et de la résistance glottique au flux d'air. Sa mesure est donc un reflet de la biomécanique laryngée, des forces musculaires

laryngées et respiratoires. Une intensité excessive peut contribuer à l'apparition de lésions des cordes vocales (nodules, polypes).

Le jitter et le shimmer sont des mesures de la stabilité du signal acoustique, en fréquence pour le premier et en intensité pour le second. Une limitation importante de ces mesures est l'impossibilité de calculer un jitter ou un shimmer fiable à partir d'un signal vocal trop instable, en cas de dysphonie sévère.[23]

L'analyse spectrale est une méthode instrumentale reposant sur le théorème de Fourier qui stipule que tout son complexe peut être décomposé en une série de sons sinusoïdaux (les harmoniques) dont la fréquence est un multiple de la F0. L'analyse du rapport signal/bruit permet de quantifier le bruit présent dans le spectre et indique ainsi le degré de sévérité de la dysphonie.

1.6 Les Pathologies Fonctionnelles

Les troubles de la voix peuvent toucher n'importe qui, mais sont plus répandus chez les individus qui font un usage vocal important dans leur vie quotidienne. Par exemple, les enseignants sont des professionnels particulièrement à risque de rencontrer des difficultés par rapport à leur voix. Les troubles de nature « fonctionnelle » sont causés par une utilisation non efficace de la voix et peuvent mener à des extinctions de voix à répétition ainsi qu'à des lésions sur les plis vocaux. Les conséquences de ce changement sont les troubles de la voix appelés dysphonies vocales.

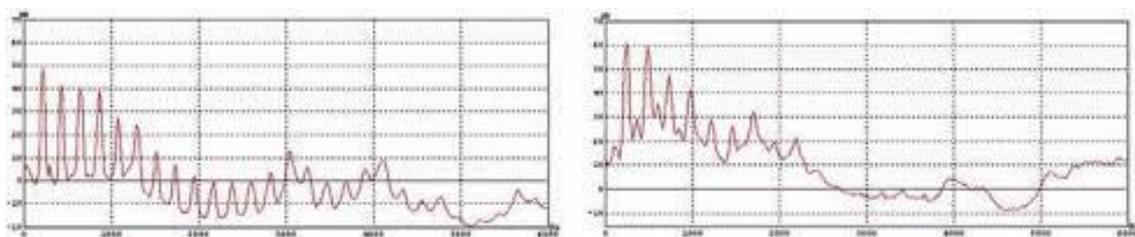


Figure 6.1 : Bilan instrumental de la dysphonie (Changement de la voix)

1.6.1 Vibrations non modales

La voix humaine est souvent ramenée à sa source sonore primaire, comme résultante de la mise en vibration des plis vocaux au sein du larynx. Bien que cette vision de la voix soit très réductrice, elle souligne le rôle fondamental que joue la vibration des plis vocaux dans la Production vocale humaine.

1.6.2 Jitter et Shimmer

Le Jitter ou Gigue mesure l'irrégularité avec laquelle les paquets d'un même signal (d'une même conversation) arrivent à destination. La gigue vocale (**vocal jitter**) désigne les petites perturbations rapides des durées des cycles glottiques. Les origines sont inconnues. [24] L'amplification de la gigue vocale est parfois observée en présence de pathologies. Le Jitter est donné par la formule :

$$jitter = \frac{1/n - 1 \sum_{k=1}^N |T_k - T_{k+1}|}{1/N \sum_{k=1}^N |T_k|}$$

T_k = La période de la k'ime cycle.

Le shimmer est une « variation d'amplitude entre périodes successives quand un individu tente de tenir la phonation à une fréquence et une intensité constantes ». Une légère modulation de 1 dB joue sur l'impression de raucité du son. Le shimmer est donné par la formule : A_k est l'amplitude de la k'ime cycle. [25]

$$shimmer = \frac{1/n - 1 \sum_{k=1}^N |A_k - A_{k+1}|}{1/N \sum_{k=1}^N |A_k|}$$

1.7 Différentes méthodes d'évaluation

L'évaluation de la voix est un élément central de la consultation phoniatrique. Elle est axée d'une part sur une évaluation perceptive, à l'oreille, de la qualité de la voix et, d'autre part, sur une analyse instrumentale basée sur des mesures acoustiques et aérodynamiques du son. L'évaluation perceptive, bien que considérée comme le gold standard, possède une certaine subjectivité ; l'analyse instrumentale a pour avantage d'être objective mais possède également une fiabilité limitée. C'est pourquoi, une analyse multiparamétrique, intégrant les deux méthodes d'évaluation, permet d'évaluer la dysphonie selon plusieurs angles d'observation.

Dans le domaine de la phonétique clinique, l'évaluation de la qualité vocale apparaît nécessaire pour établir un bilan vocal. Analyser des cas pathologiques, évaluer un bilan thérapeutique, comparer et distinguer entre différents cas pathologiques.

1.8 Conclusion

En conclusion, Après tout ce que nous avons présenté dans ce chapitre, nous pouvons dessiner que la phonation est un phénomène complexe, dont les mécanismes biomécaniques sont encore incomplètement cernés. Une altération de la voix peut entraîner d'importantes perturbations dans la communication, mais aussi dans l'identité et la représentation sociales du patient. Son évaluation clinique est l'axe central autour duquel s'organise la prise en charge phoniatrique et logopédique et comprend une évaluation perceptive et une analyse instrumentale. Ces deux méthodes sont complémentaires et irremplaçables et permettent d'évaluer la dysphonie selon plusieurs angles d'observation.

CHAPITRE 2 :

Reconnaissance De La Voix Pathologique À Partir Du Signal

De La Parole

Contenu

2.1 Introduction.....	20
2.2 Bases de données des voix pathologiques.....	21
• 2.2.1 SVD « Saarbrucken Voice Database ».....	21
2.3 Travaux selon les paramètres.....	21
• 2.3.1 Cepstre.....	22
• 2.3.1.1 Coefficients LPC.....	23
• 2.3.1.2 Coefficients MFCC.....	24
2.4 Mesures de bruit.....	24
2.5 travaux selon l’approche de classification.....	25
• 2.5.1 Séparateurs à vaste marge (SVM).....	25
• 2.5.1.1 Principe de fonctionnement.....	26
• 2.5.1.2 Méthode proposée.....	29
2.6 Méthodes De Réduction De Dimension.....	29
• 2.6.1 L’Analyse en Composantes Principales « PCA ».....	30
• 2.6.2 Analyse Discriminante Linéaire (LDA).....	32
2.7 Les systèmes hybrides.....	35
2.8 Conclusion.....	36

2.1 Introduction

La détection et la classification automatique de la voix pathologique à partir du signal de la parole est la convergence de plusieurs disciplines, aussi variées que celles de la médecine, la linguistique, la phonétique physiologique, la phonétique acoustique, et le domaine de l'intelligence artificielle. La médecine permet de définir la nature de la maladie, déterminer le grade de la dysphonie, établir la différence par rapport aux autres maladies, fournir des consignes lors de la création de la base de données. La linguistique permet de définir les unités linguistiques, phonèmes, mots ou énoncés, contenant des informations pertinentes sur la qualité de la voix. En d'autres mots, dans certains cas pathologiques, la dysphonie apparaît clairement lorsque le patient prononce une unité linguistique bien déterminée.

La phonétique physiologique est la discipline qui s'intéresse à définir comment un son est produit, quels sont les organes intervenants, durée de phonation, comme elle s'intéresse aussi à reconnaître la manière dont ce son est perçu. Puis vient la phonétique acoustique pour déterminer les paramètres caractérisant la dysphonie. Enfin, un problème de classification de la voix pathologique est peut-être résolu en utilisant les techniques de l'apprentissage automatique, commençons par le choix des paramètres, la réduction de dimension et la conception des modèles en utilisant les classificateurs les plus appropriés afin d'assurer l'efficacité et la robustesse du système. Dans la majorité des travaux, un système de détection automatique de la voix pathologique suit une architecture générale.

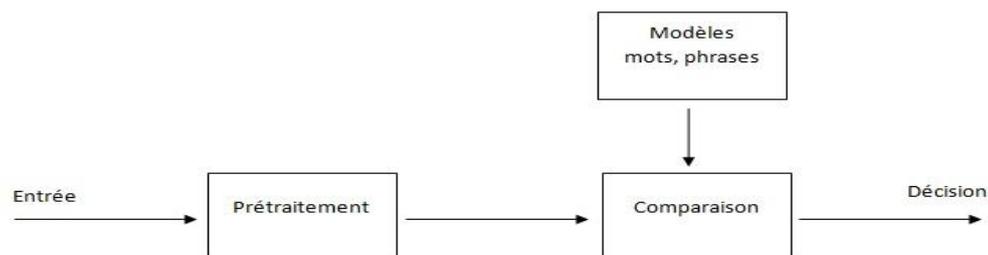


Figure 7.1: Le système de reconnaissance

2.2 Bases de données des voix pathologiques

Plusieurs bases de données ont été utilisées dans la littérature pour la tâche de classification des voix saines et pathologiques. Nous pouvons les classer en données publiques et privées. Les bases de données publiques peuvent être librement ou commercialement distribuées pour la communauté scientifique. Cependant, les bases de données privées ne sont pas disponibles puisqu'elles ont été enregistrées localement.

Comme il est impossible de comparer les résultats de la recherche lorsque les expériences sont réalisées avec des données privées, nous pouvons affirmer que les données publiques sont plus utilisées car elles permettent des comparaisons rapides.

- **2.2.1 SVD « Saarbrücken Voice Database »**

Cette base de données est une collection d'enregistrements de voix de plus de 2000 personnes [12]. Elle est disponible gratuitement en ligne. Une session d'enregistrement contient :

- Les voyelles /a/, /i/ et /u/ produit à un ton normal, haut et bas.
- Une phrase en allemand « GutenMorgen, wiegeht es Ihnen? » qui signifie « Bonjour, comment vas-tu? ».
- L'Electroglottograph (EGG) de tous les enregistrements

. La durée des voyelles varie entre 1 et 3 secondes. La Fréquence d'échantillonnage des enregistrements est de 50 kHz. Environ 71 pathologies différentes sont incluses dans cette base de données.[18]

2.3 Travaux selon les paramètres

- **Paramètres du domaine temporel**

- **Fréquence fondamentale :** La fréquence fondamentale F_0 est définie comme le nombre de vibrations des cordes vocales par seconde. Dans le domaine temporel, c'est la période d'un signal voisé à un instant donné. Pour le signal de la parole, sa fréquence fondamentale n'est rien d'autre que la fréquence du cycle d'ouverture/fermeture des cordes vocales. Elle dépend de la masse vibrante, de la tension des cordes vocales déterminée par les muscles qui contrôlent celles-ci. Elle varie d'un locuteur à un autre en fonction de son âge et de son sexe. Elle s'étend

approximativement de 80 à 200 Hz chez les hommes, de 150 à 450 Hz chez les femmes, et de 200 à 600 Hz chez les enfants [Boite 2000] La détection de F0 joue un rôle essentiel dans le domaine de traitement de la parole et notamment pour l'évaluation de la voix pathologique. C'est un témoin des propriétés biomécaniques des cordes vocales et la configuration laryngée.[14]

- **Perturbation de la fréquence fondamentale et de l'amplitude** : Les cycles glottiques naturellement ne sont pas parfaitement périodiques et la présence d'une pathologie augmente considérablement ces apériodicités. Deux phénomènes qui apparaissent : le vacillement (Jitter) et le tremblement (Shimmer). Le vacillement est défini par les variations trame par trame dans les périodes de F0. Le tremblement représente les variations cycle par cycle dans les périodes de l'énergie. Le jitter et le shimmer sont utilisés pour l'évaluation de la voix pathologique. Une étude comparative entre différents algorithmes d'évaluation a été réalisée par Darcio G. Silva et al. [Silva 2009], L'objectif c'était l'estimation exacte de ce paramètre. L'algorithme proposé jitter local présente une meilleure performance par rapport au jitter estimé par le logiciel PRAAT et MDVP voire même par rapport au jitter estimé par l'algorithme STJE (Short Time Jitter Estimation) proposé par M. Vasilakis and Y. Stylianou.

2.3.1 Cepstre

Le couplage entre l'excitation glottale et le conduit vocal, rend difficile la mesure précise du fondamental. Pour remédier à ce problème, une approche originale connue par le cepstre a été proposée par Noll (1964) pour extraire le fondamental. Oppenheim (1968) l'a exploitée pour les systèmes de codage/décodage. Finalement, on trouve les systèmes d'analyse et de la reconnaissance automatique qui ont bénéficié largement de cette technique. Le formalisme mathématique de cette méthode peut être consulté dans l'ouvrage [Oppenheim and Schafer, 1975]. Tel que mentionné, la parole peut être

considérée comme étant le résultat d'une convolution entre la source d'excitation et le conduit vocal.

Le formalisme mathématique d'homomorphisme, permet de transposer l'opérateur convolution "*", en opérateur addition "+" dans le domaine transposé. C'est le principe de base utilisé par le cepstre pour déconvoluer le signal vocal. [7]

Les $s(n)$ sont les échantillons du signal temporel avant toute transformation.

L'avantage du cepstre, est son insensibilité au problème des formants(leur confusion avec les harmoniques dans le contexte de détection de fréquence glottale) et de la distorsion de la phase. Il permet par un simple filtrage passe bas (simple fenêtre rectangulaire) de récupérer la contribution vocale dans le domaine Fréquentiel. Dans le cas des voisés, le fondamental est celui ayant le pic le plus élevé. Par contre l'inconvénient de la méthode du cepstre est sa sensibilité au bruit et à la taille de la fenêtre d'analyse. En plus la mesure du fondamental n'est pas toujours précise.

2.3.1.1 Coefficients LPC

Le signal de parole présente une corrélation à court terme induite principalement par les cavités buccales et aussi une corrélation à long terme induite par la périodicité du signal d'excitation.[19] La corrélation à court terme est traduite dans le spectre par la structure des formants (l'enveloppe du signal) et la corrélation à long terme est traduite dans le spectre par une structure fine en peigne dite harmonique [Fort 1996]. Dans la littérature nous ne trouvons pas beaucoup de travaux qui utilisent les coefficients à prédiction linéaire LPC comme paramètres pour caractériser la voix pathologique. Ces coefficients sont utilisés dans D.G. Childers 1992[29] avec la méthode de quantification vectorielle comme classificateur. Ils ont trouvé un taux de reconnaissance de 82.9%. [13]

2.3.1.2 Coefficients MFCC

Les coefficients MFCC sont calculés en utilisant l'échelle de mel (linéaire pour les fréquences inférieures à 1000 et logarithmique pour les fréquences supérieures à 1000). Cette échelle est connue pour prendre compte de la perception humaine. Ces coefficients ont la propriété d'être décorrélés. La description de l'algorithme de l'extraction de ces paramètres est bien détaillée dans la section 4.4. Les coefficients MFCC sont utilisés dans plusieurs domaines de traitement du signal de la parole tel que la reconnaissance de la parole Schafer. 2009[28], la synthèse K. Tokuda 2002[30] et la reconnaissance des émotions [Y. Attabi]. La détection et la classification de la voix pathologique ne fait pas exception à la règle, dans les travaux [Godino-Llorente 2004], [J. I. Godino-Llorente 2005] et [J. I. Godino-Llorente 2006] les coefficients MFCC présentent l'entrée des classificateurs ANN, SVM, GMM respectivement dont l'objectif était la discrimination entre la voix normale et la voix pathologique. Ces paramètres sont utilisés aussi pour l'évaluation du grade de la dysphonie selon l'échelle de GRBAS. Dans les travaux [C. Fredouille 2005], [G. Pouchoulin 2006], 16 coefficients MFCC et leurs dérivées sont proposés d'être l'entrée du GMM. Chaque grade est représenté par un modèle.

2.4 Mesures de bruit

Dans certains cas pathologiques, les cordes vocales perdent leurs caractéristiques dynamiques et leur fermeture devient incomplète. [8] L'air provenant des poumons devient turbulent et se superpose sur la composante périodique au niveau de la glotte. En terme d'évaluation perceptuelle, ce bruit présente la source de la raucité dans la voix selon les auteurs [A. Kacha 2006]. Dans la littérature nous pouvons trouver trois types de bruit :

- **Rapport harmonique/bruit (HNR)** : Yumoto [Yumoto 1982] a été l'un des premiers qui a proposé une technique simple de calcul du rapport Harmonique/Bruit (H/N en Anglais). Pour le groupe normal, ce rapport est centré sur 12 dB (harmoniques significativement plus énergétiques que le bruit) et peut devenir négatif pour des pathologies sévères (harmoniques noyées dans le bruit). Plusieurs études sont dédiées à l'estimation exacte du HNR [F. Severin 2005]. Ce paramètre est peut-être mieux estimé dans le domaine spectral que dans le domaine temporel. [26]
- **Energie de bruit normalisée (NNE)** : Ce paramètre est proposé [H. Kasuya 1986] pour quantifier l'énergie de bruit dans le signal de la parole. Les auteurs supposent que la voix pathologique est plus bruitée que la voix normale.
- **Energie de bruit glottique (GNE)** : Ce paramètre est défini comme le rapport entre la quantité du signal produit par les cordes vocales et le bruit de turbulence au niveau de la glotte [D. Michaelis]. Ce paramètre est supposé comme un bon indicateur sur le degré de la raucité. [9]

2.5 Méthodes De Réduction De Dimension

Il est évident que les performances d'un système de reconnaissance s'améliorent par l'utilisation des coefficients différentiels du premier et second ordre (Δ et $\Delta\Delta$). Cependant, ces coefficients entraînent un triplement de la taille des vecteurs acoustiques et manquent de discrimination au niveau de ces paramètres. Il est donc préférable de ne conserver que les coefficients discriminants et réduire la redondance de l'information présente.

2.5.1 L'Analyse en Composantes Principales PCA

La méthode d'analyse en composantes principales (ACP ou PCA) est une technique linéaire de réduction de dimension, qui signifie qu'il réduit la dimensionnalité en incorporant les données dans un sous-espace linéaire de dimensionnalité inférieure.

C'est une procédure mathématique qui transforme un nombre de variables corrélées en

Un nombre (plus petit) de variables non corrélées appelées composantes principales. L'objectif est de réduire l'espace d'attribut d'un plus grand nombre de variables à un plus petit nombre de facteurs.

2.5.2 Analyse Discriminante Linéaire LDA :

L'analyse discriminante linéaire est définie comme une transformation linéaire orthogonale qui sépare au mieux deux ou plus classe d'objets []. Il trouve l'ensemble de la projection des vecteurs qui maximisent le rapport entre les classes dispersion contre dispersion dans la classe. La combinaison résultante peut être utilisée pour la classification. LDA trouve un sous-espace de dimension inférieure dans lequel le rapport entre la variance interclasse et la variance intra-classe est maximisée. C'est-à-dire qu'une transformation linéaire

$$W_{opt} = \arg \max_w \frac{\text{trace}[w^T S_B W]}{\text{trace}[w^T S_w W]}$$

Discriminante **Wopt** est le maximum du critère suivant : Dont **Sb** et **Sw** sont inter-classe et intra-classe matrices de covariance, respectivement, définies comme suit :

- $S_w = \frac{1}{N} \sum_{k=1}^L \sum_{x_i \in C_k} (x_i - u_k)(x_i - u_k)^T$, $u_k = \frac{1}{N} \sum_{x_i \in C_k} x_i$
- $S_b = \frac{1}{N} \sum_{k=1}^L N_k (u_k - u)(u_k - u)^T$
- $\Gamma_c = \sum_{i \in S_c} \bar{X}_i / |S_c|$,
- $\Sigma_c = \sum_{i \in S_c} (\bar{X}_i - \Gamma_c)(\bar{X}_i - \Gamma_c)^T / |S_c|$

2.6 Classification :

Les classificateurs peuvent se diviser en deux grandes catégories fondamentalement différentes. La première englobe les méthodes dites "génératives" telles que les modèles de mélanges gaussiennes et modèles de Markov cachés et la deuxième inclue les méthodes dites "discriminantes" telles que les k-voisins plus proches, les réseaux de neurones et les séparateurs à vaste marge.

2.6.1 Séparateurs à vaste marge (SVM)

Les machines à vecteurs supports « En Anglais : Support Vector Machines (SVMs) » (ou séparateurs à vaste marge) proposées par Vapnik [27] en 1995 sont un ensemble de techniques d'apprentissage destinées à résoudre des problèmes de :

Classification (Discrimination) : c'est-à-dire décider à quelle classe appartient un échantillon
Régression : c'est-à-dire prédire la valeur numérique d'une variable. Dans notre étude nous nous focalisons sur la classification. Les SVM sont devenus un outil populaire pour la classification discriminante. Un domaine passionnant de l'application récente des SVM est dans le traitement de la parole. Ces modèles ont une stratégie de modélisation nettement différente dans la détection de troubles de la voix, par rapport à d'autres méthodes de classification reportées dans la littérature.[6]

2.6.1.1 Principe de fonctionnement

La méthode SVM suit une procédure particulière pour trouver la séparation des hyperplans avec la plus grande marge entre deux classes. Dans la forme la plus simple, où les deux classes sont linéairement séparables comme représentée sur la Figure 3.3, cette méthode essaie de trouver un hyperplan discriminatif $w \cdot x + b = 0$ Avec la plus grande distance possible entre les données d'apprentissage des deux

classes. Tel que : w : Vecteur des poids ($w_1, w_2, w_3, \dots, w_n$). X : Vecteur des attributs ($x_1, x_2, x_3, \dots, x_n$). b : Seuil du séparateur linéaire.[17]

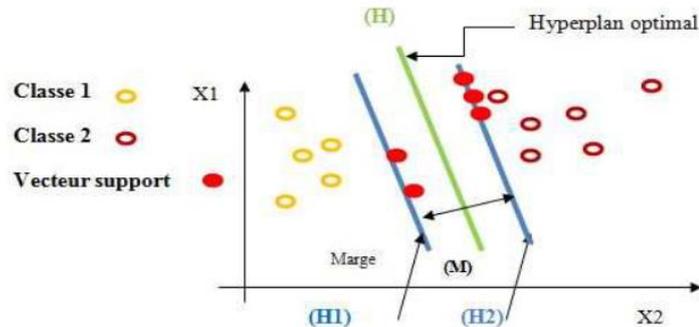


Figure 1.2 : Exemple de classification par SVM

• Hyperplan

On appelle hyperplan séparateur (H) un hyperplan qui sépare les points d'apprentissage de deux classes données (1 et 2). Les deux hyperplans (H1) et (H2) sont appelés hyperplans canoniques.

La détermination de l'hyperplan est basée seulement sur les points situés sur la frontière entre les deux classes de données (parmi l'ensemble total d'apprentissage). Ces points sont appelés vecteurs de supports.[22]

• Marge

Il existe une infinité d'hyperplans capable de séparer parfaitement les deux classes d'exemples. Le principe des SVM est de choisir celui qui va maximiser la distance minimale entre l'hyperplan et les exemples d'apprentissage (i.e la distance entre l'hyperplan et les vecteurs supports), cette distance est appelée la marge. Le fait d'avoir une marge plus large procure plus de sécurité lorsqu'on classe un nouvel.

- La distance de l'hyperplan de séparation est : $1/\|w\|$
- La marge est égale $2/\|w\|$
- La distance de l'hyperplan à l'origine est égale à : $b/\|w\|$

- $\|w\|$ fait référence a la norme du vecteur w .

SVM Linéaire

Un classificateur est dit linéaire lorsqu'il est possible d'exprimer sa fonction de décision par une fonction linéaire. Dans la suite, nous supposons que nos exemples sont donnés dans un format vectoriel. Notre espace d'entrée x est composé de n composantes.

Si les données sont linéairement séparables, alors il existe un hyperplan d'équation :

- $\langle w.x \rangle + b = 0$
- $\langle w.x \rangle + b \geq 1$ si $y_i = +1$
- $\langle w.x \rangle + b \leq -1$ si $y_i = -1$

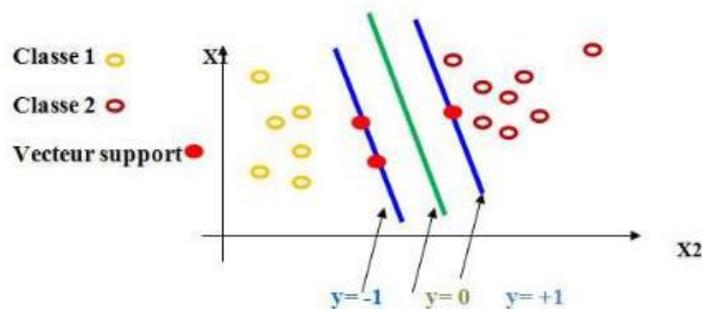


Figure 2.2 : Exemple graphique de données linéairement séparables

- **SVM Non Linéaire**

Le SVM non linéaire transforme l'espace d'entrée d'origine dans un espace caractéristique de dimension supérieure. En introduisant une fonction noyau (Kernel) (ϕ) dans le nouvel espace de dimension supérieure (Figure 3.5).

Dans cet espace, il est plus probable que les données ont une distribution linéaire. L'équation de l'hyper plans est la suivante: $H(x) = w \cdot \phi(x) + b$ (3.8)

Vapnik a montré que toute fonction satisfaisant les conditions d'admissibilité peut être utilisée comme noyau. [11]

Parmi les noyaux les plus utilisés pour la classification, nous citons : Le noyau linéaire, le noyau polynomial et le noyau Gaussien (Radial Basis Function -RBF)

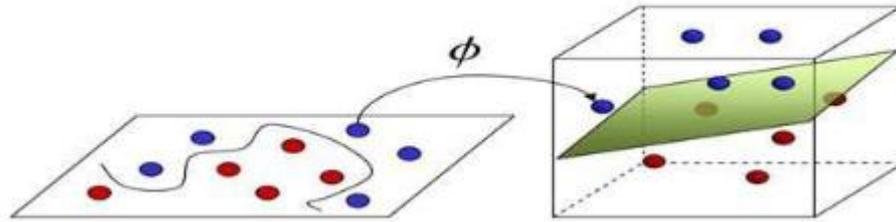


Figure 3.2 : Exemple de transformation des données non linéaires dans un espace de dimension supérieure

Les etapea suivre dans notre travail se résume en :

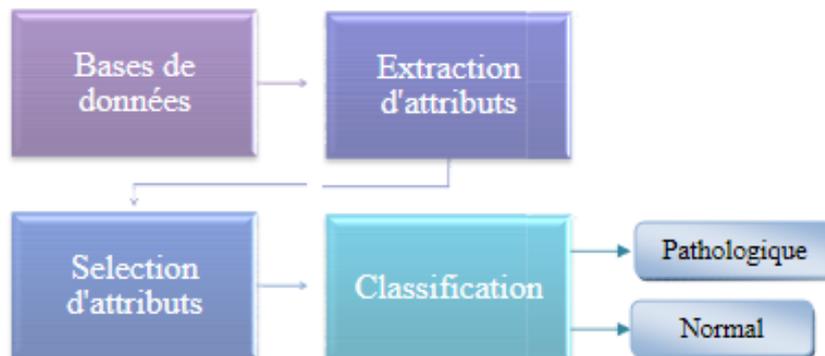


Figure 6.2 : Organigramme Du Système D'aide Au Diagnostique

2.8 Conclusion

Dans ce chapitre nous avons montré que la détection et la classification de la voix pathologique à partir du signal de la parole a attiré l'attention du chercheur le long des deux dernières décennies. Malgré qu'il y ait un grand nombre de travaux et qu'il y a une variété dans les méthodes l'objectif était l'amélioration des performances du système de détection. La première partie de ce chapitre concerne les travaux qui comportent des contributions sur la classification et l'autre partie traite les travaux qui s'appuient sur l'importance de la réduction de la dimension. Présenter l'architecture typique d'un détecteur de la voix pathologique. Nous décrivons en détails les différentes étapes qui composent ce système.

Chapitre 3 :

Expérimentations, Résultats Et Interprétations

Contenu

3.3 Introduction.....	37
3.4 Logiciel de développement.....	37
• 3.2.1 Conditions d'implémentation.....	37
• 3.2.2 Les Toolboxes.....	37
3.3 Méthodologie.....	38
3.4 Corpus de données.....	38
3.5 Extraction Des Paramètre MFCC.....	39
3.6 Résultats et discussion.....	40-48
3.7 Conclusion.....	49
3.8 Conclusion Générale	50

Chapitre 3 : Expérimentations, Résultats Et Interprétations

3.1 Introduction

Dans ce chapitre, nous exposerons les résultats obtenus par l'application de la méthode proposée en vue de la distinction entre la voix normale et la voix pathologique par le système de détection automatique. Deux bases de données différentes ont été utilisées SVD et HUPA. La première partie de notre travail consiste à faire une classification avec les SVM's. Dans la deuxième partie, et afin d'améliorer les résultats nous avons utilisés les deux méthodes de réduction LDA et PCA.

Dans ce qui suit, nous décrivons la méthodologie suivie et nous présenterons les expérimentations et les résultats obtenus.

3.2 Logiciels De Développement

- **Conditions d'implémentation**

Nos simulations et tests ont été faites sous MATLAB version 2010 (Matlab R2010a) qui est un environnement de calcul technique conçu pour le calcul numérique et la visualisation à haute performance. Cette version contient plusieurs fonctions et outils dédiés à la classification des signaux.[16]

- **Les Toolboxes suivantes ont été utilisées :**

Drtoolbox: Actuellement « Matlab Toolbox for DimensionalityReduction » contient les techniques suivantes : LDA, PCA. En plus boîte à outils contient des implémentations de 6 techniques d'estimation de la dimensionnalité intrinsèque, ainsi que des fonctions d'extension hors échantillon, de pré blanchiment des données et de génération d'ensembles de données de jouets.[20]

Libsvm : « Library for SVM » : c'est une bibliothèque pour les machines à vecteurs de support.[21]

3.3 Méthodologie

Nous avons utilisé deux différentes bases des données SVD et Hupa, qui ont des ensembles de voix (maladie et voix pour la base SVD), mais n'a pas déterminé le sexe de la voix, mais la deuxième fois que nous avons utilisé une base de données HUPA, elle a également un ensemble de voix et le genre de la voix qu'il contient.

3.4 Corpus de données

Dans notre travail nous avons choisis deux classes : une classe normale et une classe pathologique. Pour la classe pathologique nous avons utilisé des patients qui souffrent de la dysphonie spasmodique, la laryngite et la dysodie. Ces maladies sont des maladies qui touchent le larynx. Toutes les données sont divisées en : données d'apprentissage et données de test. Chaque locuteur prononce la lettre « A » comme le montre les tableaux

Genre	Nature	Apprentissage	Test	UnitéD'analyse
Femme	Normale	34	20	Lettre "A"
	Pathologique	30	20	

Tableau 3.2: Corpus de données HUPA « Femme »

Genre	Nature	Apprentissage	Test	UnitéD'analyse
Homme	Normale	25	10	Lettre "A"
	Pathologique	20	15	

Tableau3.3: Corpus de données – Base de données SVD

3.5 Extraction Des Paramètres MFCC :

Pour l'extraction des paramètres acoustiques, nous envisageons l'utilisation d'une plus large gamme d'information spectrale du signal acoustique permettant de mieux caractériser la classe pathologique, en utilisant un vecteur des 13 premiers

coefficients, augmentés par les dérivées (Δ) premières et deuxièmes ($\Delta\Delta$) représentant la vitesse et l'accélération respectivement. Sachant que, les treize premiers coefficients MFCC, leurs dérivées, fournissent toute l'information et aucun gain dans les performances n'est enregistré par l'augmentation du nombre des coefficients (Huang, Acero et Hon, 2001). Les différents calculs pour cette sélection sont effectués sur un PC Lenovo d'un processeur I5 3eme génération et d'une mémoire RAM de 8Go.

Pour l'extraction des coefficients MFCC sous MATLAB, les paramètres suivants sont utilisés :

- Une durée de trame d'analyse de 25 ms
- Un coefficient $\alpha = 0.95$ de préaccentuation
- Un nombre des coefficients MFCC = 13, 26, 39
- Un nombre $M = 20$ de banc de filtres à échelle de Mel

3.6 Résultats et discussion

➤ Base de données SVD

Décision \	Normale	Pathologique
Normale	77%	23%
Pathologique	25%	75%

**Tableau 3.4: résultat obtenu sans réduction en utilisant SVM pour la base SVD
« 13 Coefficients MFCC »**

➤ Pour 13 coefficients :

- Les cas reconnus normales avec un taux de 77% et les cas non reconnus avec un taux de 23%
- Les cas reconnus pathologique avec un taux de 75% et les cas non reconnus avec un taux de 25%

Décision \	Normale	Pathologique
Normale	78%	22%
Pathologique	25%	75%

**Tableau 3.5: résultat obtenu sans réduction en utilisant SVM pour la base SVD
« 26 Coefficients MFCC »**

➤ Pour 26 coefficients :

- Les cas reconnus normales avec un taux de 78% et les cas non reconnus avec un taux de 22%
- Les cas reconnus pathologique avec un taux de 75% et les cas non reconnus avec un taux de 25%

Décision \	Normale	Pathologique
Normale	80%	20%
Pathologique	20%	80%

Tableau 3.6: résultat obtenu sans réduction en utilisant SVM pour la base SVD
« 39Coefficients MFCC »

➤ **Pour 39 coefficients :**

- Les cas reconnus normales avec un taux de 80% et les cas non reconnus avec un taux de 20%
- Les cas reconnus pathologique avec un taux de 80% et les cas non reconnusavec un taux de 20%.

• **Les résultats après la réduction**

Décision \	Normale	Pathologique
Normale	78%	22%
Pathologique	20%	80%

Tableau 3.7: résultat obtenu avec réduction « LDA » pour la base SVD
« 13Coefficients MFCC »

➤ **Pour 13 coefficients :**

- Les cas reconnus normales avec un taux de 78% et les cas non reconnus avec un taux de 22%
- Les cas reconnus pathologique avec un taux de 80% et les cas non reconnus avec un taux de 20%

Décision \	Normale	Pathologique
Normale	80%	20%
Pathologique	20%	80%

Tableau 3.8: résultat obtenu avec réduction « LDA » pour la base SVD
« 26Coefficients MFCC »

➤ **Pour 26 coefficients :**

- Les cas reconnus normales avec un taux de 80% et les cas non reconnus avec un taux de 20%
- Les cas reconnus pathologique avec un taux de 80% et les cas non reconnus avec un taux de 20%

Décision \	Normale	Pathologique
Normale	84%	16%
Pathologique	14%	86%

Tableau 3.9: résultat obtenu avec réduction « LDA » pour la base SVD
« 39Coefficients MFCC »

➤ **Pour 39 coefficients :**

- Les cas reconnus normales avec un taux de 84% et les cas non reconnus avec un taux de 16%
- Les cas reconnus pathologique avec un taux de 86% et les cas non reconnus avec un taux de 14%.

Décision \	Normale	Pathologique
Normale	72%	28%
Pathologique	26%	74%

Tableau 3.10: résultat obtenu avec réduction « PCA » pour la base SVD
« 13Coefficients MFCC »

➤ **Pour 13 coefficients :**

- Les cas reconnus normales avec un taux de 72% et les cas non reconnus avec un taux de 28%
- Les cas reconnus pathologique avec un taux de 74% et les cas non reconnus avec un taux de 26%

Décision \	Normale	Pathologique
Normale	77%	23%
Pathologique	25%	75%

Tableau 3.11: résultat obtenu avec réduction « PCA » pour la base SVD
« 26Coefficients MFCC »

➤ **Pour 26 coefficients :**

- Les cas reconnus normales avec un taux de 77% et les cas non reconnus avec un taux de 23%
- Les cas reconnus pathologique avec un taux de 75% et les cas non reconnus avec un taux de 25%

Décision \	Normale	Pathologique
Normale	80%	20%
Pathologique	25%	75%

Tableau 3.12: résultat obtenu avec réduction « PCA » pour la base SVD
« 39Coefficients MFCC »

➤ **Pour 39 coefficients :**

- Les cas reconnus normales avec un taux de 80% et les cas non reconnus avec un taux de 20 %
- Les cas reconnus pathologique avec un taux de 75% et les cas non reconnus avec un taux de 25%.

• **Base de données HUPA**

Décision \	Normale	Pathologique
Normale	85%	15%
Pathologique	18%	82%

Tableau 3.13: résultat obtenu sans réduction pour la base HUPA
« 13Coefficients MFCC »

➤ **Pour 13 coefficients :**

- Les cas reconnus normales avec un taux de 85% et les cas non reconnus avec un taux de 15%

- Les cas reconnus pathologique avec un taux de 82% et les cas non reconnus avec un taux de 18%.

	Normale	Pathologique
Décision		
Normale	90%	10%
Pathologique	18%	82%

Tableau 3.14: résultat obtenu sans réduction pour la base HUPA

« 26Coefficients MFCC »

➤ **Pour 26 coefficients :**

- Les cas reconnus normales avec un taux de 90% et les cas non reconnus avec un taux de 10%
- Les cas reconnus pathologique avec un taux de 82% et les cas non reconnus avec un taux de 18%

	Normale	Pathologique
Décision		
Normale	90%	10%
Pathologique	10%	90%

Tableau 3.15: résultat obtenu sans réduction pour la base HUPA

« 39Coefficients MFCC »

➤ **Pour 39 coefficients :**

- Les cas reconnus normales avec un taux de 90% et les cas non reconnus avec un taux de 10 %
- Les cas reconnus pathologique avec un taux de 90% et les cas non reconnus avec un taux de 10%.

• **Les résultats après la réduction**

	Normale	Pathologique
Décision		
Normale	90%	10%
Pathologique	10%	90%

Tableau 3.16: résultat obtenu avec réduction « LDA » pour la base HUPA
« 13Coefficients MFCC »

➤ **Pour 13 coefficients :**

- Les cas reconnus normales avec un taux de 90% et les cas non reconnus avec un taux de 10%
- Les cas reconnus pathologique avec un taux de 90% et les cas non reconnus avec un taux de 10%

	Normale	Pathologique
Décision		
Normale	92%	8%
Pathologique	10%	90%

Tableau 3.17: résultat obtenu avec réduction « LDA » pour la base HUPA
« 26Coefficients MFCC »

➤ **Pour 26 coefficients :**

- Les cas reconnus normales avec un taux de 92% et les cas non reconnus avec un taux de 8%

- Les cas reconnus pathologique avec un taux de 90% et les cas non reconnus avec un taux de 10%

Décision \	Normale	Pathologique
Normale	98%	2%
Pathologique	4%	96%

Tableau 3.18: résultat obtenu avec réduction « LDA » pour la base HUPA
« 39Coefficients MFCC »

➤ **Pour 39 coefficients :**

- Les cas reconnus normales avec un taux de 98% et les cas non reconnus avec un taux de 2 %
- Les cas reconnus pathologique avec un taux de 96% et les cas non reconnus avec un taux de 4%.

Décision \	Normale	Pathologique
Normale	88%	12%
Pathologique	10%	90%

Tableau 3.19: résultat obtenu avec réduction « PCA » pour la base HUPA
« 13Coefficients MFCC »

➤ **Pour 13 coefficients :**

- Les cas reconnus normales avec un taux de 88% et les cas non reconnus avec un taux de 12%

- Les cas reconnus pathologique avec un taux de 90% et les cas non reconnus avec un taux de 10%

Décision \	Normale	Pathologique
Normale	88%	12%
Pathologique	8%	92%

Tableau 3.20: résultat obtenu avec réduction « PCA » pour la base HUPA
« 26Coefficients MFCC »

➤ **Pour 26 coefficients :**

- Les cas reconnus normales avec un taux de 88% et les cas non reconnus avec un taux de 12%
- Les cas reconnus pathologique avec un taux de 92% et les cas non reconnus avec un taux de 8%

Décision \	Normale	Pathologique
Normale	94%	6%
Pathologique	8%	92%

Tableau 3.21: résultat obtenu avec réduction « PCA » pour la base HUPA
« 39Coefficients MFCC »

➤ **Pour 39 coefficients :**

- Les cas reconnus normales avec un taux de 94% et les cas non reconnus avec un taux de 6%

- Les cas reconnus pathologique avec un taux de 92% et les cas non reconnus avec un taux de 8%.

➤ **Decision Finale**

- Les résultats obtenus en utilisant la base de données HUPA sont mieux que les résultats obtenus avec la base de données SVD
- La méthode de réduction LDA est mieux que la méthode de PCA.
- Les résultats des 39 coefficients sont mieux que les coefficients de 26 et 13

3.7 Conclusion

Dans ce chapitre nous avons proposé une méthode de discrimination entre la voix pathologique et la voix normale à base des machines à vecteurs de supports et en utilisant deux base de données SVD et HUPA.

L'objectif principal est de trouver les meilleurs systèmes donnant de bons résultats et pour aider à les avoir nous avons exploité deux méthodes de réduction de dimension qui sont la LDA et PCA. Le model LDA a présenté le meilleur taux de reconnaissance pour les deux bases de données.

Les résultats que nous avons obtenus sont Les résultats obtenus sont très acceptables et comparables à ceux reportés dans la littérature.

3.8 Conclusion Générale

L'objectif de notre travail était la détection et la classification des pathologies de la voix par la méthode de classification SVM et l'évaluation de l'impact des paramètres sélectionnés du signal vocal sur cette identification. Plusieurs variantes du classificateur SVM et les deux techniques de réduction de dimension PCA et LDA ont été utilisés pour comparer leurs performances.

Le système proposé peut être utilisé comme un outil précieux par les chercheurs et les orthophonistes pour détecter si la voix est normale ou pathologique et également pour détecter un type spécifique de pathologie. Le système proposé et étudié emploie des mesures non invasives, peu coûteuses et entièrement automatisées des caractéristiques du signal vocal. À partir des résultats obtenus, nous avons remarqué que les paramètres MFCC sont plus pertinents pour discriminer les pathologies des cordes vocales, lorsqu'ils sont évalués individuellement et surtout lorsqu'ils sont combinés. Nous avons montré que le classificateur SVM et les méthodes de réduction PCA et LDA peuvent détecter efficacement les voix pathologiques et normales.

Les résultats de la validation simple confirment que la méthode de réduction LDA assure les meilleures performances. Il est conclu à partir des résultats expérimentaux que la LDA offre de bonnes performances pour ces données vocales en fonction des fonctionnalités utilisées nous avons prouvé que les performances de la réduction de dimension LDA sont très bonnes.

Dans les travaux futurs, pour améliorer le taux de classification obtenu, nous proposons d'améliorer la phase de classification par l'utilisation d'un système hybride en combinant plusieurs techniques d'apprentissage automatique, d'augmenter le nombre de paramètres utilisés dans la phase d'extraction des paramètres, d'étudier l'influence de la sélection de certains paramètres sur les performances du système, d'optimiser certaines valeurs et coefficients du système, d'étendre cette étude sur d'autres bases de données et de développer un système de diagnostic en ligne.

Bibliographie

- [1]<http://www.claudegabriel.be/Cine%20acoustique%209.pdf>
« Appareil Phonatoire » **Page 4**
- [2] D. Docteur massot, "Voix normale et pathologique," 2014. [Online]. Available: <http://www.hubert-massot.com/#!/-propos1/cd3a>. **Page 12**
- [3] Lion's voice clinic, "Types of voice disorders," [Online]. Available: <http://www.lionsvoiceclinic.umn.edu/page3b.html>. **Page 12**
- [4] Clinique d'orthophonie de Montréal, "Traitement des troubles de la voix," 2016. Available: <http://www.arsenault-orthophonie.com/fr/lesproblematiques/troubles-de-la-voix>. **Page 13**
- [5] Institut Numérique., "L'évaluation subjective et objective de la voix," 2013 Available: <http://www.institut-numerique.org/chapitre-3-levaluationsubjective-et-objective-de-la-voix-5194afdf80cc9>. **Page 14**
- [6] J. I. Godino-Llorente, P. Gómez-Vilda, N. Sáenz-Lechón, M. Blanco-Velasco, F. CruzRoldán, Ferrer-Ballester and M. Angel, "Support vector machines applied to the detection of voice disorders," in Nonlinear Analyses and Algorithms for Speech Processing, 2005. **Page 24**
- [7] G. De Krom, "A cepstrum-based technique for determining a harmonics-to-noise ratio in speech signals," Journal of Speech and Hearing Research, vol. 36, no. 02, pp. 254-266, App 1993. **Page 21**
- [8] H. Kasuya, S. Ogawa, K. Mashima and S. Ebihara, "Normalized noise energy as an acoustic measure to evaluate pathologic voice," Journal of the Acoustical Society of America, vol. 80, no. 5, pp. 1329-1334, Nov 1986. **Page 22**
- [9] D. Michaelis, T. Gramss and H. Strube, "Glottal-to-Noise Excitation ratio- a new measure for describing pathological voices," Acustica/Actaacustica, vol. 83, pp. 700- 706, 1997. **Page 23**

- [10] M. Hariharan, K. Polat, R. Sindhu and S. Yaacob, "A hybrid expert system approach for telemonitoring of vocal fold pathology," Applied Soft Computing, vol. 13, p. 4148–4161, 2013.**Page 32**
- [11] M. K. Arjmandi and M. Pooyan, "An optimum algorithm in pathological voice quality assessment using wavelet-packet-based features, linear discriminant analysis and support vector machine," Biomedical Signal Processing and Control, vol. 7, pp. 3-19, 2012.**Page 26**
- [12] Massachusetts. Eye and Ear Infirmary, "Voice disorders database," Kay Elemetrics Corp, 1994.P. Manfred and J. B. William, "Saarbruecken Voice Database,"[Online]**Page 19**
Available:http://www.stimmdatenbank.coli.uni-saarland.de/help_en.php4.
- [13] J. Godino-Llorente, V. Osma-Ruiz, N. Sáenz-Lechón, I. Cobeta-Marco, R. GonzálezHerranz and C. Ramírez-Calvo, "On combining information from modulation spectra and mel-frequency cepstral coefficients for automatic detection of pathological voices," Logopedics Phoniatrics Vocology, vol. 36, pp. 1-10, 2011. **Page 21**
- [14] P.-P. Béatrice and P. Jean-Christophe, Le traitement de signal et ses applications, Techniques de l'Ingénieur, 2001.**Page 20**
- [15] P.FLANDRIN, Temps-fréquence, Paris:Hermès, 1998.**Page 16**
- [16] "Continuous Wavelet Transform and Scale-Based Analysis," The MathWorks, 1994.**Page 36**
- [17] G. Mak, "The implementation of support vector machines using the sequential minimal optimization algorithm," Montreal, Canada, 2000.**Page 24**
- [18] N. Marref, "Apprentissage incrémental et machines à vecteurs supports," Batna, Algérie, 2013.**Page 19**
- [19] E. Nemer, R. Goubran and S. Mahmoud, "Robust voice activity detection using higherorder statistics in the LPC residual domain," IEEE Trans. Speech Audio Process, vol. 3, no. 9, pp. 217-231, 2001.**Page 21**

- [20] "MathWorks," [Online]. **Page 36**
Available: <https://www.mathworks.com/matlabcentral/linkexchange/links/797-voicebox-speechprocessing-toolbox-for-matlab>.
- [21] C. Chih-Chung and L. Chih-Jen, "LIBSVM A library for support vector machines," 14 December 2015. [Online]. Available: <https://www.csie.ntu.edu.tw/~cjlin/libsvm/> **Page 36**
- [22] M. Vahid and K. Igor, "The SVM-based feature reduction in vocal fold pathology diagnosis," International Journal of Future Generation Communication and Networking, vol. 06, no. 01, February, 2013. **Page 25**
- [23] Mounir Boudjerda, "Analyse du signal de parole pour l'évaluation automatique des voix pathologiques," Thèse de Doctorat, Université Mohammed Seddik BENYAHIA-Jijel, Algérie, 2018. **Page 16**
- [24] J. Paulo Teixeiraa and A. Gonçalves, "Accuracy of jitter and shimmer measurements," in Procedia Technology, International Conférence on Project Management / HCIST 2014 - International Conférence on Health and Social Care Information Systems and Technologies, Portugal, vol.16, pp. 1190-1199, 2014. **Page 18**
- [25] M. Farrús, J. Hernando and P. Ejarque, "Jitter and Shimmer measurements for speaker recognition," in INTERSPEECH, Antwerp, Belgium, August 27-31, 2007, pp. 778-781. **Page 18**
- [26] <https://www.institut-numerique.org/chapitre-3-levaluation-subjective-et-objective-de-la-voix-5194afdf80cc9> **Page 23**
- [27] www.boowiki.info/art/analytique-2/analyse-des-composantes-principales.html
- [28] [Schafer. 2009] R. W. Schafer. Homomorphic Systems and Cepstrum Analysis of Speech. Springer handbook of speech processing, pages 161–180, 2009. **Page 22**
- [29] D.G. Childers 1992 K. Sung-Bae D.G. Childers. Detection of laryngeal function using speech and electroglottographic data. IEEE Trans. Biomed. Eng., vol. 39, no. 1, pages 19–25, 1992. **Page 21**

- [30] **K. Tokuda 2002** H. Zen K. Tokuda and A. W. Black. An HMM-Based Speech Synthesis System Applied to English. IEEE Workshop on Speech Synthesis, pages 227–230, 2002. **Page 22**
- [31] Développement d'un système de détections de la voix pathologique (D. Fethi Amara, 2014)**Page 2.**
- [32] Deep Learning basé sur les méthodes de réduction pour la reconnaissance de visage