

**République Algérienne Démocratique et Populaire**  
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique  
Université Djilali Bounaama Khemis Miliana



**Faculté des Sciences et de la Technologie**  
**Département de Technologie**

Mémoire du Projet de Fin d'Etudes  
Pour l'obtention du diplôme de

**Master**

En

« Télécommunications »

Option :

« Systèmes de Télécommunications »

**Titre :**

**Débruitage du signal de la parole par la  
transformée en ondelette**

**Réalisé par :**

KASSAA Ahmed

DJERMOUNI Abdelhamid

**Encadré par :**

Mme BOUNIF Aouda

Année Universitaire: 2016/2017

# *Dédicace*

## **Je dédie ce mémoire à :**

*Mon père et Ma mère, qui a œuvré pour ma réussite, de par son amour, son soutien, tous les sacrifices consentis et ses précieux conseils, pour toute son assistance et sa présence dans ma vie, reçois à travers ce travail aussi modeste soit-il, l'expression de mes sentiments et de mon éternelle gratitude.*

*Mes frères et sœurs qui n'ont cessé d'être pour moi des exemples de persévérance, de courage et de générosité.*

*Mes professeurs de l'UKM qui doivent voir dans ce travail la fierté d'un savoir bien acquis. Toutes mes amies et mes collègues*

**Ahmed**

# *Dédicace*

## **Je dédie ce mémoire à :**

*A la mémoire de mon père qui a souhaité vivre pour longtemps juste pour nous voir Qu'est-ce que nous allons devenir.*

*A ma mère et mes frères et sœurs*

*A tous mes amis dans classe de la MASTER 2 OPTIONS SYSTEMES DE TELECOMUNICATIONS DE L'UNIVERSITE DJILALI BONAAMA.*

*A tous mes amis avec lesquels j'ai partagé mes moments de joie et de bonheur*

*Que toute personne m'ayant aidé de près ou de loin, trouve ici l'expression de ma reconnaissance.*

**Abdelhamid**

# Remerciements

Nous remercions tout d'abord le grand Dieu pour l'achèvement de ce mémoire.

Nous exprimons nos gratitude à Monsieur le président de jury d'avoir accepté examiné ce mémoire.

Nous remercions Messieurs les membres de jury, d'avoir accepté de prendre part à ce jury ainsi que pour l'intérêt qu'ils l'ont portés à ce travail.

Nous remercions Mme BOUNIF AOUDA, notre promotrice, pour ses conseils et suggestions avisés qui nous ont aidés à mener à bien ce travail, et d'avoir rapporté à ce mémoire ces remarques et conseils.

A tous mes professeurs à l'université de DJILALI BOUNAAMA khemis miliana.

## **Résumé**

Ce mémoire de fin d'études présente la débruitage de la parole avec l'ondelette de Haar. Le signal de parole (communication parlée) n'est pas un signal ordinaire, il est le vecteur d'un phénomène extrêmement complexe. Il est difficile de modéliser ce signal car ses propriétés statistiques varient au cours du temps. Sa redondance lui confère une robustesse à certains types de bruit. Il possède aussi une très grande variabilité. Le traitement de la parole est aujourd'hui une composante fondamentale des sciences de l'ingénieur, située au croisement du traitement du signal numérique et du traitement du langage. Nous nous sommes intéressés à la débruitage d'un signal de parole par l'ondelette de Haar car cette méthode offre plus de flexibilité en ce qui concerne le choix d'une base ortho-normale pour le représenter.

L'évaluation de la qualité de compression a été faite avec les critères subjectifs et objectifs.

## **Mots clés :**

Parole, débruitage, ondelette, ondelette de Haar, signal, reconstruction.

---

## **Abstract**

This dissertation presents the de-noising of speech with Haar wavelet. The speech signal (spoken communication) is not an ordinary signal, it is the vector of an extremely complex phenomenon. It is difficult to model this signal because its statistical properties vary over time. Its redundancy gives it a robustness to certain types of noise. It also has a great variability. Speech processing is today a fundamental component of the engineering sciences, located at the intersection of digital signal processing and language processing. We are interested in the denoising of a speech signal by the Haar wavelet since this method offers more flexibility in the choice of an ortho-normal basis to represent it.

The evaluation of the quality of compression was made with the subjective and objective criteria

## **Keyword**

Speech, noising, wavelet, Haar wavelet, PSNR, MSE reconstruction signal.

## Liste des abréviations

API : Alphabet Phonétique International

SNR: Signal to Noise Ratio

TFD : transformée de Fourier discrète

ANC : Adaptative Noise Cancellation

AVC : Addaptatife voice codor

LMS : Least Mean Square

RIF : Réponse Impulsionel Finie

MSE : Mean Square Error

LPC : linear predictive coding

MFCC : Mel Frequency Cepstral Coefficients

MOS: Mean Opinion Score

PESQ : Perceptual Evaluation of Speech Quality

TCD : transformée en cosinus discret

TF : Transformée de Fourier

TO : Transformée en Ondelettes

TOC : Transformées en Ondelettes Continue

TOD : Transformées en Ondelettes discrète

STFT : Short Time Fourier Transforme

MRA : Multi Résolution Analyse

EQM: Erreur quadratique moyenne

## Liste des figures

figure1.1 : (a) : L'appareil phonatoire, (b) : glotte.....	04
Figure.1.2 : Diagramme synoptique de la production de la parole humaine (modèle de canalisations) .....	06
Figure.1.3 : (a) : les symboles de l'alphabet phonétique international (IPA).  (b) : Modèle de la production de la parole.....	09
Figure.1.4 : Représentation temporelle d'un son : (a) voisé, (b) non voisé.....	11
Figure. 1.5 : Forme d'onde d'une trame de signal vocal voisé et son spectre d'amplitude.....	12
Figure. 1.6 : Forme d'onde d'une trame de signal vocal non voisé et son spectre d'amplitude.....	13
Figure. 1.7 : Fonction d'auto corrélation pour un son voisé (a) et (b) ; un son non voisé (c) avec une fenêtre rectangulaire de N=401.....	14
Figure1.8 : Harmoniques de la source glottale, formant du conduit Vocal et enveloppe spectrale.....	15
Figure1.9 : Exemple de forme d'onde de parole (mâle) du mot " problèmes ".....	15
Figure 1.10 : Spectrogramme d'un signal de parole.....	16
Figure .1.11 : Interprétation d'un Spectrogramme.....	17
Figure1.12 : Application du traitement de la parole.....	17
Figure1.13 : Système auditif humain.....	20
Figure1.14 : Les oreilles externe, moyenne et interne.....	21
Figure1.15.Sensibilité fréquentielle le long de la membrane basilaire.....	21
Figure 2.1: illustre les différentes approches de filtrage appliquées dans la réduction du bruit.....	27
Figure 2.2 La transformée de Fourier discrète.....	28
Figure 2.3 diagramme de la méthode d'élimination adaptative du bruit(AVC).....	30
Figure 2.4 : Le filtre adaptif LMS.....	30
Figure 2.5 : Principe du filtrage perceptuel .....	32

Figure 2.6 : Seuil absolu d'audition et masquage .....	33
Figure 2.7: Le masquage fréquentiel (a) à 1kHz (b) à plusieurs fréquences .....	34
Figure 3.1.Représentation temps-échelle de la STFT et de la transformée en ondelette .....	39
Figure 3.2. Translation et dilatation des ondelettes .....	42
Figure 3.3 : Exemples d'ondelettes continues.....	44
Figure 3.4 : Familles des ondelettes discrètes (Daubechies ,Coiflets et symlet) .....	45
Figure 3.5 : La fonction d'échelle (a) et Ondelette (b) de Haar .....	46
Figure 3.6 : Un signal (traits gras) et son approximation par la fonction d'échelle de Haar ...	46
Figure 3.7 : Coefficients d'échelles et d'ondelettes d'une analyse multi-résolution.....	47
Figure 3.8 : famille d'ondelettes de Haar.....	48
Figure 3.9 : Analyse multi-résolution de Haar d'un signal .....	49
Figure .4.1:Analyse d'un signal de parole avec l'ondelette Haar de profondeur .....	54
Figure.4.2: Débruitage d'un signal de parole par Ondelette de Haar de profondeur .....	55
Figure.4.3 : Représentation d'un signal débruité .....	55
Figure.4.4 : Audiogramme du signal original avec bruit .....	56
Figure.4.5: signal de parole bruité avec différents SNR .....	56
Figure.4.6 : Signal de parole bruité avec SNR =40dB.....	57
Figure 4.7: Les coefficients de détail et d'approximation de la décomposition par ondelette de Haar niveau 3 du signal .....	57
Figure.4.8: Signal de parole bruité reconstruit.....	58
Figure.4.9:Signal sinusoïdal original utilisé.....	59
Figure.4.10: Représentation de tous les détails du signal analysé .....	59
Figure.4.11: Reconstruction du signal .....	60
Figure.4.12 : coefficients de détails du signal bruité avec 40 dB .....	61
Figure.4.13 : coefficients d'approximation du signal bruité avec 40 dB .....	61
Figure.4.14 : Signal original, sa version bruitée avec 40 dB et le signal débruité .....	62
Figure .4.15 : coefficients de détails du signal bruité avec 20 dB .....	63

Figure.4.16 : coefficients d'approximation du signal bruité avec 20 dB .....	63
Figure.4.17 : Signal original, sa version bruitée avec 20 dB et le signal débruité .....	64
Figure.4.18 : coefficients de détails du signal bruité avec 10 dB. ....	65
Figure.4.19 : coefficients d'approximation du signal bruité avec 10 dB .....	65
Figure.4.20 : Signal original, sa version bruitée avec 10 dB et le signal débruité .....	66
Figure.4.21 : débruitage d'un signal de parole quantifié à 8 bits .....	68
Figure.4.22 : débruitage d'un signal de parole quantifié à 6 bits .....	68



## Liste des tableaux

Tableau 1.1 : Fréquences moyennes des formants pour les voyelles.....	9
Tableau 2.1 : Différentes classes du bruit.....	26
Tableau 3.1 : L'échelle MOS.....	51
Tableau 4.1 : Comparaison des résultats obtenus par deux ondelettes (Haar et coiffman).....	67
Tableau 4.2 : débruitage d'un signal quantification à 8 bits et 6 bits .....	69

# Tableau de Matière

**Remerciement**

**Résumé**

**Introduction générale**..... 01

## **Chapitre1 : Généralité sur le signal de la parole**

1.1. Introduction .....	03
1.2. Catégorisation des sons vocaux .....	03
1.3. Le mécanisme de la production de la parole .....	04
1.3.1. Processus de la production vocale .....	04
1.3.2. Le phénomène de la coarticulation .....	06
1.3.3. Intérêt de la coarticulation .....	07
1.3.4. L'assimilation .....	08
1.3.5. Exemples d'assimilation .....	08
1.3.6. Production de voyelles .....	10
1.4. Caractéristiques acoustiques du signal de la parole .....	10
1.4.1. Onde acoustique .....	10
1.4.2. Analyse spectrale de la parole .....	11
1.4.2.1. Les sons voisés .....	11
1.4.2.2. Les sons non voisés .....	12
1.4.3. Harmoniques .....	13
1.4.4. Conduit vocal et formants .....	13
1.4.5. Spectrogrammes .....	15
1.5. Application du traitement de la parole .....	17
1.5.1. Modifications .....	17
1.5.2. Amélioration du discours .....	18
1.6. Variabilité et non linéarité du signal vocal .....	18
1.7. Observation et modélisation du conduit vocal .....	19
1.8. Système auditif humain .....	19

1.8.1. Fonctionnement du système auditif humain .....	19
1.8.2. A l'intérieur de la cochlée .....	21
1.9. Conclusion .....	22

## **Chapitre 2 : les techniques de débruitage d'un signal**

2.1. Introduction .....	23
2.2. L'inconvénient du bruit .....	23
2.3. Types et sources de bruits.....	24
2.3.1. Types de bruits.....	24
2.3.2. Sources de bruits.....	25
2.3.3 .Nature et caractéristiques du bruit.....	25
2.4. Nécessité de l'opération de débruitage.....	26
2.5. Familles des techniques de débruitage.....	26
2.5.1 Méthode de soustraction spectrale.....	27
2.5.2. Filtrage de Wiener.....	29
2.5.3. Elimination adaptive de bruit .....	30
2.5.4. Techniques basé sur le modèle de la parole.....	31
2.5.5. Méthode de masquage .....	32
2.6. Modèles de bruit et évaluation de la qualité .....	34
2.6.1. Modèles de bruit.....	34
2.6.2. Evaluation de la qualité du débruitage.....	34
2.6.3. Valeurs typiques de la ligne téléphonique.....	35
2.6. Conclusion.....	35

## **Chapitre 3 : La transformée en ondelette et son application pour debruiter le signal de son**

3.1. Introduction .....	37
3.2. Transformée de Fourier .....	37
3.3. Transformée de Fourier à court terme .....	38
3.4 La multi résolution .....	39
3.5. Transformée en ondelettes continue .....	40
3.6. Transformée en ondelettes discrète (TOD).....	42
3.7. Propriétés des ondelettes .....	43
3.8. Familles d'ondelettes .....	44

3.8.1 Exemples d'ondelettes continues.....	44
3.8.2 Exemples d'ondelettes discrètes.....	45
3.9. Analyse d'un signal avec l'Ondelette de Haar .....	46
3.10. Quelques applications d'ondelettes.....	50
3.11. Critères d'évaluation de qualité de la compression.....	50
3.11.1. L'erreur quadratique moyenne .....	51
3.11.2. le rapport signal sur bruit.....	51
3.11.3. Le MOS (Mean Objective Score ).....	51
3.12. Conclusion.....	52

#### **CHAPITRE 4 : Simulation et résultats expérimentaux (ondelette de haar)**

4.1. Introduction .....	53
4.2. Environnement matériel et logiciel .....	53
4.3. Prise en main de la boîte à outils .....	53
4.3.1. La commande « wavemenu ».....	54
4.3.2. Analyse d'un signal 1D .....	54
4.3.3. visualisation des signaux .....	55
4.4. Génération de la parole bruitée avec divers niveaux de RSB.....	56
4.5. Décomposition d'un signal de parole par ondelette de Haar.....	57
4.5.1. Analyse d'un signal ayant un SNR=40dB.....	57
4.5. 2. Reconstruction du signal traité.....	58
4.6. Le débruitage d'un signal par les ondelettes.....	58
4.6.1. Débruitage d'un signal sinusoïdal par ondelette de Haar.....	58
4.6.2. Débruitage d'un signal de parole par ondelette de Haa.....	60
4.6.2.1. signal de parole SNR = 20 db .....	60
4.6.2.2. Signal de parole avec SNR = 10 dB .....	62
4.6.2.3. Signal de parole SNR = 0 db .....	64
4.7. Résultats récapitulatifs et évaluation de la qualité du débruitage.....	67
4.8. Débruitage du bruit quantification .....	68
4.9. Conclusion .....	69

### Introduction générale

Le traitement du signal est une discipline qui est méconnue car elle se fonde à l'intersection de plusieurs domaines scientifiques. Les premiers développements trouvent leurs origines dans l'électronique et l'automatique. Les signaux de télécommunications transportent l'information par modulation avec une porteuse, le plus souvent sous forme numérique. Ces signaux sont simples à étudier car ils présentent des propriétés permettant de faciliter les opérations de modulation, de démodulation. Le but du traitement du signal est, en effet d'extraire le maximum d'information utile sur un signal perturbé par le bruit.

Les signaux de communication, tels la parole ou l'écriture, à l'extrême opposé des signaux de télécommunications, ces signaux réalisent une modulation très complexe de l'information qu'ils portent. Le modulateur et le démodulateur sont en effet un cerveau humain. On peut également classer les signaux en deux catégories signaux stationnaires et non stationnaires dont la parole fait partie.

Dans de telles situations, les méthodes classiques comme la transformée de Fourier n'est pas l'outil approprié pour mener cette analyse puisqu'elle masque l'évolution temporelle du signal. Elle ne nous permet pas d'obtenir de bons résultats car la représentation fréquentielle (TF) ne localise pas les lieux de changement de fréquence.

De là, il est apparu la transformée de Fourier à fenêtre glissante et la transformée de Gabor comme solution à la limitation de Fourier. Mais ces deux méthodes donnent une même résolution temporelle pour les hautes et les basses fréquences ce qui rend cette analyse non idéale.

Pour remédier à ce problème, la transformée en Ondelettes a été créée. Elle permet d'analyser les différentes propriétés des signaux. Il s'agit d'une fonction qui oscille comme une onde mais qui est rapidement atténuée. Elle permet à la fois une localisation dans l'espace-temps (résolution temporelle) et dans l'espace fréquentiel (la résolution fréquentielle). Par rapport aux techniques de Fourier, les ondelettes donnent plus de flexibilité dans le choix des représentations appropriées pour des applications particulières.

Notre mémoire est composé de quatre chapitres dont la première partie est consacrée à l'analyse du signal de parole. Ainsi, les mécanismes de production et caractéristiques acoustiques du signal vocal.

Dans le deuxième chapitre nous présenterons les méthodes de débruitage d'un signal de parole. Nous essayons de décrire le signal de bruit et les traitements qui permettent sa suppression / réduction.

Le troisième chapitre décrit de façon détaillée la transformée en ondelette et son application pour débruiter le signal de son. La première section concerne les transformées fondamentales en traitement du signal. La section suivante décrit la transformée d'ondelette et quelques types de cette dernière. Dans la troisième section, nous présenterons l'analyse d'un signal avec l'Ondelette de Haar .

Le dernier chapitre sera consacré à la simulation de l'algorithme de débruitage d'un signal de parole, par ondelettes de Haar , avec le logiciel MATLAB.

En fin, nous terminerons notre travail par une conclusion générale.

### 1.1. Introduction

Le traitement de la parole est une science située au croisement du traitement du signal numérique et du traitement du langage. La parole a la particularité, par rapport aux autres signaux du traitement de l'information, à être produite et perçue instantanément par le cerveau, et pour cela le traitement de la parole tend à remplacer ces fonctions par des systèmes automatiques [1].

Dans ce chapitre, nous présenterons les différents des sons vocaux et le mécanisme de la production. Ensuite, nous donnerons quelques caractéristiques du signal de la parole.

Dans la troisième section, nous aborderons les mécanismes de l'audition et les propriétés perceptuelles de l'appareil auditif humain.

### 1.2. Catégorisation des sons vocaux

Les sons de la parole sont étudiés et classés selon les perspectives suivantes:

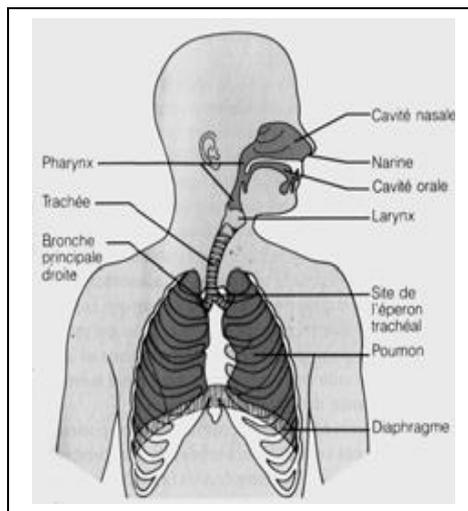
- La nature de la source: périodique, bruyante, ou impulsif, et des combinaisons des trois.  
Autres classes facultatives:
- La forme de l'appareil vocal.
- La forme d'onde du domaine temporel, qui donne la variation de pression avec le temps à la sortie des lèvres.
- Les caractéristiques spectrales variant dans le temps révélées par le spectrogramme.

Ainsi, on distingue sept niveaux de descriptions possibles pour caractériser le signal de la parole : acoustique, phonétique, phonologique, morphologique, syntaxique, sémantique, pragmatique.

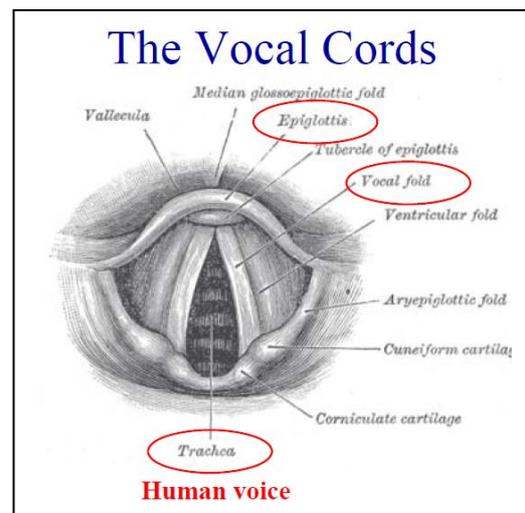
Exemple: Dans le mot «shop», les «sh», «o» et «p» sont générés à partir d'une source de bruit, périodique et impulsive, respectivement.

### 1.3. Le mécanisme de la production de la parole

C'est la représentation phonétique de la parole. La production du son est accomplie par l'appareil phonatoire humain (figure.1.1). Les sons vocaux sont produits soit par la vibration quasi-périodique de cordes vocales pour les sons sonores, ou par turbulence à un certain point de constriction sur le conduit vocal : le larynx, le pharynx et les cavités buccales et nasales [1].



(a)



(b)

figure1.1 : (a) : L'appareil phonatoire, (b) : glotte

#### 1.3.1. Processus de la production vocale

La production vocale est réalisée grâce à quatre processus :

- **La respiration** : C'est le mécanisme de puissance qui fournit l'énergie pour le son. C'est un processus aérodynamique, une différence de pression d'air est créée entre la cavité thoracique et la pression atmosphérique à l'extérieur du corps. Le processus de respiration se fait en deux étapes :
  - Inspiration : les muscles de la cavité thoracique élargissent la poitrine; la pression atmosphérique diminue et l'air entre. La pression atmosphérique est moindre dans le thorax que dans l'atmosphère.

- Expiration : lorsque la pression est plus grande dans le thorax, l'air s'échappe. Il y a une diminution de la cavité thoracique et une augmentation de la pression atmosphérique. La respiration pour la parole utilise les mêmes muscles que la respiration pour la vie, mais elle est plus contrôlée. La phase expiratoire est plus longue que la phase inspiratoire pour la parole. Les phases d'inspiration et d'expiration de respiration (pour la vie) sont identiques.
- **La phonation** : C'est le mécanisme vibratoire. La phonation est l'ouverture et la fermeture rapide des plis vocaux. Les plis vocaux se trouvent horizontalement dans le larynx. Ils se fixent antérieurement au cartilage thyroïdien et postérieurement aux cartilages arythnoïdes. Ils sont libres au milieu. l'ouverture au milieu des plis vocaux est la **glotte**. L'adduction des arythnoïdes ferme les plis vocaux. Le milieu des plis vocaux vibre vers et depuis la ligne médiane. La vibration des cordes vocales produit des sons. La position des plis vocaux : ouvert (enlevé) ou fermé (adducted).

Trois aspects majeurs de la voix sont influencés par le mouvement des cordes vocales :

- Pitch (fréquence des vibrations) : le taux de vibration des plis vocaux est appelé **fréquence fondamentale**. Elle est mesurée en Hz ou cps. La fréquence de vibration des cordes vocales dépend de la masse, de la longueur et de la tension des plis vocaux.
- Loudness (intensité du son) : représente la quantité d'énergie générée par les mouvements du pli vocal. Elle augmente avec la pression atmosphérique des poumons et l'amplitude de la vibration du pli vocal
- Qualité (la qualité sonore de la voix) : elle est affectée par le modèle du mouvement des plis vocaux.
- **La résonance (la qualité caractéristique de la voix)** : Le mécanisme de la parole est un résonateur. Il ressemble à un tube rempli d'air (fermé à une extrémité et ouvert à l'autre extrémité). La fréquence fondamentale générée au niveau des plis vocaux est résonnée dans le conduit vocal qui vibre également. Cette vibration donne la fréquence harmonique. Chaque partie a sa propre fréquence naturelle. Les fréquences qui correspondent à la fréquence naturelle de la voix seront amplifiées : c'est la résonance. Le Pharynx est le premier résonateur dont les trois parties sont : nasopharynx, oropharynx, laryngopharynx. Le nasopharynx et l'oropharynx sont les deux systèmes qui fournissent la résonance aux sons qui passent par les cavités orales et nasales.

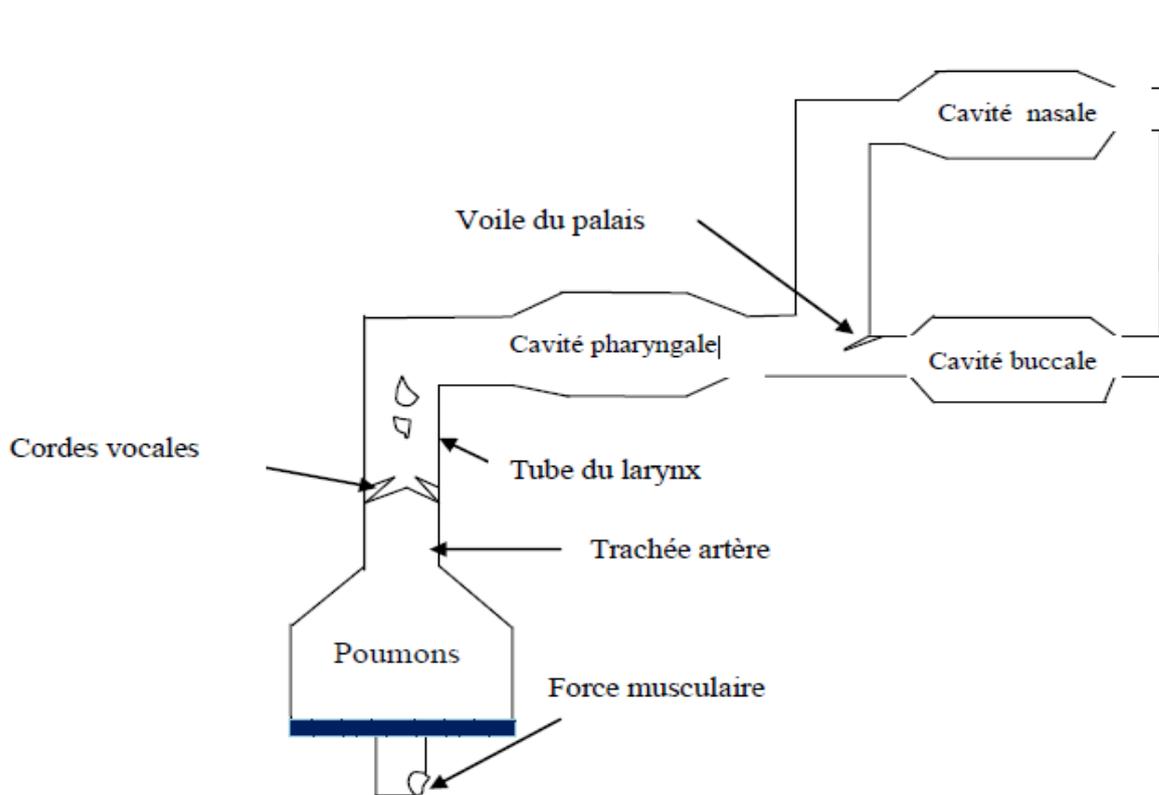


Figure.1.2 : Diagramme synoptique de la production de la parole humaine (modèle de canalisations)

- **L'articulation** : Sert à produire les différentes configurations qui composent les différents sons vocaux. La parole est un phénomène dynamique plutôt que statique. Les articulateurs sont dans un état de mouvement constant pendant la parole. Les sons vocaux individuels sont décrits en termes de positions cibles :
  - Les sons d'un mot ne se suivent pas comme des lettres imprimées
  - les sons parlés se chevauchent : c'est ce qu'on appelle **coarticulation**.

### 1.3.2. Le phénomène de la coarticulation

Lorsque les articulateurs sont en train de former un son, leur positionnement les prépare pour le son qui suit :

- les sons de la parole influencent et sont influencés par d'autres sons dans un environnement phonétique

- la nature dynamique de la parole signifie que nous planifions des événements avant qu'ils ne se produisent
- on peut articuler un segment avant de compléter un autre.
- Exemples de coarticulation :

Un exemple concernant les voyelles. Le /u/ dans "doute" a une réalisation phonétique qui se rapproche du son [y], car son articulation est plus avancée que lorsque la voyelle /u/ est prononcée de façon hyper-articulée ou isolément, alors que la réalisation de /u/ dans le mot "Ruhr" se rapproche de celle du son [o], car son articulation est plus ouverte du fait de la présence des consonnes environnantes. Un exemple concernant les consonnes. La réalisation des consonnes /s/ and /ch/ sont maximale dans les mots "si" [si] et "chou" [Su], mais elles tendent à se rapprocher dans la prononciation des séquences "chichi" ou "sousou": en particulier, dans le cas de "sousou", la consonne /s/ tend à être réalisée avec des lèvres arrondies, ce qui abaisse les résonances dues à la cavité antérieure à la constriction, et les rendent plus proches de celle de la consonne S (exemples tirés de J. Vaissière, Que Sais-Je?, La phonétique).

### 1.3.3. Intérêt de la coarticulation

Nous avons la coarticulation, ce n'est pas parce que nous sommes paresseux. Il y a trop de demandes sur le mécanisme de la parole qu'il n'y a de temps pour les remplir.

La conversation normale est produite à un taux compris entre 10-20 segments / seconde, ce qui implique un maximum de 100msec / segment. Cependant, il faut beaucoup plus de temps pour terminer tout geste de la parole et revenir ensuite au point de départ. Donc, quelque chose doit donner - ainsi, un compromis des gestes qui sont faits. Il est vrai que les dialectes et les standards de la parole (formels vs informels) diffèrent dans la quantité de coarticulation qui peut être faite.

Les locuteurs non-natifs veulent apprendre cette accommodation admissible afin de parler plus normal. L'apprenant de la langue seconde a souvent une prononciation hypercorrecte qui semble étrange car elle manque des raccourcis habituels que les locuteurs natifs prennent.

On a besoin de distinguer entre la coarticulation et l'assimilation : la coarticulation signifie que deux sons différents sont articulés simultanément c.à.d. un mouvement de la langue a été fait pour les deux sons.

#### 1.3.4. L'assimilation

L'assimilation concerne les changements qui traversent les frontières phonémiques et aboutissent à un changement phonétique important. Le changement résultant est à un son différent (phonème). La coarticulation entraîne des différences non phonémiques. Dans l'assimilation, il y a un changement important dans le lieu de l'articulation, de la manière d'articuler ou d'exprimer un segment, de sorte qu'il tombe dans une catégorie phonémique différente.

#### 1.3.5. Production de voyelles

Dans la production de voyelles, l'air est expulsé des poumons par la contraction des muscles autour de la cavité pulmonaire. L'air circule à travers les cordes vocales, qui sont deux masses de tissus, provoquant une vibration périodique des cordes dont le rythme donne la fréquence du son. Les bouffées d'air périodiques résultantes agissent comme une entrée d'excitation, ou une source, dans le conduit vocal.

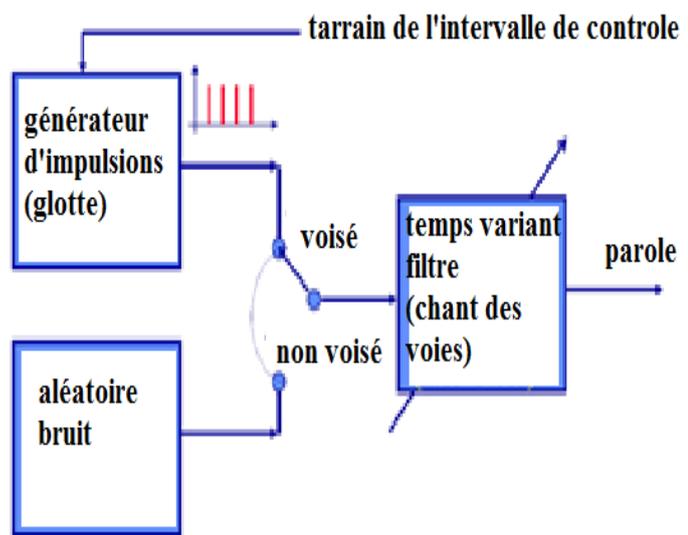
On présente dans la table ci-dessous les valeurs typiques des formants F1 et F2 pour les voyelles du français. Si on reporte ces valeurs dans un graphique représentant F1 et F2 en abscisse et en ordonnée, on obtient le fameux triangle vocalique de la discipline phonétique.

Fréquences des formateurs pour les voyelles		
Voyelle (API)	F1 (HZ)	F2 (HZ)
[u]	320	800
[o]	500	1000
[ɑ]	700	1150
[a]	1000	1400

[ø]	500	1500
[y]	320	1650
[ɛ]	700	1800
[e]	500	2300
[i]	320	3200

Tableau 1.1 : Fréquences moyennes des fermiers pour les voyelles.

API	exemple	
[i]	pire	[piʁ]
[e]	pré	[pʁe]
[ɛ]	père	[pɛʁ]
[a]	mal	[mal]
[y]	Vu	[vy]
[ø]	peu	[pø]
[ə]	Je	[ʒə]
[œ]	peur	[pœʁ]
[u]	mou	[mu]
[o]	zéro	[zɛʁo]
[ɔ]	sort	[sɔʁ]
[ɑ]	pàle	[paʁ]



(a)

(b)

Figure.1.3 : (a) : les symboles de l'alphabet phonétique international (IPA). (b) : Modèle de la production de la parole

### 1.3.6. Les phonèmes

Les sons de base d'une langue (par exemple "p" dans le mot "père") sont appelés phonèmes. Un segment de la parole consiste en une chaîne de voyelles et de consonnes dont les caractéristiques temporelles et spectrales changent avec le temps.

De plus, la source et le système variables dans le temps peuvent également interagir de manière non linéaire d'une manière complexe. Notre modèle simple est correct pour une voyelle régulière. Mais les sons de la parole ne sont pas toujours bien représentés par des systèmes linéaires invariants dans le temps.

## 1.4. Caractéristiques acoustiques du signal de la parole

A ce niveau le signal de la parole est décrit par la forme d'onde dans l'espace temps. La parole est un signal continu, d'énergie finie, non stationnaire. La structure complexe, la variabilité et l'invariabilité dans le temps du signal vocal résulte de :

- La périodicité (plus exactement pseudopériodique) des sons voisés.
- L'aspect aléatoire des sons fricatifs.
- L'aspect impulsionnel des phases explosives pour les sons occlusifs [3].

Le spectrogramme permet caractériser le signal vocal dans le domaine fréquentiel.

### 1.4.1. Onde acoustique

La figure 1.4.(a) ci –dessous montre la variation d'un son sonore (1Sec, 10 000 échantillons, 8kpbs, exprimé ("Ahhhhh")) : un signal quasi-périodique. Un signal non voisé est donné sur la figure 1.4.(b) dont l'allure ressemble à un bruit c.-à-d. manque de périodicité.

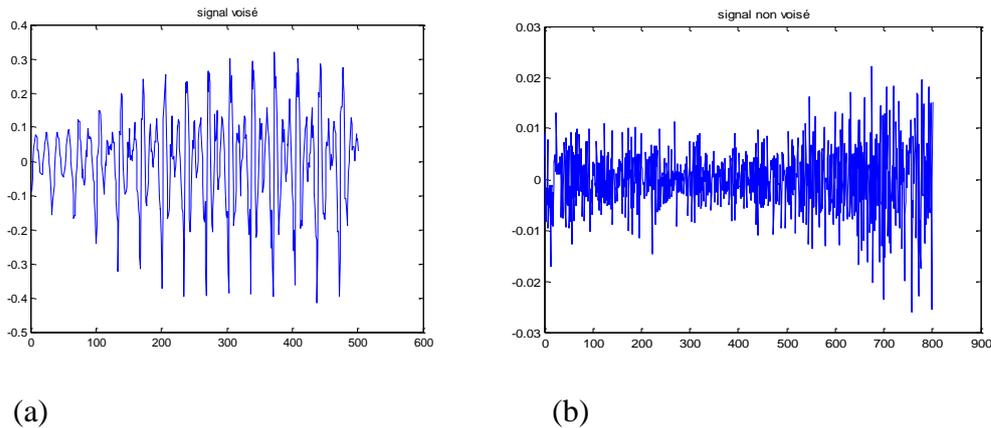


Figure.1.4 : Représentation temporelle d'un son : (a) voisé, (b) non voisé

Il est important de remarquer que le signal vocal n'est pas stationnaire, son évolution suit les déformations du conduit vocal. Cependant, ces déformations sont suffisamment lentes pour que le signal vocal puisse être considéré comme stationnaire sur des périodes allant de 20 à 40ms. On considère donc ce signal comme quasi-stationnaire sur de tels intervalles ce qui motive le choix du traitement par trames (de 20 à 40ms) du signal vocal pour de nombreuses applications, notamment pour la réduction de bruit et le codage de la parole [2].

### 1.4.2 .Analyse spectrale de la parole

L'analyse spectrale des sons voisés et non voisés donne des informations fondamentales dans le cadre du traitement de la parole et en particulier pour la réduction du bruit. Nous allons donc illustrer les principales caractéristiques spectrales des sons voisés et non voisés.

#### 1.4.2.1. Les sons voisés

La hauteur d'un son voisé est liée à la fréquence de vibration des cordes vocales. Cette fréquence est appelée fréquence fondamentale ou pitch. Un son voisé est par définition un signal quasi-périodique qui possède un spectre fréquentiel très caractéristique, comme le montre la figure 1.5. La première raie de ce spectre correspond au fondamental ( $F_0$ ) et les raies suivantes à ses harmoniques (multiples de  $F_0$ ) [2].

L'enveloppe de ces raies possède des maxima locaux appelés formants (localisés en  $F_i$  avec  $i = 1, \dots, 4$ ). Si les trois premiers formants sont indispensables pour assurer l'intelligibilité du signal vocal, les formants d'ordres supérieurs jouent quant à eux un rôle important pour la

qualité du signal vocal. Deux sons de même intensité et de même hauteur se distinguent par le timbre qui est déterminé par les amplitudes relatives des harmoniques du fondamental. Les positions du fondamental ( $F_0$ ) et des formants ( $F_1$ ,  $F_2$ ,  $F_3$  et  $F_4$ ) sont indiquées.

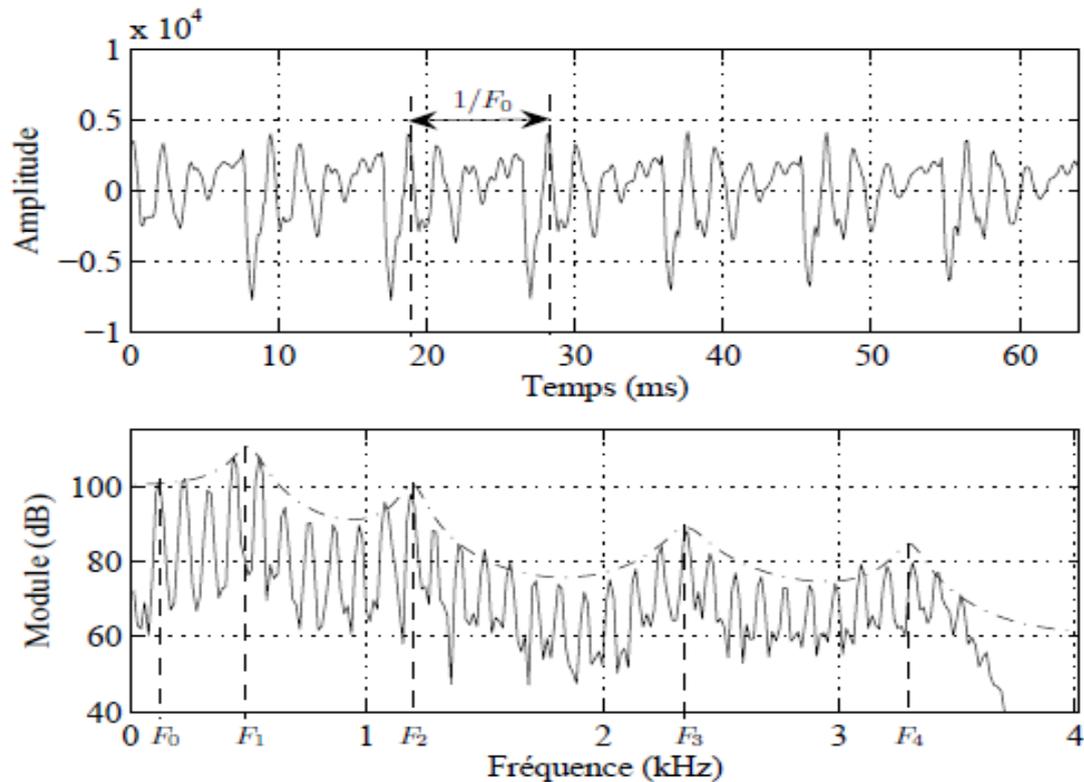


Figure. 1.5 : Forme d'onde d'une trame de signal vocal voisé et son spectre d'amplitude.

La fréquence fondamentale ( $F_0$ ) peut varier [2] :

- de 80 à 200 Hz pour une voix masculine,
- de 150 à 450 Hz pour une voix féminine,
- de 200 à 600 Hz pour une voix d'enfant.

#### 1.4.2.2. Les sons non voisés

Un son non voisé ne présente pas de structure périodique, il peut être considéré comme un bruit blanc filtré par le conduit vocal. Son spectre ne possède donc pas de structure particulière et c'est souvent dans les hautes fréquences que le spectre est le plus énergétique comme l'illustre la figure 1.6 [2].

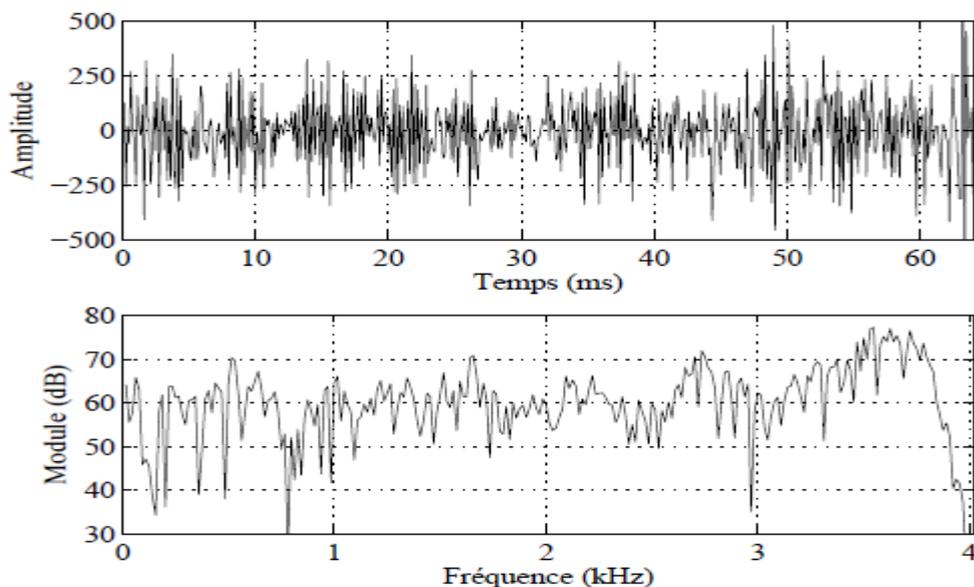


Figure. 1.6 : Forme d'onde d'une trame de signal vocal non voisé et son spectre d'amplitude.

### 1.4.3. Harmoniques

Les fréquences  $\omega = \frac{2\pi}{P}k$ , sont appelées harmoniques de la forme d'onde du signal glottique. ( $\frac{2\pi}{P}$ : Est la fréquence fondamentale = le pitch). Lorsque la période P diminue, l'espacement entre les harmoniques augmente. La détection de la période du pitch et les décisions V / NV sont élémentaires pour de nombreux codeurs vocaux. Parmi de nombreuses méthodes de calcul du pitch, la fonction d'auto-corrélation est utilisée pour la détection du fondamental comme indiquée sur la figure ci-dessous.

Pour un signal de parole, échantillonné à 8KHz, la plage des échantillons varie entre 20-130 (2.5-16ms).La fonction d'auto-corrélation est définie comme suit:

$$r(m) = \sum_{i=-\infty}^{\infty} x(i)x(i + m) \quad \text{Eq (1.1)}$$

### 1.4.4. Conduit vocal et formants

Il peut être approché par un filtre linéaire avec des résonances appelées formants (comme les résonances des tubes d'orgue et des instruments à vent). Les formants changent

avec différentes configurations de voies vocales correspondant à différentes cavités résonantes et donc à différents phonèmes.

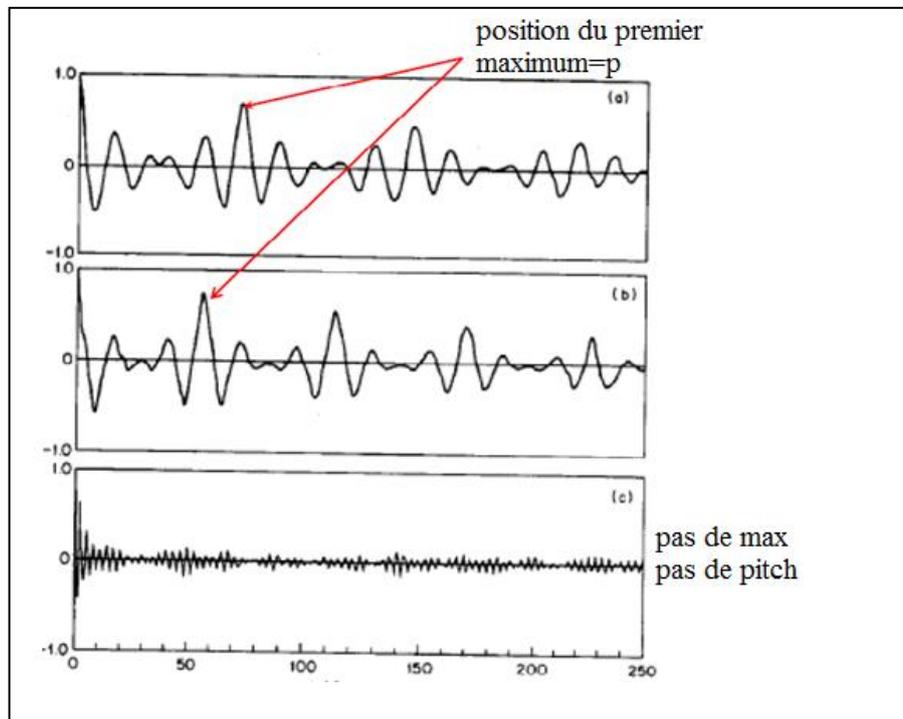


Figure. 1.7 : Fonction d'auto corrélation pour un son voisé (a) et (b) ; un son non voisé (c) avec une fenêtre rectangulaire de  $N=401$

En général, les fréquences des formants diminuent à mesure que la longueur du conduit vocal augmente.

- Par conséquent, un locuteur masculin a tendance à avoir des formants inférieurs à une femme qui a des formants inférieurs à un enfant.
- Avec une hypothèse de linéarité et d'invariance temporelle du champ vocal, et lorsque la source sonore se produit par la glotte, la forme d'onde vocale (c'est-à-dire la vitesse d'écoulement de l'air à la sortie du système vocal) peut être exprimée comme une convolution entre l'écoulement glottal et la réponse impulsionnelle des voies vocales (voir figure ci-dessous).

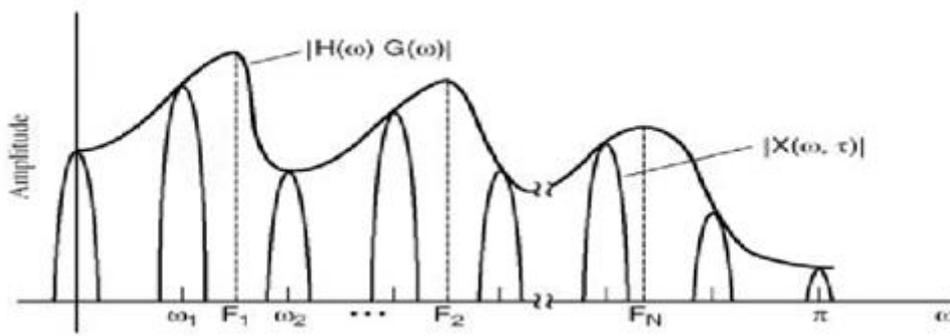


Figure 1.8 : Harmoniques de la source glottale, formant du conduit Vocal et enveloppe spectrale.

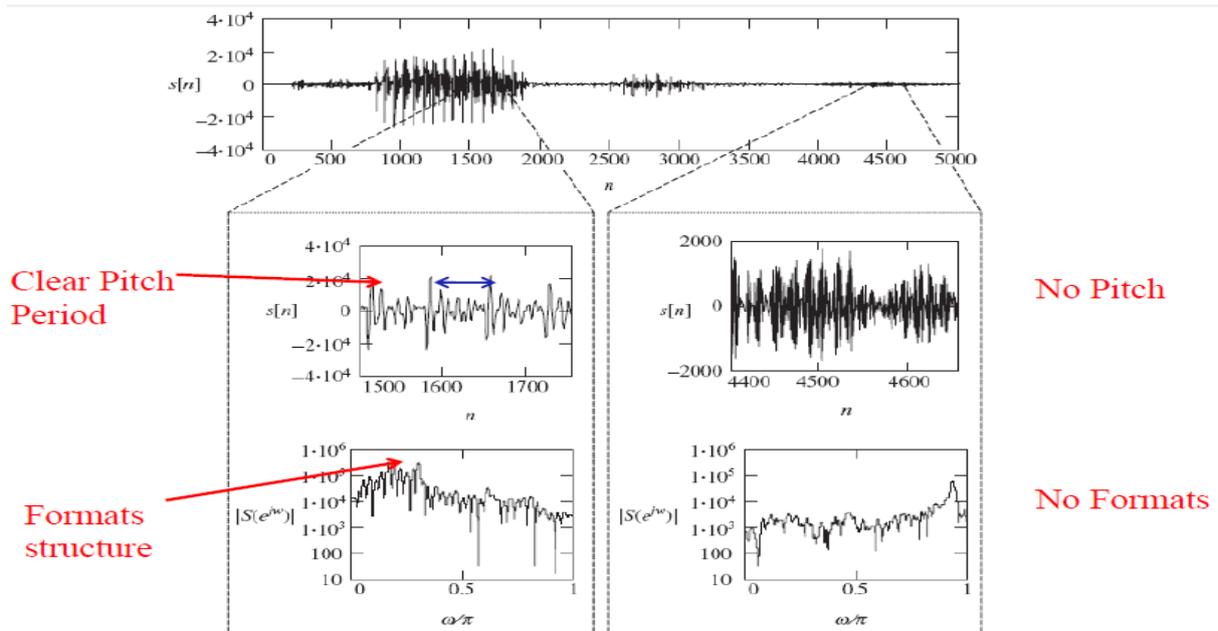


Figure 1.9 : Exemple de forme d'onde de parole (mâle) du mot "problèmes".

### 1.4.5. Spectrogramme

Les caractéristiques spectrales variables dans le temps du signal de parole peuvent être représentées graphiquement à l'aide d'un motif attirant comme le montre la Figure 1.10

- Axe vertical: fréquence, axe horizontal: temps

- La pseudo-couleur du motif est proportionnelle à l'énergie du signal (rouge: haute énergie)
- Les fréquences de résonance du conduit vocal apparaissent comme des "bandes d'énergie"
- Les intervalles sonores caractérisés par l'aspect strié (périodicité du signal)
- Les intervalles non voisés sont plus solidement remplis

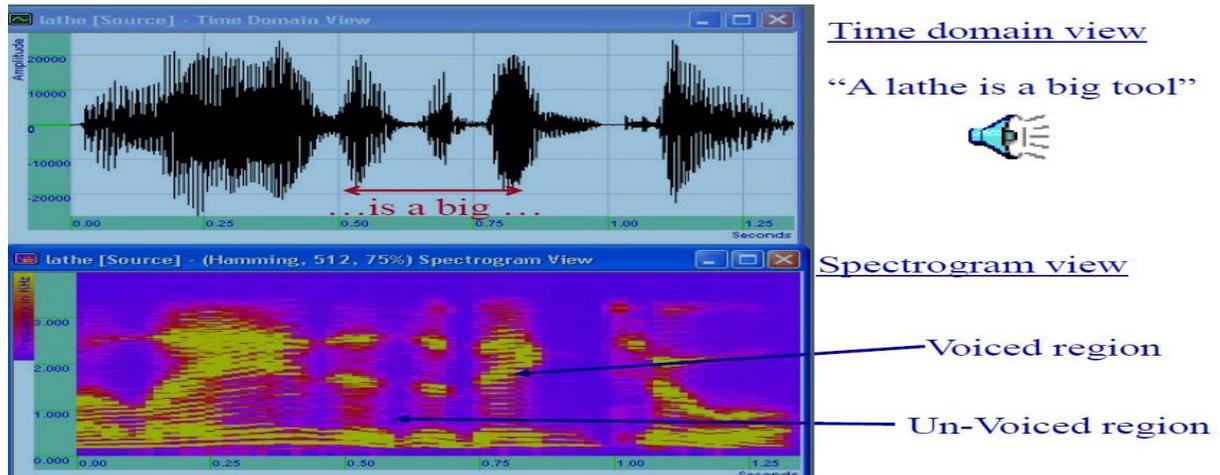


Figure 1.10 : Spectrogramme d'un signal de parole[12]

Pour comprendre l'utilité d'un spectrogramme, on utilise différentes fenêtres d'analyse. On retrouve les cas suivants :

- Spectrogramme à bande large: fenêtre d'analyse étroite (plus étroite que la période de la hauteur tonale) - les lignes verticales étroites correspondent aux périodes de hauteur successives.
- Spectrogramme à bande étroite: Une large fenêtre d'analyse (comprend plusieurs périodes de hauteur) - les lignes horizontales étroites sont des harmonies de hauteur.
- Les bandes jaunes décrivent les formants changeant dans le temps.

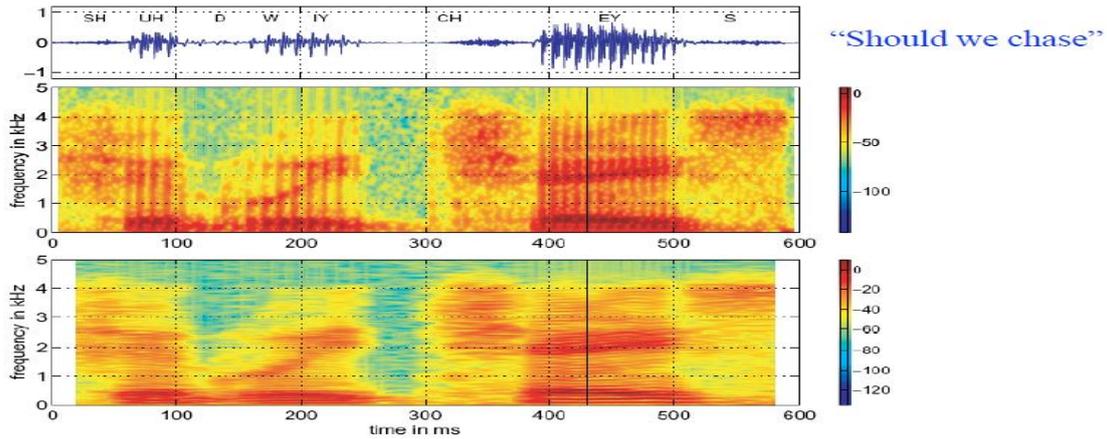


Figure 1.11 : Interprétation d'un Spectrogramme[12]

## 1.5. Application du traitement de la parole

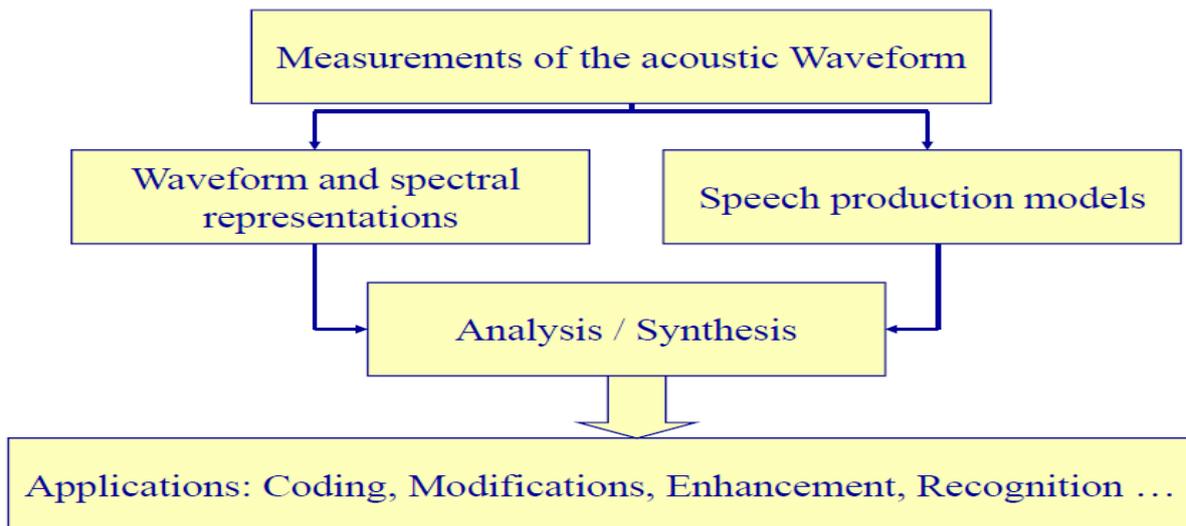


Figure 1.12 : Application du traitement de la parole.

### 1.5.1. Modifications

Le but est de modifier le signal de parole pour avoir une propriété souhaitée: échelle de temps, pas et changements spectrales.

Applications: montage de publicités de radio et de télévision dans un intervalle de temps alloué, synchronisation des présentations audio et vidéo, etc.

En outre, l'accélération de la parole est utilisée dans la lecture de messages, la messagerie vocale, et les machines de lecture et des livres pour les aveugles, tout en ralentissant la parole a une application à l'apprentissage d'une langue étrangère.

### 1.5.2. Amélioration du discours

L'objectif est d'améliorer la qualité de la parole dégradée. Une approche consiste à prétraiter la forme d'onde de parole (analogique) avant qu'elle ne soit dégradée. Une autre approche est le post-traitement: l'amélioration après la dégradation du signal.

- Réduction des bruits additifs dans la téléphonie numérique, communication des véhicules et communications aériennes.
- Réduction des milieux parasites et des haut-parleurs pour les malentendants.

### 1.6. Variabilité et non linéarité du signal vocal

La variabilité du signal de parole est due à la dynamique spécifique aux articulateurs. Nous soulignons le fait bien connu que le système phonatoire est variable dans le temps et, par conséquent, le signal de parole est un processus non stationnaire d'ordre supérieur.

Globalement, le signal de parole est non stationnaire.

Prenant comme base le modèle du filtre source, la variabilité de la parole vocale Peut être expliqué approximativement en considérant la variation temporelle de ce qui suit:

- La forme du train d'impulsion glottique et le pas, tous deux étant des fonctions non linéaires dans le conduit vocal.
- Fréquences fondamentales, largeurs de bande et amplitudes relatives des formats.
- Couplages entre les parties " source " et " filtre ", si elles sont modélisées avec précision.

Dans le cas des phonèmes vocaux, la variabilité non stationnaire est causée par:

1. la nature temporelle du système qui génère la parole ( le système phonatoire ).
2. la nature non linéaire et dynamique de la source de la parole (glotte).
3. la nature variable et non linéaire du système phonatoire, le générateur glottal a dynamique non linéaire, le conduit vocal est en continu évoluant dans le temps[1].

## 1.7. Observation et modélisation du conduit vocal

L'essentiel des études portant sur le conduit vocal a pour objectif de décrire sa forme générale ou la fonction de transfert acoustique qui en résulte. Pour cela, différentes méthodes d'imagerie ont été utilisées. Dans un premier temps, les imageries par rayon X ont beaucoup été utilisées.

Cette méthode permettait de visualiser les contours du conduit vocal, depuis le larynx jusqu'aux extrémités buccale ou nasale.

## 1.8. Système auditif humain

L'oreille effectue une analyse de fréquence du signal reçu et permet à l'auditeur de discriminer de petites différences de temps et de fréquence dans le son. La plage d'écoute s'étend dans l'intervalle : 16Hz -18KHz mais la sensibilité est dans la gamme : 2KHz -4KHz. La plage dynamique (du plus silencieux au plus fort) est d'environ 96 dB. La voix humaine normale perçue par notre système auditif est située environ dans l'intervalle 500 Hz à 2 kHz

### 1.8.1. Fonctionnement du système auditif humain

Il est constitué de trois oreilles comme illustré par la figure 1.13

- **L'oreille externe :** appelée aussi pavillon ou oreillette, permet d'attraper l'onde sonore. Après que les ondes pénètrent dans l'oreille externe, elles traversent le conduit auditif et se dirigent vers l'oreille moyenne. Le rôle du canal auditif externe est de protéger l'oreille en faisant du cérumen. Cette cire spéciale contient des produits chimiques qui combattent les infections qui pourraient blesser la peau à l'intérieur du canal auditif. Il recueille également la saleté pour aider à garder le canal auditif propre.
- **Le tympan :** est un morceau de peau mince tendue serré. Il est attaché au premier osselet, un petit os appelé le marteau, attaché lui aussi, à un autre minuscule os appelé l'enclume, qui est attaché à l'os le plus petit dans le corps, appelé l'étrier (étrier). Lorsque les ondes sonores se déplacent dans l'oreille et atteignent le tympan, elles font vibrer le tympan. Ces vibrations sonores sont portées aux trois minuscules os de l'oreille moyenne.

- **L'oreille moyenne :** Le principal rôle de l'oreille moyenne est de recevoir les ondes sonores, de les transformer en vibrations et de les transmettre à l'oreille interne. Elle aide également le tympan à gérer la pression: l'oreille moyenne est reliée à l'arrière du nez par un tube étroit appelé la trompe d'Eustache. Ensemble, la trompe d'Eustache et l'oreille moyenne maintiennent la pression d'air égale sur les deux côtés du tympan. Garder la pression de l'air égale est important afin que le tympan puisse fonctionner correctement et ne pas se blesser.
- **L'oreille interne :** Les vibrations de l'oreille interne vont dans la cochlée. C'est un petit tube enroulé dans l'oreille interne. Elle est remplie de liquide et doublée de cellules qui ont des milliers de petits poils microscopiques (15 000 à 20 000) à leur surface. Lorsque les vibrations sonores frappent le liquide dans la cochlée, il commence à vibrer. Différents types de sons donnent différents modèles de vibrations. Les vibrations provoquent le mouvement des poils sensoriels dans la cochlée. Ces vibrations sonores se transforment en signaux nerveux et sont transmis au cerveau par le nerf auditif, appelé aussi le « huitième nerf ».

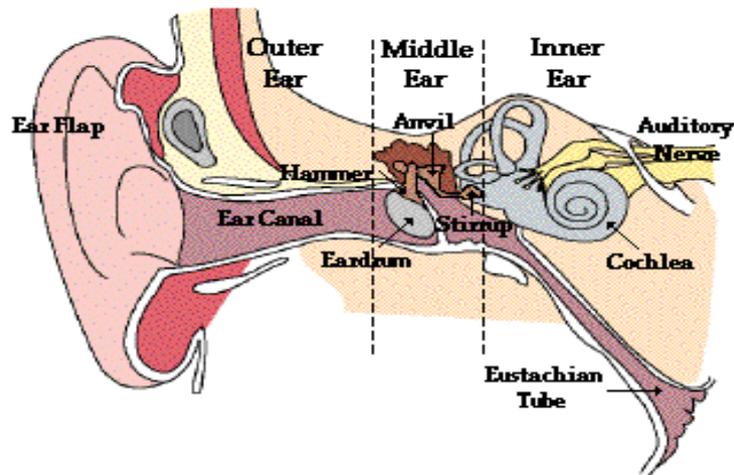


Figure1.13 : Système auditif humain.

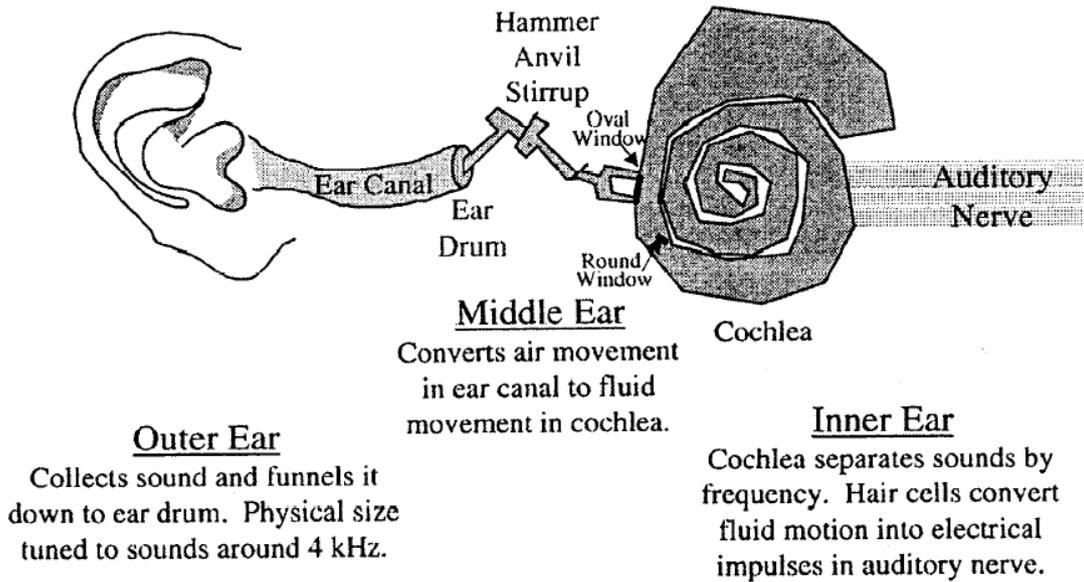


Figure1.14 : Les oreilles externe, moyenne et interne [4].

### 1.8.2. A l'intérieur de la cochlée

L'intensité minimale à laquelle les sons peuvent être perçus est appelée seuil d'audition. L'oreille est moins sensible au ton et au premier formant (F1) qu'aux formants supérieurs, au sens de l'intelligibilité.

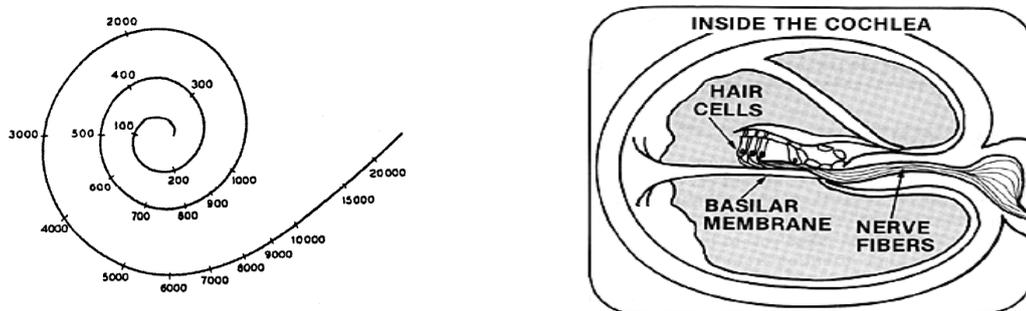


Figure1.15.Sensibilité fréquentielle le long de la membrane basilaire [5].

## **1.9. Conclusion**

La parole est un support d'information spécifique aux êtres humains pour se communiquer entre eux. Le signal vocal est produit par les fluctuations de la pression de l'air qui sont émises par l'appareil phonatoire. Le contenu d'un message parlé est analysé par le système auditif humain et les informations résultantes sont transmises au cerveau qui les interprète. Un signal vocal est perçu uniquement par son intelligibilité.

### 2.1. Introduction

La réduction de bruit est devenue une nécessité dans les communications modernes ainsi que dans bien d'autres applications. L'amélioration de la parole dégradée par le bruit, ou réduction du bruit, est un champ plus important dans le traitement de la parole. Il est utilisé dans de nombreuses applications telles que les téléphones mobiles, les systèmes de téléconférence, la reconnaissance de la parole et les aides auditives.

Dans ce chapitre nous essayons de décrire le signal de bruit et les traitements qui permettent sa suppression / réduction. Premièrement, nous expliquons la gêne que pose un bruit lorsqu'il s'attache à un signal utile par exemple la parole. Après, les types et les sources de bruits sont exposés brièvement.

Après avoir parlé de la nécessité de l'opération de débruitage, la section 2.5 présente les différentes familles de techniques de réduction de bruit et explique l'essor des techniques dites : soustraction spectrale, adaptative, à modèles, avec masquage fréquentiel.

Dans la partie 2.6 nous verrons les modèles de bruit utilisés dans l'élaboration des algorithmes de réduction de bruit. L'évaluation de la qualité d'un débruiteur est nécessaire, ainsi des critères subjectifs et objectifs pour juger le signal débruité sont mentionnés.

Enfin, nous terminerons notre chapitre par une conclusion qui résume les points abordés.

### 2.2. L'inconvénient du bruit

Le bruit est un signal indésirable dans la communication parlée. Dans un environnement bruyé, l'auditeur est le plus gêné car, contrairement au locuteur, il n'a aucun contrôle sur le milieu acoustique. Le locuteur garde un certain contrôle sur le bruit ambiant. Il a la possibilité de hausser la voix ou de se réfugier dans un endroit moins influencé par le bruit.

Dans le cas d'une communication mains-libres, il peut aussi focaliser son oreille sur le signal utile grâce aux capacités de localisation spatiale de l'oreille alors que l'auditeur subit complètement la perturbation sonore. L'ensemble du champ sonore (des sons provenant de

toutes les directions) est donc intégré et restitué en une somme des perturbations. L'information de spatialisation ayant disparu, l'auditeur ne peut donc pas séparer l'information utile des différentes sources de bruit bien que l'oreille humaine en soit normalement capable. De plus, la superposition du bruit et de la parole réduit l'intelligibilité du message ce qui demande un effort constant de la part de l'auditeur et le fatigue rapidement.

Il faut souligner aussi que, dans tout réseau de communication, il y a une étape de codage du signal de parole qui est largement basée sur ses propriétés. En présence de bruit le signal à coder s'éloigne du modèle utilisé pour l'opération de codage-décodage et il en résulte une dégradation d'autant plus gênante que le niveau de bruit est élevé. Le timbre de la voix est altéré et le bruit est codé de façon peu naturelle ce qui le rend encore plus désagréable [9].

### **2.3. Types et sources de bruits**

La réduction de bruit est un traitement de la parole qui vise à obtenir la parole sans bruit. Pour concevoir des algorithmes efficaces et généralisés pour l'amélioration de la parole, il est nécessaire d'avoir des connaissances sur le bruit qui pourrait être présent dans la parole.

#### **2.3.1 Types de bruits**

Mathématiquement, le bruit peut être parfois modélisé par sa distribution de probabilité de ses valeurs. Il peut être représenté par une distribution de type gaussien, Laplacien ou Gama dans le cas de la modélisation de parole.

Le bruit peut être classé dans l'une des catégories suivantes : adaptatif, additif, additif aléatoire, aéroport, arrière-plan, voiture, Cross-Noise, Salle d'exposition, usine, babble multi-locuteurs (Balbutiements multiples), musical, naturel, babble non-stationnaire, bureau, quantile-based, restaurant, rue, ambiance, aléatoire, station de train, gaussien blanc, etc [8].

Le niveau du SNR peut être un indicateur important de la stratégie à utiliser dans les algorithmes d'amélioration. En présence d'un SNR de haut niveau et d'un SNR de faible niveau, l'algorithme peut être commuté d'un schéma d'amélioration à un autre après avoir estimé le SNR présent dans la parole.

En fonction des caractéristiques de stationnarité du bruit, la technique doit être conçue pour traiter séparément le bruit stationnaire et le bruit non stationnaire.

Dans le cas du débruitage et de la reconnaissance de la parole pour les applications de téléphonie mobile et de téléphonie mains libres à bord de véhicules, on s'intéresse aux bruits additifs, stationnaires ou non stationnaires, et decorrélés avec la parole (indépendance au sens statistique), tels que le bruit de conversation appelé Babble, le bruit de voiture appelé (car) et le bruit blanc gaussien (ce dernier est souvent utilisé mais peu réaliste)[9].

### 2.3.2 Sources de bruits

Il y a des bruits dont la source est connue et facile à modéliser et à analyser. En dépit des sources connues de bruit, il devient difficile de réduire le bruit quand plus qu'un seul type de bruits est présent. Dans des applications comme le casque sans bruit, il peut devenir facile de surmonter le problème du bruit quand les oreilles sont isolées de la région externe dans une large mesure.

Dans l'application de la voiture, les bruits générés par le moteur, le vent et les essuie-glaces doivent être manipulés intelligemment. Les appareils mains libres, le bruit le plus dominant est l'écho de la voix. Dans une application de réunion de salle fermée, les caractéristiques acoustiques des murs peuvent générer de multiples échos de voix. Dans les réunions de foule ou en plein air, beaucoup de gens peuvent parler simultanément et il y a chevauchement avec la parole principale. Dans une zone de circulation, le bruit peut être créé à partir des autres véhicules et buffle impulsif. En voiture ou à la maison, parfois la musique de fond peut devenir source de bruit pour la parole utile. Dans le fonctionnement mains libres des téléphones cellulaires dans les véhicules, le signal vocal à transmettre peut être contaminé par la réverbération et le bruit de fond [7].

### 2.3.3 Nature et caractéristiques du bruit

Le bruit est un signal nuisible qui se superpose au signal utile en un point quelconque d'une chaîne de mesure ou d'un système de transmission. Il constitue donc une gêne dans la compréhension du signal utile, qui est dans notre cas, la parole. En physique, en acoustique et en traitement du signal, bien que le bruit soit, par nature, aléatoire, il possède certaines caractéristiques statistiques, spectrales ou spatiales. Le tableau 2.1, représente les différentes classes auxquelles un bruit peut appartenir [9].

Propriétés	Types
Structure	Continu/Impulsif/Périodique
Type d'interaction	Additif/Multiplicatif/Convolutif
Comportement temporel	Stationnaire/Non-stationnaire
Bande de fréquence	étroite/large
Dépendance	Corrélé/ Décorrélé
Propriétés statistiques	Dépendant/Indépendant
Propriétés spatiales	Cohérent/Incohérent

Tableau 2.1 : Différentes classes du bruit

## 2.4. Nécessité de l'opération de débruitage

Pour transmettre le signal à la partie distante avec une certaine intelligibilité, il est important de développer un algorithme de réduction du bruit, qui vise à extraire des informations utiles du signal bruité. La complexité de cette tâche est toujours liée à la nature du bruit et au type du signal. Cette complexité est liée la plupart du temps à la stationnarité ou à la non stationnarité et du signal et du bruit. Un bruit blanc est souvent plus simple à éliminer qu'un bruit coloré [6].

La réduction du bruit est principalement appliquée dans le domaine du traitement du son (de la parole ou de la musique), tel que :

- La téléconférence et la vidéoconférence dans un environnement bruyant.
- La téléphonie: traitement dans l'usine (stations classiques et des appareils portables, etc.)
- Les appareils mains-libres (ordinateurs de bureau, téléphone mobile, les terminaux fixes embarqués dans les voitures, etc.)
- L'enregistrement du son dans les lieux publics (chemins de fer, les aéroports, la gare de la rue, etc.)

## 2.5. Familles des techniques de débruitage

Les algorithmes d'amélioration de la parole pour la réduction du bruit peuvent être classés en trois catégories fondamentales: les techniques de filtrage, de restauration spectrale et des méthodes basées sur des modèles [3].

Le bruit se divise principalement en quatre catégories: bruit additif, d'interférence, de réverbération et d'écho.

Ces quatre types de bruit ont conduit à l'évolution de quatre grandes catégories de techniques de traitement du signal acoustique. Elles comprennent, la réduction du bruit / amélioration de la parole, la séparation de source, la déréverbération de la parole et annulation / Suppression d'écho.

Le choix d'une technique de réduction de bruit dépend en premier lieu du nombre d'observations du signal disponibles. Dans le cas monovoie, seule l'observation bruitée est disponible. Une étape commune et indispensable à la majorité des techniques de réduction de bruit consiste alors à extraire des connaissances statistiques sur le bruit qui seront ensuite exploitées pour le supprimer. Ceci peut d'ailleurs se révéler particulièrement difficile lorsque le bruit est non-stationnaire et que son niveau est important.

Les méthodes de réduction du bruit sont divisées en quatre classes d'algorithmes: soustraction spectrale, sous espace signal, modèle statistique et filtrage de Wiener [8]. La figure 2.1 illustre les différentes approches de filtrage appliquées dans la réduction du bruit.

Single Channel Speech Enhancement Techniques

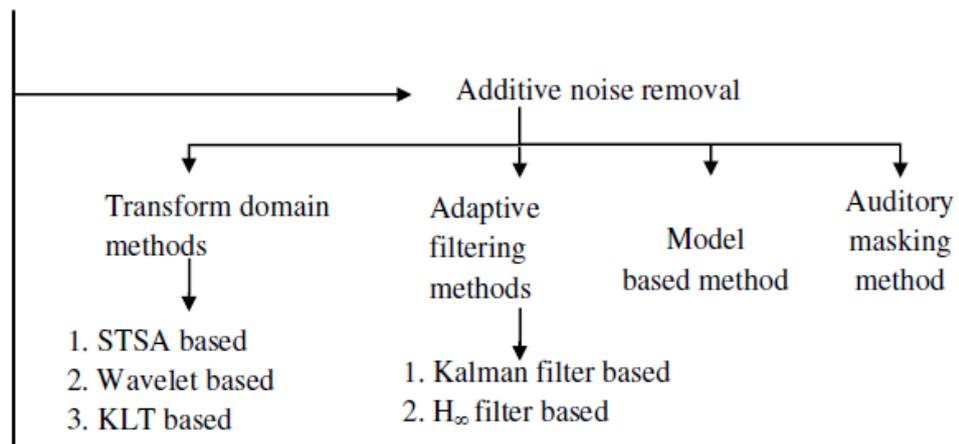


Figure 2.1: illustre les différentes approches de filtrage appliquées dans la réduction du bruit

### 2.5.1 Méthode de soustraction spectrale

Dans les approches par atténuation spectrale à court terme, le traitement se fait dans un domaine transformé qui est généralement celui de Fourier. La TFD (transformée de

Fourier discrète) reste la plus utilisée puisque des algorithmes optimisés de calcul très efficaces existent (FFT: transformée de Fourier rapide).

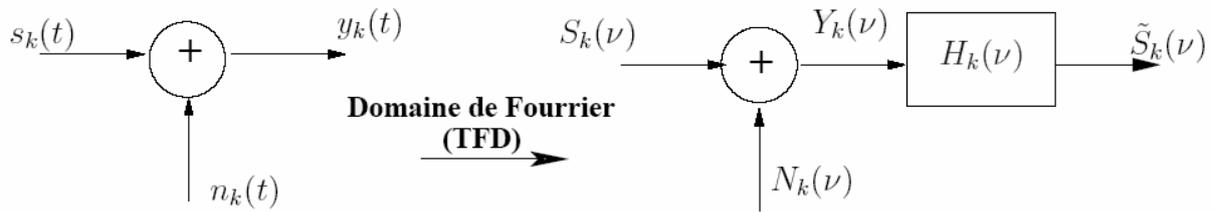


Figure 2.2 La transformée de Fourier discrète

Le principe du débruitage par soustraction spectrale [6], consiste à éliminer le bruit dans le domaine spectral. En effet, lorsqu'un signal  $x(t)$  est dégradé par un bruit additif  $n(t)$ , le signal bruité est exprimé par:

$$y(t) = x(t) + n(t) \quad \text{Eq (2.1)}$$

Dans le domaine fréquentiel, l'équation (2.1) devient :

$$Y(f) = X(f) + N(f) \quad \text{Eq (2.2)}$$

Où ;  $X(f)$  et  $N(f)$  sont respectivement les transformées de Fourier du signal et du bruit.

L'estimation de l'amplitude fréquentielle du bruit doit être effectuée afin de la soustraire de la valeur du signal bruité. La soustraction spectrale est traduite par l'équation suivante :

$$\tilde{X}(f) = Y(f) - \tilde{N}(f) \quad \text{Eq (2.3)}$$

Si on utilise le spectre d'énergie, l'énergie du signal bruité est évaluée par l'expression ci-dessous:

$$|Y(f)|^2 = |X(f)|^2 + |N(f)|^2 + X(f)N^*(f) + X^*(f)N(f) \quad \text{Eq (2.4)}$$

En utilisant l'hypothèse que le bruit et le signal sont indépendants, alors l'équation (2.4) s'écrit:

$$|Y(f)|^2 = |X(f)|^2 + |N(f)|^2 \quad \text{Eq (2.5)}$$

Et par conséquent, on déduit l'équation :

$$|\tilde{X}(f)|^2 = |\tilde{Y}(f)|^2 - |\tilde{N}(f)|^2 \quad \text{Eq (2.6)}$$

Pour la reconstruction du signal débruité dans le domaine temporel, il est nécessaire d'avoir la phase du signal original. Cette dernière est obtenue par extraction de la phase du signal bruité lui-même qui est utilisée avec la transformée de Fourier inverse sur l'estimé de l'amplitude fréquentielle pour récupérer le signal débruité.

Cette technique permet d'éliminer « le bruit de fond »

### 2.5.2 Filtrage de Wiener

Une autre approche de débruitage consiste à appliquer au signal bruité à l'entrée d'un filtre optimal appelé filtre de Wiener . Cette approche utilise également le spectre d'énergie et par conséquent, on a besoin de l'estimé du signal dans le domaine spectral. Comme dans le cas de la soustraction spectrale, cet estimé peut être obtenu par l'équation (2.6).

Ce type de filtrage consiste à obtenir une estimation du signal original suivant le critère « minimum de l'erreur quadratique moyenne ». Ce critère conduit au filtre de Wiener dont la réponse fréquentielle est donnée par :

$$H(f) = \frac{|X(f)|^2}{|Y(f)|^2} = \frac{|X(f)|^2}{|Y(f)|^2 + |N(f)|^2} \quad \text{Eq (2.7)}$$

Pour cette méthode, il est essentiel de connaître un estimé du bruit pour évaluer  $|N(f)|^2$ . Concernant  $|X(f)|^2$ , cela est obtenu par soustraction de l'estimation de  $|N(f)|^2$  du spectre d'énergie du signal bruité et conduit au résultat suivant :

$$H(f) = \begin{cases} 1 - \frac{1}{SNR} & \text{si } 1 - \frac{1}{SNR} > 0 \\ 0 & \text{sinon} \end{cases} \quad \text{Eq (2.8)}$$

où: SNR est le rapport signal-bruit en entrée.

### 2.5.3 Elimination adaptative de bruit

Cette technique générale de l'élimination adaptative du bruit (ANC : Adaptive Noise Cancellation) a été utilisée avec succès pour résoudre un grand nombre de problèmes liés à des signaux tels que les signaux de la parole, les interférences dans les lignes de transmission téléphonique ainsi que la théorie des antennes. Elle est utilisée surtout pour éliminer « le bruit impulsif ». Cette technique utilise des algorithmes nécessitant une source d'entrée primaire et une source de référence secondaire (voir la figure 2.3). La source primaire contient le signal de parole bruité, tandis que la source de référence est une source de bruit utilisée comme référence pour avoir une estimation plus exacte du bruit grâce à un filtre adaptatif. La sortie de ce filtre est ensuite soustraite du signal de la source primaire pour produire le signal débruité. La sortie globale est utilisée pour contrôler les ajustements effectués sur les coefficients du filtre adaptatif. Le critère utilisé pour ajuster ces coefficients est celui minimisant l'erreur quadratique moyenne.

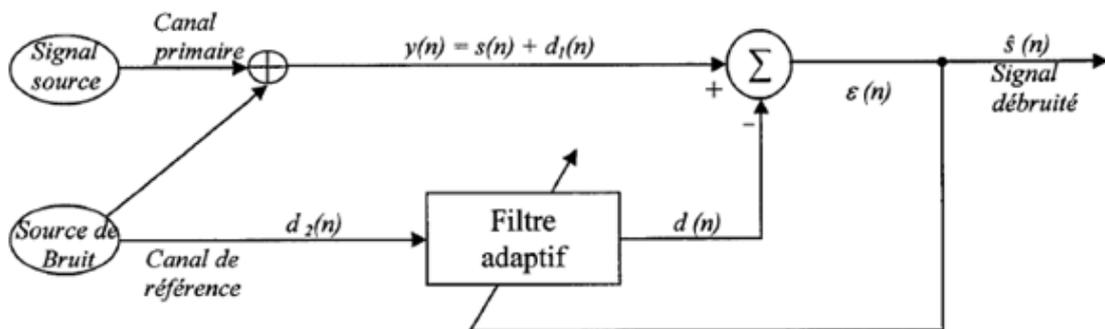


Figure 2.3 diagramme de la méthode d'élimination adaptative du bruit(AVC)

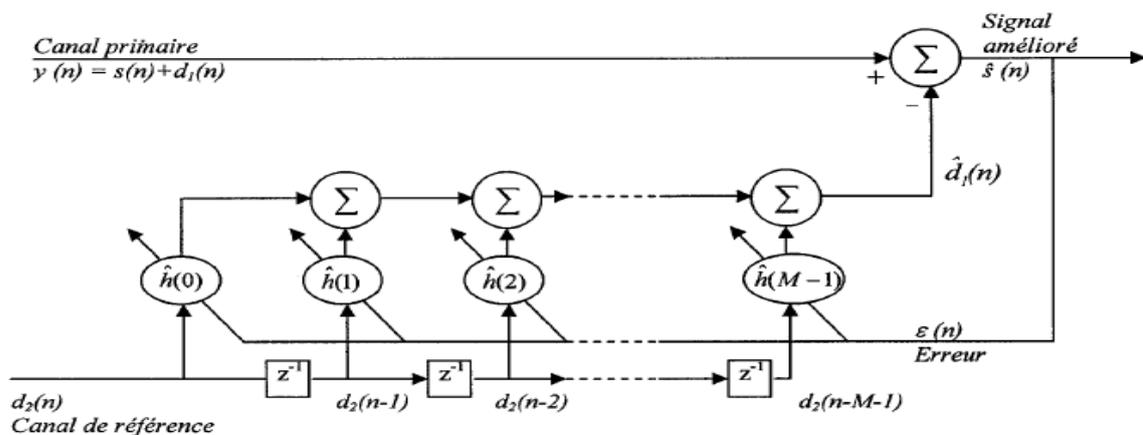


Figure 2.4 : Le filtre adaptatif LMS

Contrairement aux techniques de la soustraction spectrale et au filtrage adaptif de Wiener, cette technique requiert un canal de référence. Le filtre utilisé est un filtre FIR utilisant l'algorithme LMS "Least Mean Square" (voir la figure 2.4). Reprenons le schéma ci-dessus pour illustrer le problème. Le but est de faire un estimé fidèle du bruit ' $d(n)$ ' qui se rapproche de la vraie valeur du bruit ' $d_1(n)$ ' et cela en utilisant une référence  $d_2(n)$  et le critère du minimum MSE "Mean Square Error" entre  $d_1(n)$  et  $d(n)$  qui revient en final à estimer le signal  $y(n)$ . Soient  $\hat{h}(i)$  les estimées des coefficients du filtre,  $i = 0, 1 \dots M-1$ , ainsi on peut écrire l'équation :

$$\hat{d}_1(n) = \sum_{i=0}^{M-1} \hat{h}(i) d_2(n-1) \quad \text{Eq(2.9)}$$

$$\begin{aligned} \hat{h} &= \arg \min_h E\{[y(n) - \hat{y}(n)]^2\} \\ &= \arg \min_h E\left\{[y(n) - \sum_{i=0}^{M-1} \hat{h}(i) d_2(n-1)]^2\right\} \end{aligned} \quad \text{Eq(2.10)}$$

Où  $E(\cdot)$  désigne l'espérance mathématique d'une expression.

#### 2.5.4. Techniques basé sur le modèle de la parole

Plusieurs techniques de traitement des signaux ont été développées pour modéliser la manière dont l'oreille humaine réagit aux stimuli acoustiques [3]. Ces modèles auditifs prouvent la suppression significative des effets du bruit.

Les systèmes de reconnaissance vocale utilisent des techniques d'appariement de formes ayant des paramètres dérivés de la forme d'onde et le spectrogramme à court terme de la parole.

Jusqu'à présent, les systèmes qui utilisent des coefficients de prédiction linéaire (LPC) de modèles de filtre tout-polaire appropriés du conduit vocal et des coefficients cepstraux de fréquence à échelle Mel(MFCC), sont les plus réussis. En reconnaissance de signaux bruyants, les valeurs de ces coefficients seront perturbées, trame par trame, par leurs valeurs idéales et sans bruit. Les approches de modèles auditifs sont en effet des techniques de paramétrage alternatives à la représentation LPC ou MFCC de la parole. L'approche la plus

simple, est de précéder tout traitement LPC ou MFCC standard d'un signal par un processus de réduction du bruit adaptatif. Le problème de base est de réduire le bruit externe sans perturber les composantes de bruit non voisées et de faible intensité du signal de parole lui-même [3].

### 2.5.5. Méthode de masquage

Les méthodes classiques, comme la soustraction spectrale ou le filtrage de Wiener, réussissent à réduire le bruit additif, mais en contrepartie, introduisent un bruit résiduel (bruit musical) gênant pour la perception humaine. Le signal observé est donc considéré comme la somme du signal de parole et du bruit ambiant. Ce modèle omet tout bruit convolutif, électrique ou de quantification.

En codage de la parole, les progrès ont exploité les propriétés de la perception auditive de l'être humain à travers le phénomène de masquage. Ce phénomène montre que notre cerveau n'est pas objectif dans l'interprétation des sons. En effet, l'existence physique d'un son ne garantit pas sa perception auditive par notre cerveau. Ceci est dû à la possibilité qu'il soit masqué par la présence d'un autre son plus puissant. Il est donc possible de conserver un bruit tant que celui-ci reste inaudible. Il existe deux types de masquage : fréquentiel et temporel. Dans le cadre des applications audio, notamment le débruitage, le phénomène le plus exploité est le masquage fréquentiel. Le masquage temporel est rarement utilisé car difficile à modéliser et à exploiter ; en plus il est de faible influence par rapport au masquage fréquentiel [9].

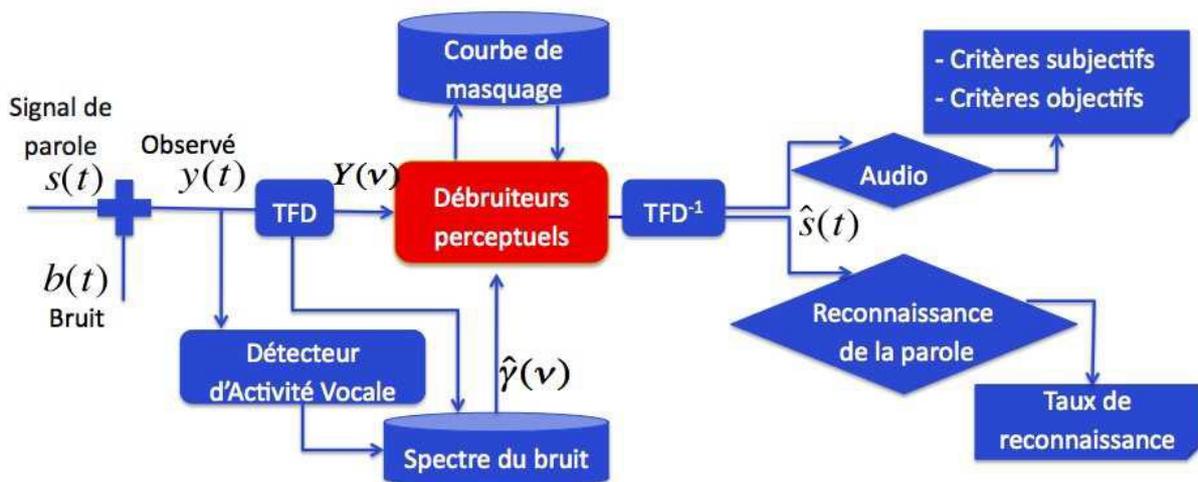


Figure 2.5 : Principe du filtrage perceptuel [9]

Cette technique utilise les propriétés de l'audition humaine via la modélisation de certaines de ces propriétés déterminantes. Elle distingue entre ce qui est audible et non audible. La modélisation du phénomène de masquage est réalisée par le calcul de la courbe de masquage. La Figure 2.6 montre comment un signal dit masquant rend inaudible un signal d'énergie plus faible appelé signal masqué (masquage psychoacoustique). Ce phénomène s'explique par une augmentation du seuil d'audition due au signal masquant. Les limites extrêmes du système auditif humain sont environ de 20 et 20000 Hz.

Si nous entendons un son fort, puis il s'arrête, il faut un peu de temps jusqu'à ce que nous puissions entendre un ton doux à proximité. Ce phénomène est appelé **Masquage temporel**.

L'objectif du débruitage perceptuel est de réduire le bruit et ne pas apporter plus de distorsion sur le signal de parole. Pour éviter les distorsions, il opère uniquement dans les fréquences où le bruit est audible.

Si le bruit est initialement inaudible, il n'est pas pris en compte par le débruiteur perceptuel, donc risque de devenir audible et gênant si les masquant de ce bruit sont filtrés.

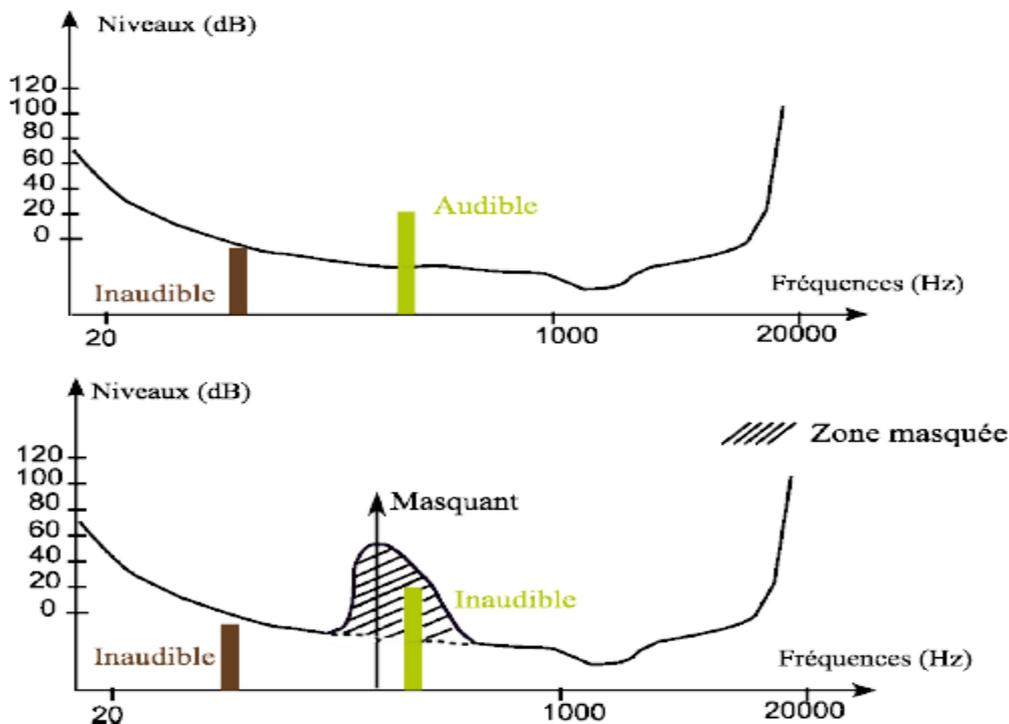
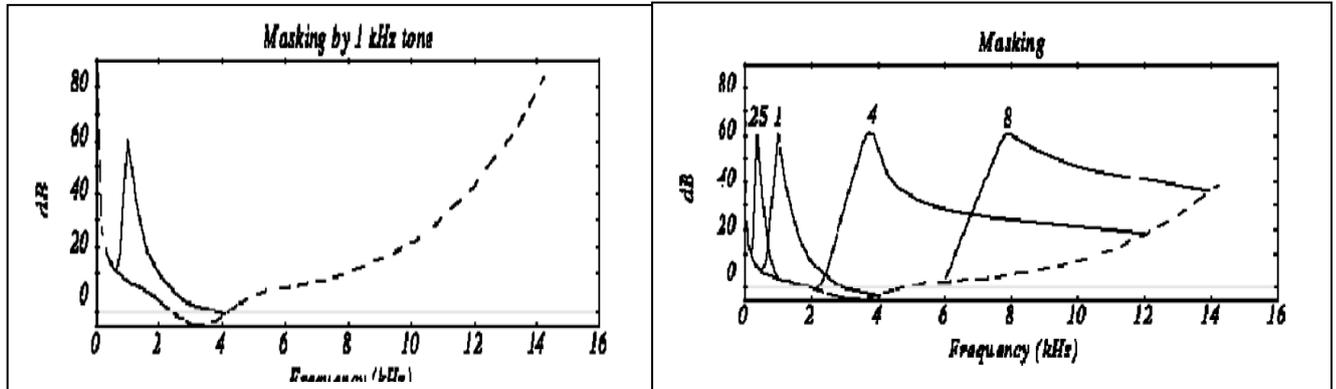


Figure 2.6 : Seuil absolu d'audition et masquage [9].

La figure 2.7 montre le seuil d'audition en fonction de la fréquence du ton d'essai. On applique une tonalité de 1kHz (tonalité de masquage) au niveau fixe (60dB). Après, la tonalité

de test est utilisée à un niveau différent (par exemple, 1,1 kHz) et on relève le niveau jusqu'à ce qu'on puisse le distinguer. Si on répète l'opération pour différentes fréquences de masquage on obtient le graphe suivant :



(a)

(b)

Figure 2.7: Le masquage fréquentiel (a) à 1kHz (b) à plusieurs fréquences

## 2.6. Modèles de bruit et évaluation de la qualité

### 2.6.1 Modèles de bruit

La plupart des travaux analytiques sur la réduction du bruit sont basés sur le bruit blanc gaussien. C'est principalement parce qu'il est relativement facile à modéliser et les résultats peuvent être exprimés en termes simples. Le bruit blanc gaussien est cependant aussi un modèle raisonnable des caractéristiques de bruit de nombreux canaux de communication.

L'environnement de reconnaissance vocale téléphonique est également soumis à d'autres sources de bruits et d'interférences non gaussiennes. Idéalement, le bruit de bureau et les interférences des haut-parleurs (en arrière-plan) seraient annulé par des techniques différentes.

### 2.6.2 Evaluation de la qualité du débruitage

La mesure générale de la qualité d'un système de réduction du bruit est basée sur l'amélioration du rapport signal / bruit (amélioration SNR). C'est une mesure utile, bien que grossière, car elle ne tient pas compte de l'effet du système de réduction du bruit sur les sons particulièrement sensibles, tels que les sons non-voisés susmentionnés.

Le SNR peut être calculé de plusieurs façons pour un signal de parole nominale silencieuse (dit «propre») auquel le bruit est ajouté, selon que des périodes de silence sont ou non incluses, et sur le bilan énergétique des échantillons de parole [3].

On entend par qualité de la parole l'intelligibilité (en téléphonie), le naturalisme, l'identification du locuteur et la perception.

La dégradation de la qualité peut être quantifiée par des critères objectifs ou subjectifs tels que :

- SNR, MSE etc. ne sont pas des mesures de perception.
- Critères subjectifs (tests d'écoute: MOS) sont coûteux et prennent du temps.
- PESQ: un critère objectif standard.

### 2.6.3 Valeurs typiques de la ligne téléphonique

- Bande limitée: 300 - 3.3 KHz
- La sonnerie est un signal sinusoïdal de fréquence 50Hz et de valeur efficace comprise entre 50V et 80V superposé au signal continu  
La tonalité est un signal sinusoïdal de fréquence 440Hz
- En partant d'une qualité assez limitée liée à la bande passante très réduite du téléphone, il est envisageable, grâce aux progrès rapides des technologies de traitement numérique du signal en temps réel
- Échantillonnage=nombre d'échantillons codés = 8000khz
- temps de propagation des terminaux de téléphonie 3,1 kHz.

### 2.7. Conclusion

La parole, étant un moyen de communication fondamental pour les humains. Elle a été intégrée dans diverses applications telles que la reconnaissance vocale, la voix-locuteur-distant et d'autres formes de communications personnelles. Il y a tellement d'applications de la parole encore à être loin de la réalité juste en raison de l'absence de mécanisme efficace et fiable de suppression du bruit et de préserver ou d'améliorer l'intelligibilité pour les signaux de la parole. Les grandes catégories de techniques d'amélioration de la parole peuvent être

énumérées comme des techniques de filtrage de la parole, des méthodes adaptatives d'annulation de bruit. Dans ce chapitre, nous sommes intéressés à la présentation des techniques classiques et récentes de débruitage. Nous avons également exposé les types et les sources de bruit qui peuvent être considérés dans les améliorations de la parole.

Il est à noter que d'autres transformées (autre que la TF) sont exploitables, entre autres la TCD (transformée en cosinus discret) et la transformée en ondelettes. Cette dernière permet une analyse multirésolution du signal. Le troisième chapitre sera consacré à l'étude et à l'utilisation de cette transformée pour le débruitage d'un signal de parole.

### 3.1. Introduction

La plupart des signaux du monde réel ne sont pas stationnaires, et dans l'évolution de leurs caractéristiques (statistiques, fréquentielles, temporelles, spatiales) que réside l'information pertinente qu'ils contiennent. A titre d'exemple, on cite les signaux vocaux et les images.

L'analyse avec la Transformée de Fourier(TF) propose une approche globale de ce type de signaux. La première section de ce chapitre présente des rappels sur la TF et discute la notion de localisation temporelle (ou spatiale pour des images) qui disparaît dans l'espace de Fourier. Ensuite, nous allons parler sur une transformation qui renseigne sur le contenu fréquentiel tout en préservant la localisation temporelle afin d'obtenir une représentation temps/fréquence ou espace/échelle du signal, qui est la Transformée en Ondelettes(TO). Dans la quatrième section, nous discutons la notion de multi résolution. Ensuite, nous présentons une brève théorie des Transformées en Ondelettes Continue(TOC) et discrète ((TOD).

Après, les propriétés et les familles des ondelettes seront abordés. Dans la section dix, nous essayerons de développer une analyse d'un signal avec l'Ondelette de Haar.

Avant de conclure ce chapitre, nous mentionnons quelques applications des ondelettes.

### 3.2. Transformée de Fourier

Souvent, l'information qui ne peut pas être distinguée dans le domaine temporel est facilement visible dans le domaine fréquentiel. Notons que la transformée de Fourier indique, qu'une certaine composante de fréquence existe ou pas. Cette information est indépendante de celle qui cherche où dans le temps cette composante apparaît. Il est donc très important de savoir si un signal est stationnaire ou pas, avant de le traiter avec la FT.

La transformée de Fourier d'un signal  $f(t)$  est donnée par :

$$F(\omega) = \int_{-\infty}^{+\infty} f(t) e^{-j\omega t} dt \quad \text{Eq (3.1)}$$

Pour limiter le domaine d'intégration temporel (ou spatial), on utilise une fonction «fenêtre» que l'on pourra faire glisser pour explorer le signal ; on obtient ainsi la transformée de Fourier à fenêtre glissante.

### 3.3. Transformée de Fourier à court terme

Cette transformation consiste à faire des analyses locales d'un signal. Nous balayons ce dernier par des fenêtres étroites, tel que les parties du signal vu à travers ces fenêtres sont en effet stationnaires. Dans le but de réaliser le meilleur compromis entre les résolutions temporelles et fréquentielles, Gabor(1946) eut l'idée de proposer la STFT (Short Time Fourier Transforme), une transformée dont l'expression est la suivante [6] :

$$F_G(\omega) = TF^{glisse}(f(t, \omega)) = \int_{-\infty}^{+\infty} f(s) g^*(s - t) e^{-j\omega s} ds \quad \text{Eq (3.2)}$$

Si on pose  $\psi_{t,\omega}(s) = g(s - t)e^{-j\omega s}$  Eq (3.3)

$\psi_{t,\omega}^*(s)$  étant la conjuguée de  $\psi_{t,\omega}(s)$ .

La fonction  $\psi_{t,\omega}(s)$  est une fonction fenêtre qui satisfait la condition suivante :

$$\hat{\psi}(0) = \int_{-\infty}^{+\infty} \psi(s) ds \neq 0 \quad \text{Eq (3.4)}$$

Où  $\hat{\psi}(\omega)$  est la transformée de Fourier de  $\psi(t)$ .

On peut interpréter cette transformée comme la projection de  $f(t)$  sur la «base» des fonctions fenêtres glissantes  $\psi$  :

$$TF^{glisse}(f(t, \omega)) = \langle f, \Psi_{t,\omega} \rangle \quad \text{Eq (3.5)}$$

Pour des fonctions réelles, le produit scalaire  $\langle f, g \rangle$ , de  $f(t)$  et la fenêtre glissante est défini par:

$$\langle f, g \rangle = \int_{-\infty}^{+\infty} f(x)g(x)dx \quad \text{Eq (3.6)}$$

Malgré son immense succès, cette technique a plusieurs défauts, en particulier

• son manque évident de *localisation temporelle*. La TF, comme la TO est une transformation réversible, c'est à dire qu'elle permet des « aller-retour » entre le signal bruit et le signal traité (transformé). Cependant, seulement l'un des deux est disponible à un instant donné. Aucune information de fréquence n'est disponible dans le domaine temporel et aucune information temporelle n'est disponible dans la FT du signal.

L'analyse de Fourier ne permet pas l'étude de signaux dont la fréquence varie dans le temps. De tels signaux nécessitent la mise en place d'une analyse temps-fréquence qui permettra une localisation des périodicités dans le temps et indiquera donc si la période varie d'une façon continue, si elle disparaît puis réapparaît par la suite, etc.

En 1982, Morlet ouvre la voie conduisant à la solution en construisant la Transformation en Ondelettes

### 3.4 La multi résolution

Les ondelettes s'adaptent en fonction des caractéristiques recherchées, ainsi pour les hautes fréquences (l'ondelette est très contractée) et pour les basses fréquences (l'ondelette s'étire). On parlera alors de multi résolution. La figure 3.1 illustre ce phénomène et donne un aperçu sur la représentation temps-échelle de la STFT et de la transformée en ondelette.

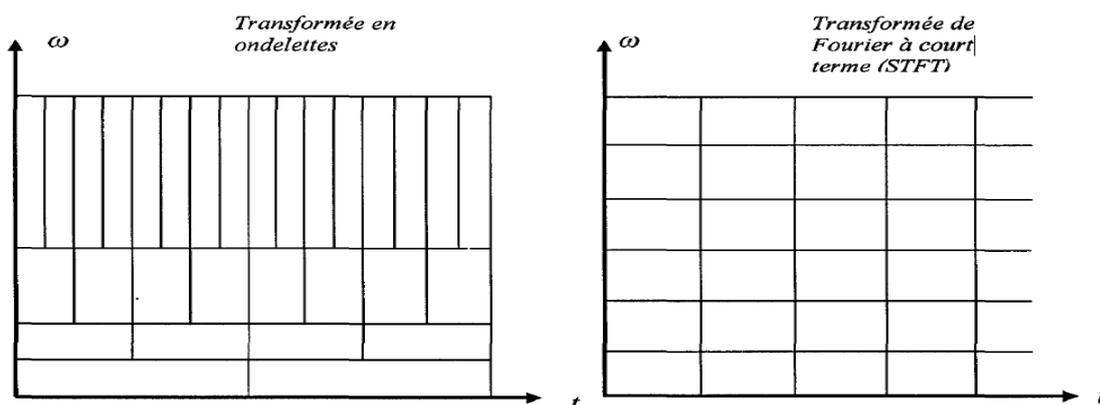


Figure 3.1. Représentation temps-échelle de la STFT et de la transformée en ondelette

Le **principe d'incertitude de Heisenberg** ( $\Delta t \cdot \Delta f \geq \frac{1}{4\pi}$ ) peut être appliqué à l'information de temps-fréquence d'un signal : on ne peut pas connaître la représentation exacte de temps-fréquence d'un signal, i.e., on ne peut pas connaître quelles composantes spectrales existent à quels instants du

temps. Ce qu'on peut savoir sont les intervalles de temps dans lesquels certaine bande des fréquences existent, qui est un problème de **résolution** [10].

Le problème avec le STFT a un rapport avec la largeur de la fonction de fenêtre qui est employée. Pour être techniquement correcte, cette largeur de la fonction de fenêtre est connue comme le support de la fenêtre. Si la fonction de fenêtre est étroite, alors elle est connue comme «**support compact**».

Fenêtre étroite →bonne résolution temporelle, mauvaise résolution fréquentielle

Fenêtre large →bonne résolution fréquentielle, mauvaise résolution temporelle

Bien que les problèmes de résolution de temps et de fréquence soient des résultats d'un phénomène physique (principe d'incertitude de Heisenberg) et existent indépendamment de la transformation utilisée, il est possible d'analyser n'importe quel signal en employant une approche alternative appelée l'**analyse multi résolution (Multi résolution Analyse (MRA))**.

La MRA est conçue pour donner une bonne résolution de temps et une faible résolution de fréquence aux hautes fréquences et une bonne résolution de fréquence et une faible résolution de temps aux basses fréquences.

Cette approche a un sens particulièrement quand le signal réel a des composantes à haute fréquence pour des courtes durées et des composants de basse fréquence pour de longues durées [10].

### 3.5. Transformée en ondelettes continue

Une ondelette est un outil mathématique utilisé pour le traitement du signal. Elle est considérée comme un opérateur linéaire, invariant par translation et par dilatation. L'analyse du signal se fait à l'aide d'une fonction appelée ondelette mère.

L'idée est de construire une représentation qui fait à la fois apparaître des informations temporelles et fréquentielles. Avec les ondelettes, le signal est découpé en différents morceaux qui sont des versions translatées et dilatées d'une même fonction (l'ondelette mère). Il en résulte une superposition d'ondelettes décalées et dilatées qui ne diffèrent entre elles que par leur taille. On obtient une transformée en ondelettes, fonction composée de deux variables qui sont le temps et la fréquence.

La transformation en ondelette d'une fonction  $f(t)$  est exprimée par l'équation suivante :

$$TOC_{a,b}(f(t)) = \frac{1}{\sqrt{a}} \int_{-\infty}^{+\infty} f(t) \psi^*\left(\frac{t-b}{a}\right) dt \quad \text{Eq (3.7)}$$

où  $\psi_{a,b}^*(t)$  est la fonction conjuguée de  $\psi_{a,b}(t)$ . Les coefficients  $TOC_{a,b}(t)$  sont nommés les coefficients d'ondelettes de la fonction  $f(t)$ . Tout signal est reconstruit à partir d'une superposition des ondelettes translatées et dilatées :

$$f(t) = \frac{2}{C_\psi} \int_{-\infty}^{+\infty} \left[ \int_{-\infty}^{+\infty} TOC_{a,b}(f) \psi_{a,b}(t) \frac{da}{a^2} \right] db \quad \text{Eq (3.8)}$$

où ;  $C_\psi$  est un nombre constant qui ne dépend que de l'ondelette choisie.  $C_\psi$  est défini par l'équation :

$$C_\psi = \int_{-\infty}^{+\infty} \frac{|\Psi(\omega)|^2}{|\omega|} d\omega < \infty \quad \text{Eq (3.9)}$$

où  $\Psi(\omega)$  est la transformée de Fourier de  $\psi(t)$ .

Ceci veut dire que la moyenne de  $\psi(t)$  est nulle. La transformée en ondelettes continue est inversible si l'ondelette mère vérifie la condition d'admissibilité donnée par l'équation (3.10)

Les conditions de régularité facultatives concernent les moments multiples nuls et sont exprimées par :

$$\int_{-\infty}^{+\infty} t^k \psi(t) dt = 0 \quad \text{Eq(3.10)}$$

L'équation ci-dessus exprime les propriétés de décroissance et de convergence vers 0 à l'infini de  $\psi(t)$  et de  $\Psi(\omega)$ .

Les familles d'ondelettes sont obtenues à partir de l'ondelette mère par la relation suivante :

$$\psi_{a,b}(t) = \frac{1}{\sqrt{a}} \psi\left(\frac{t-b}{a}\right) \quad \text{Eq (3.11)}$$

Dans cette expression,  $a$  est le facteur d'échelle et  $b$  le paramètre de translation. La variable «  $a$  » joue le rôle de l'inverse de la fréquence c-à-d plus  $a$  est petit moins l'ondelette (la fonction analysante) est étendue temporellement, donc plus la fréquence centrale de son spectre est élevée.

On notera que la norme est conservée lors du changement de facteur d'échelle, alors on a :

$$\begin{aligned} \|\psi_{a,b}\|^2 &= \int_{-\infty}^{+\infty} \frac{1}{a} \left| \psi\left(\frac{t-b}{a}\right) \right|^2 dt \\ &= \frac{1}{a} \int_{-\infty}^{+\infty} |\psi(x)|^2 a dx \end{aligned} \quad \text{Eq (3.12)}$$

On pourra noter que :

$$T^{ond} f(a, b) = \langle f, \psi_{a,b} \rangle \quad \text{Eq (3.13)}$$

La figure 3.2 montre les ondelettes obtenues par translation et dilatation de l'ondelette  $\psi(t)$ [6].

### 3.6. Transformée en ondelettes discrète (TOD)

Cette opération consiste à discrétiser la Transformée en Ondelette Continue pour obtenir les *séries d'ondelettes*. A l'heure actuelle, la plupart des calculs sont réalisés par les machines. Il est évident que ni la TO, ni le STFT, ni le CWT ne puissent être pratiquement calculés en employant des équations analytiques, des intégrales, etc... Il est donc nécessaire de discrétiser les transformées.

La discrétisation de l'échelle est  $a = a_0^j$  et celle de translation est  $b = k \cdot a_0^j \cdot b_0$ , où les paramètres entiers  $a_0 > 1$  et  $b_0 > 0$ . Remarquons que la discrétisation de la translation dépend de la discrétisation de l'échelle avec  $a_0$ . Dans ce cas, de l'équation (3.14), devient :

$$\psi_{j,k}(t) = a_0^{-j/2} \psi(a_0^{-j}t - kb_0) \quad \text{Eq (3.14)}$$

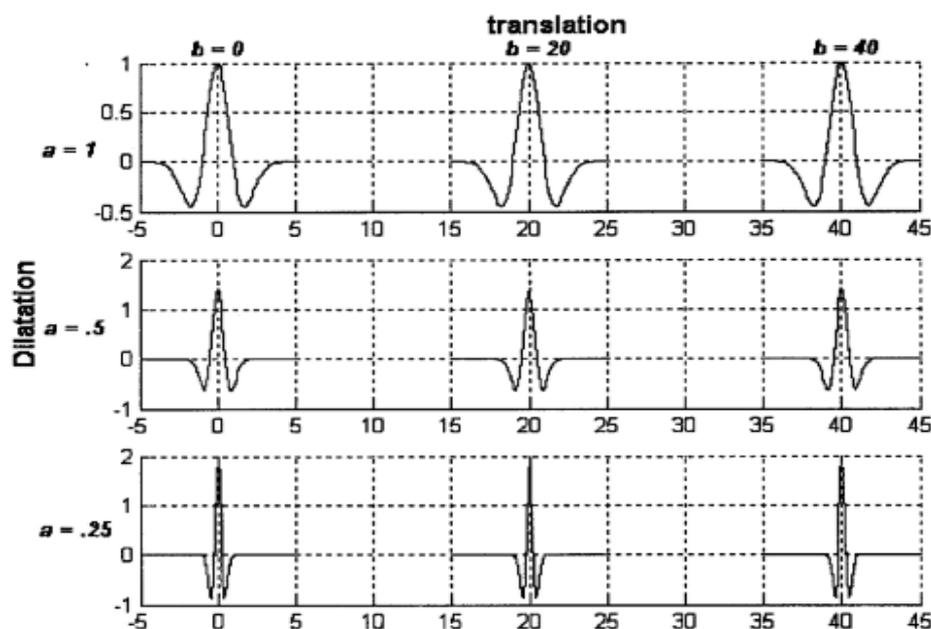


Figure 3.2. Translation et dilatation des ondelettes [6]

Alors toute fonction  $x(t)$  peut être développée dans les bases :

$$\text{TOD}(x(t)) = \sum_{j,k=-\infty}^{\infty} C_{j,k} \psi_{j,k}(t) \quad \text{Eq (3.15)}$$

Cette représentation est appelée la **série d'ondelette**.

Où  $j, k \in \mathbb{Z}$ . Alors, les coefficients des ondelettes  $C_{j,k}$  sont :

$$C_{j,k} = CWT_x^\psi(j, k) = \Psi_x^\psi(j, k) = \int_{-\infty}^{+\infty} x(t) \psi_{j,k}^*(t) dt \quad \text{Eq (3.16)}$$

Si on choisit  $a_0 = 2$  et  $b_0 = 1$ , on parle alors de transformée dyadique. Ce choix est plus commode (en termes de programmation).

Dans ce cas  $b = k2^j$  (dilation dyadique ou dilatation binaire) et  $a = 2^j$  (échelle dyadique ou échelle binaire), alors équation (3.17) devient :

$$\psi_{j,k}(t) = 2^{-j/2} \psi(2^{-j}t - kb_0) \quad \text{Eq (3.17)}$$

### 3.7. Propriétés des ondelettes

Les propriétés les plus importantes des ondelettes sont :

- **Support compact** : La plupart des ondelettes sont à support compact dans le domaine temporel, ce qui veut dire qu'elles sont à durée finie (filtres FIR) et se distinguent par leur atténuation rapide. Un support compact permet une complexité de calcul réduite, une meilleure résolution dans le domaine du temps mais donne une résolution pauvre en fréquence. Comme exemple, on peut citer les ondelettes de Daubechies, Symlets, Coiflets, etc. Par dualité, les ondelettes à bande étroite sont des ondelettes à support compact dans le domaine fréquentiel mais pas dans le domaine temporel (filtres IIR). Les ondelettes de Meyer en sont un exemple.
- **Symétrie** : Les ondelettes symétriques donnent naissance à des filtres à phase linéaire. Daubechies a montré que, pour qu'une ondelette soit symétrique, le filtre  $h$  doit être à phase complexe linéaire, et que le seul filtre miroir conjugué symétrique à support fini est le filtre de Haar qui correspond à une ondelette discontinue à un seul moment nul. Mis à part l'ondelette de Haar, il n'y a donc pas d'ondelette réelle orthogonale symétrique à support compact.
- **Moments nuls** : Par définition, une fonction  $\psi$  possède  $p$  moments nuls si et seulement si la formule  $\int_{-\infty}^{+\infty} t^k \psi(t) dt = 0$  pour  $0 \leq k \leq p$  est vérifiée [11]. L'intérêt de cette propriété est que si une ondelette a un nombre de moments nuls suffisant, on obtiendra alors plus de coefficients à petites valeurs.

- **La régularité :** La régularité d'une ondelette est la propriété permettant de localiser les singularités dans un signal. La régularité est une propriété importante pour obtenir des coefficients d'ondelettes les plus petits possibles afin de les annuler.
- Condition d'admissibilité :

$$C_\psi = \int_0^{+\infty} \frac{\|\Psi(w)\|^2}{w} dw < +\infty \quad \text{Eq (3.18)}$$

Où  $\Psi$  est la transformée de Fourier de  $\psi \Rightarrow \psi(t)$  de moyenne nulle.

- Conditions de régularité facultatives : moments multiples nuls

$$\int_{-\infty}^{+\infty} t^k \psi(t) dt = 0 \quad \text{Eq (3.19)}$$

$\Rightarrow$  Propriétés de décroissance et de convergence vers 0 à l'infini de  $\psi(t)$  et de  $\Psi(t)$ .

### 3.8. Familles d'ondelettes

#### 3.8.1 Exemples d'ondelettes continues

Il existe plusieurs types d'ondelettes qui diffèrent par leurs caractéristiques, expressions et applications. Nous allons présenter brièvement quelques unes d'entre elles. L'ondelette de Morlet et la fonction de chapeau mexicain sont deux candidats. Ces deux ondelettes sont illustrées dans la figure 3.3. L'ondelette de Haar, qui peut être considérée comme la plus simple, sera détaillée et employée pour l'analyse d'ondelette des exemples qui seront présentés plus tard en ce chapitre.

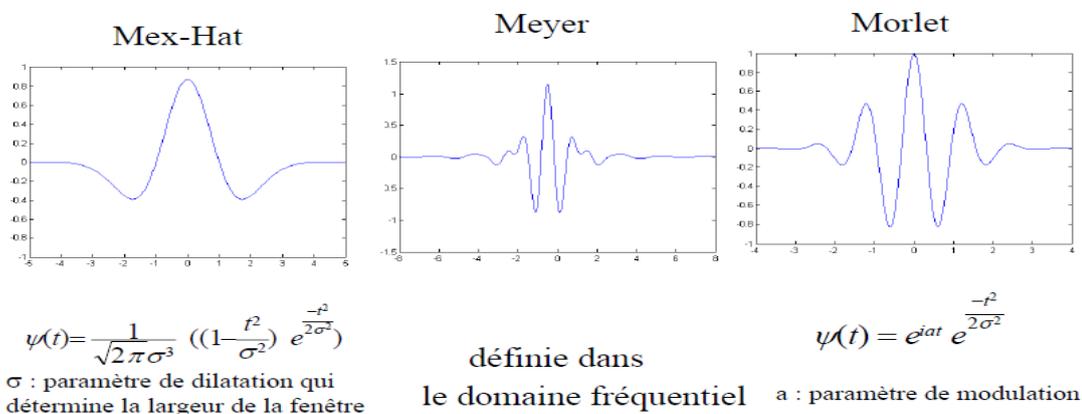
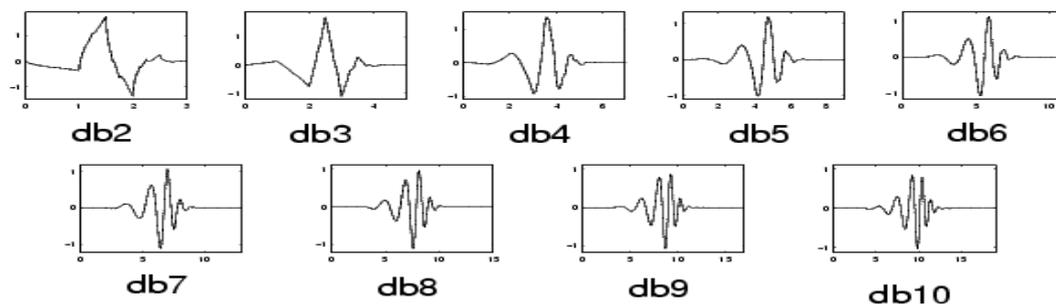


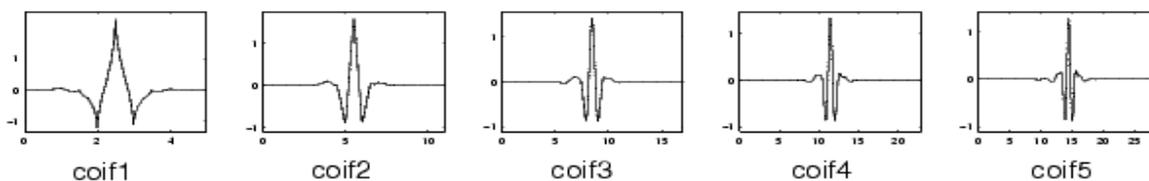
Figure 3.3 : Exemples d'ondelettes continues

### 3.8.2 Exemples d'ondelettes discrètes

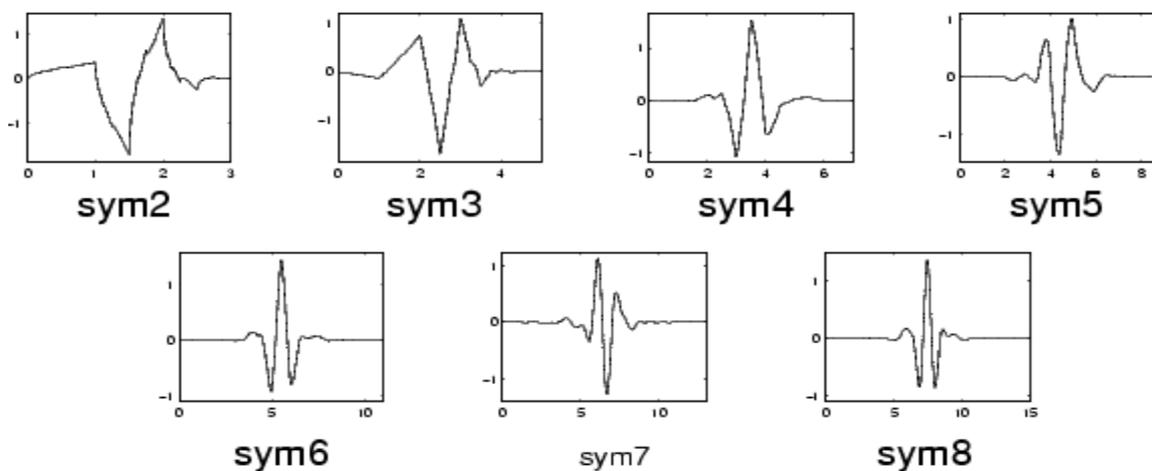
Les ondelettes de Daubechies (dbN), de Coiflets(coifN) et de symlets(symN), dont l'ordre est N, sont conçues par Ingrid Daubechies. Ces ondelettes sont orthogonales, biorthogonales et presque symétriques respectivement. Les propriétés de ces familles d'ondelettes sont semblables et différent seulement par le nombre de moments nuls. Sur la figure 3.4, sont tracées les familles de ces trois ondelettes.



(a)



(b)



(c)

Figure 3.4 : Familles des ondelettes discrètes (Daubechies ,Coiflets et symlet) [12]

### 3.9. Analyse d'un signal avec l'Ondelette de Haar

Nous prendrons comme point de départ de la décomposition les fonctions ondelette mère et échelle de Haar illustrée par les figures 3.6.b et 3.6.a respectivement. Le signal à traiter comme exemple est une fonction en escalier sur les intervalles  $[k, k + 1)$  ( $k \in \mathbb{N}$ ), soit  $f(t)$  représentée sur la figure 3.7.

Chaque coefficient de la fonction d'échelle est donné par l'intégrale du produit du signal par la fonction d'échelle. Dans notre exemple, elle donne les valeurs moyennes pour chacun des quatre intervalles  $[0,1]$   $[1,2]$   $[2,3]$   $[3,4]$ .

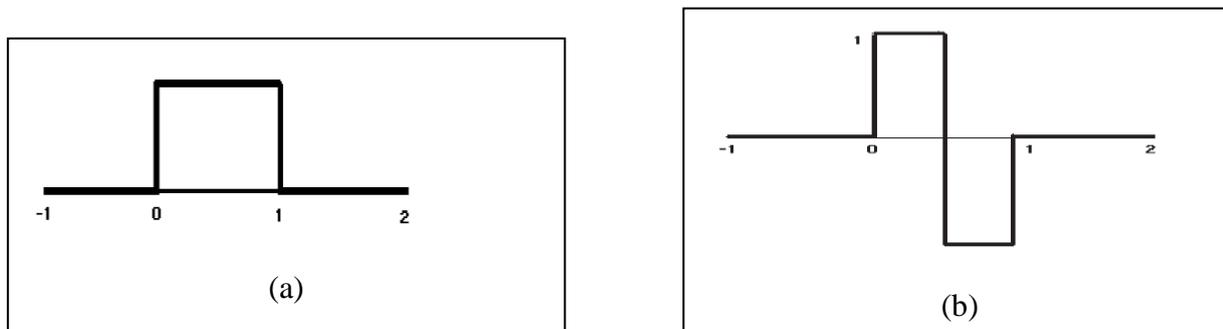


Figure 3.5 : La fonction d'échelle (a) et Ondelette (b) de Haar

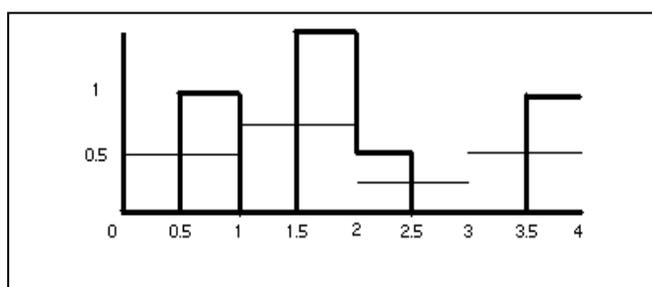


Figure 3.6 : Un signal (traits gras) et son approximation par la fonction d'échelle de Haar (traits fins)

Les coefficients des ondelettes contiennent donc l'information qu'on doit ajouter à celle encodée par la fonction d'échelle pour retrouver le signal à une résolution 2 fois plus fine comme le montre la figure 3.8.

**La fonction d'échelle** donne une suite d'images du signal, chacune à des résolutions qui diffèrent de la précédente par un facteur de deux. Quand on augmente la résolution, les images

successives approximent le signal de mieux en mieux, quand on la diminue, la quantité d'information contenue dans les images diminue, puis s'annule. Elle détermine donc la résolution la plus fine à laquelle le signal est étudié.

Les **fonctions ondelettes** encodent la différence d'information entre deux images successives, c'est à dire les détails qu'une image acquiert quand on augmente ou on diminue sa résolution. Un signal peut être composé de structures de tailles très différentes. Une analyse à plusieurs résolutions est donc particulièrement adéquate (voir figure 3.9).

La formalisation du traitement d'un signal avec l'ondelette de Haar se fait selon les étapes suivantes :

- **Étape 1 :**

Sur chaque intervalle  $[k, k + 2]$ , la fonction  $f(t)$  sera calculée par l'expression suivante :

$$f(t) = \begin{cases} a & \text{sur } [k, k + 1] \\ b & \text{sur } [k + 1, k + 2] \end{cases} \quad \text{Eq(3.22)}$$

L'approximation grossière du signal original est donnée par la moyenne  $(a + b)/2$  et elle est notée par **approx1(f)**.

Les détails du signal représentent la différence  $(b - a)/2$  et ils sont notés **details1(f)**.

<b>Image du signal</b>	<hr/>							
	0	1	0	1.5	0.5	0	0	1
<b>Coefficient d'échelle</b>	<hr/>							
	0.5		0.75		0.25			0.5
<b>Coefficient d'ondelette</b>	<hr/>							
	-0.5		-0.75		0.25			-0.5

Figure 3.7: Coefficients d'échelles et d'ondelettes d'une analyse multi-résolution

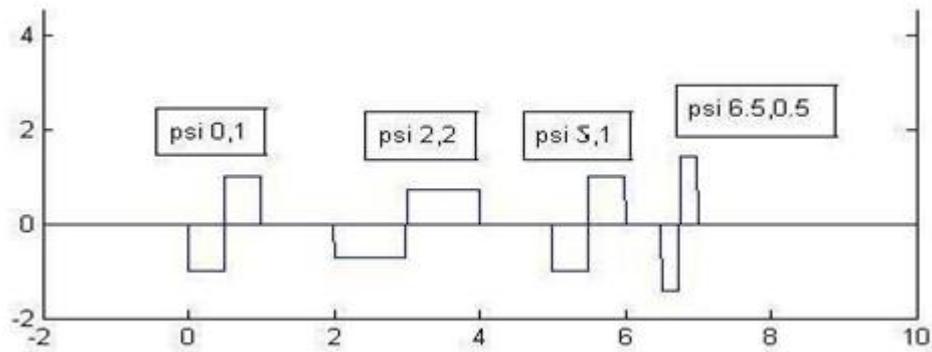


Figure 3.8 : famille d'ondelettes de Haar[13]

La reconstruction du signal  $f(t)$  se fait par la somme des coefficients d'approximation et de détails .Pour un seul niveau de décomposition on aura :

$$f(t) = approx1(f) + details1(f) \quad Eq(3.23)$$

Et

$$\frac{a+b}{2} - \frac{-a+b}{2} = a \quad \frac{a+b}{2} + \frac{-a+b}{2} = b \quad Eq (3.24)$$

L'ondelette mère de Haar est la base du processus de différence qui est décrit par

$(-1.a + 1.b)/2$ .

$$\psi(x) = \begin{cases} -1 & \text{si } 0 \leq x < 1/2 \\ 1 & \text{si } 1/2 \leq x < 1 \\ 0 & \text{sinon} \end{cases} \quad Eq (3.25)$$

Le processus de moyenne est lié au processus de différence c-à .d la fonction d'échelle  $\phi$  est liée à l'ondelette mère  $\psi$ .

$$\phi(x) = \begin{cases} 1 & \text{si } 0 \leq x < 1/2 \\ 1 & \text{si } 1/2 \leq x < 1 \\ 0 & \text{sinon} \end{cases} \quad Eq (3.26)$$

Cela correspond au processus de moyenne  $(1.a + 1.b)/2$ .

• **Etape 2 :**

Nous recommençons le processus de l'étape 1 sur l'**approx1(f)** comme suit :

- Calcul de moyenne et différence sur chaque intervalle de type  $[j \cdot 2^j, (j + 1) \cdot 2^j]$  ce qui donne :  $\text{approx2}(f)$  et  $\text{details2}(f)$ .
- $f(t) = \text{approx2}(f) + \text{details2}(f) + \text{details1}(f)$
- Si nous stoppons le processus de décomposition à cette étape,  $2^2$  est l'échelle grossière. C'est l'échelle de la dernière approximation.

La figure 3.10 ci-dessous montre l'analyse multi-résolution d'un signal avec trois niveaux de décomposition. L'approximation du dernier niveau contient l'information pertinente (les basses fréquences du signal) alors que les détails ressemblent à du bruit (les hautes fréquences du signal)

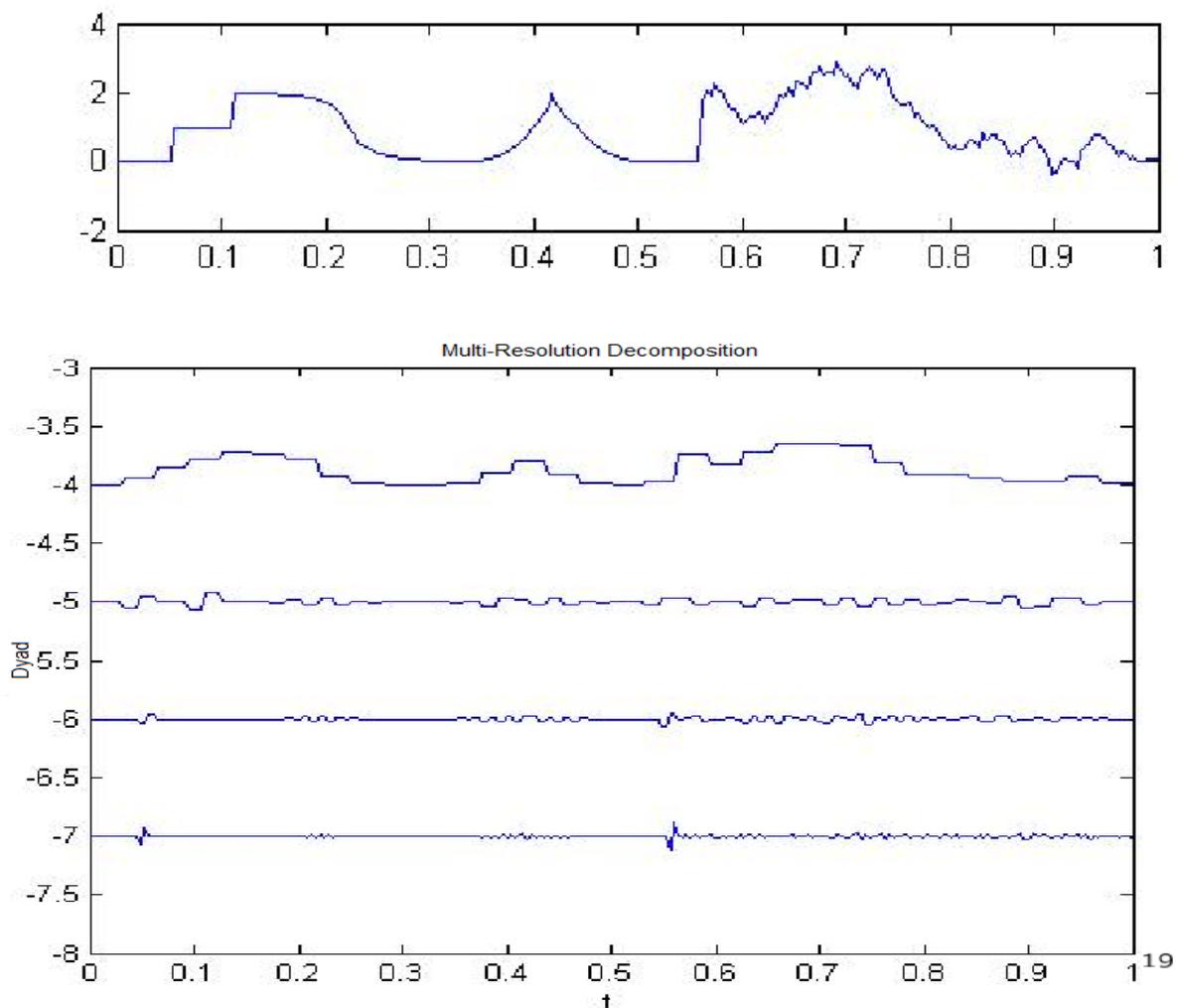


Figure 3.9 : Analyse multi-résolution de Haar d'un signal [13].

Les coefficients d'ondelettes sont le résultat d'un produit scalaire entre le signal et les différentes ondelettes  $\psi_{u,s}$  de la famille d'ondelettes :

$$\text{approx}1(j) = \int \phi_{1,j}(x)f(x)dx \text{ et } \text{details} 1(j) = \int \psi_{1,j}(x)f(x)dx \quad \text{Eq (3.27)}$$

La mesure de la similarité entre le signal et les ondelettes de la famille choisie donne une bonne information temps-fréquence

### 3.10. Quelques applications d'ondelettes

Les ondelettes sont un outil statistique puissant qui peut être utilisé pour un large éventail d'applications, à savoir :

- Traitement du signal de parole (compression, débruitage,.....)
- Traitement d'images (lissage, suppression du bruit et compression des images .
- Vérification de l'empreinte digitale.

Par exemple, le principe de la compression d'un MP3 repose sur la décomposition en ondelettes du signal qui nous permet de séparer les basses fréquences et les hautes fréquences. Les fréquences inaudibles pour l'homme, sont détruites et le signal est reconstruit à partir des coefficients restants. On obtient une musique compressée qui produit le même son pour l'homme.

Le débruitage d'un signal 1D consiste à enlever les parasites d'une conversation bruitée. Les ondelettes fournissent une méthode assez simple pour extraire les signaux du bruit blanc (bruit existant à toutes les fréquences). La difficulté du débruitage réside dans le fait que le bruit peut aussi bien être de haute fréquence (analyses médicales), que de basse fréquence (bruit produit par un sous-marin), ou encore des deux natures à la fois (images bruitées).

Pour débruiter un signal, on le transforme en ondelettes, on élimine, à toutes les résolutions, les coefficients supérieurs à un certain seuil. Ensuite, on reconstruit le signal avec les coefficients restants.

### 3.11. Critères d'évaluation de qualité de la compression

On trouve parmi les critères d'évaluation, ce qui :

### 3.11.1. L'erreur quadratique moyenne :

Elle est définie par l'équation :

$$EQM(S, \hat{S}) = \frac{1}{N_d} \|S - \hat{S}\|^2 \quad \text{Eq(3.28)}$$

S : Signal original (initial).

$\hat{S}$  : Signal débruité.

$N_d$  : Dimension du signal

### 3.11.2. le rapport signal sur bruit

Ce critère est défini par :

$$SNR = 10 \log_{10} \frac{\sum_{n=0}^{N-1} x^2(n)}{\sum_{n=0}^{N-1} (x(n) - x_r(n))^2} \quad \text{Eq(3.29)}$$

Ou, N: taille de la trame.

$x(n)$  : Parole originale ,  $x_r(n)$  : parole reconstruite.

### 3.11.3. Le MOS (Mean Objectif Score )

Ce critère se fait expérimentalement par des experts qui donnent des notes comme illustré par le tableau

Note	Qualité de la parole	Niveau de dégradation
5	Excellente	imperceptible
4	Bonne	perceptible mais non gênant
3	Moyenne	un peu ennuyeux
2	Mauvaise	Ennuyeux
1	Très mauvaise	très ennuyant

Tableau.3.1: L'échelle MOS

### 3.12. Conclusion

La transformée en ondelettes est devenue en quelques années un champ de recherche très actif. On ne compte plus aujourd'hui les applications qui utilisent cette technique. Il s'agit d'un algorithme permettant de calculer une représentation d'un signal en bandes de fréquences indépendantes.

Dans ce chapitre nous avons discuté l'handicape que présentait la TF pour analyser un signal non stationnaire. Ainsi, nous avons présenté une transformée qui résout ce problème en donnant en même temps une bonne localisation temporelle et fréquentielle.

La décomposition d'un signal par l'Ondelette de Haar a été développée dont le but de l'utiliser pour débruiter un signal de parole.

Le signal à débruiter, sera analysé par multi résolution de Haar, les coefficients supérieurs à un certain seuil seront éliminés, en les considérant comme bruit. Le signal est reconstruit avec les coefficients restants. Nous verrons cette procédure dans le chapitre suivant.

### 4.1. Introduction

Dans ce chapitre nous présenterons les résultats expérimentaux de notre travail, en référence avec le chapitre précédent.

Premièrement, nous allons présenter dans un court paragraphe l'environnement matériel et logiciel utilisés. Ainsi, la transformation de ondelette de Haar est appliquée à l'aide de la commande « wavemenu », à un signal de parole afin de débruitage.

Ensuite, nous allons réaliser des différentes expériences, en appliquant la décomposition en ondelette de Haar sur un signal de parole afin de voir la performance des ondelettes de Haar.

Enfin, nous présentons les résultats de débruitage obtenus. Nous évoluons la qualité de débruitage par EQM et SNR.

### 4.2. Environnement matériel et logiciel :

La simulation a été effectuée dans un environnement MATLAB™ version 7.10. Plusieurs fonctions y sont implantées, et la manipulation graphique se distingue par sa facilité et sa portabilité. Les graphes des résultats sont exportés directement vers les documents Word. Les formes d'ondelettes sont faciles à obtenir de même que les coefficients des filtres associés.

### 4.3. Prise en main de la boîte à outils :

Le logiciel Matlab contient une boîte spécifique aux ondelettes. La boîte à outils ondelettes se décompose en plusieurs familles de fonctions : familles d'ondelettes, ondelettes continues, ondelettes discrètes, ondelettes de Haar. Nous avons profité de cette boîte à outil pour simuler le débruitage du signal de la parole, comme suit :

### 4.3.1. La commande « wavemenu »

La commande « wavemenu » ouvre l'interface graphique de prise en main de la boîte à outils. Cette interface sert à utiliser de façon visuelle les différentes fonctions proposées sur des signaux (1D) ou des images (2D) sans connaître les commandes Matlab nécessaires à la création d'un programme. Cette boîte nous a permis de s'entraîner à analyser des signaux avec les ondelettes et la transformée en ondelette.

### 4.3.2. Analyse d'un signal 1D

Nous avons commencé par analyser un signal de parole «news : » à l'aide du menu « wavelet 1-D » qui offre plusieurs possibilités de traitement tel que : le débruitage, la compression, la visualisation (des audiogrammes, les coefficients de décomposition). Toutes ces options et autres apparaissent sur les fenêtres de dialogue ci-dessous.

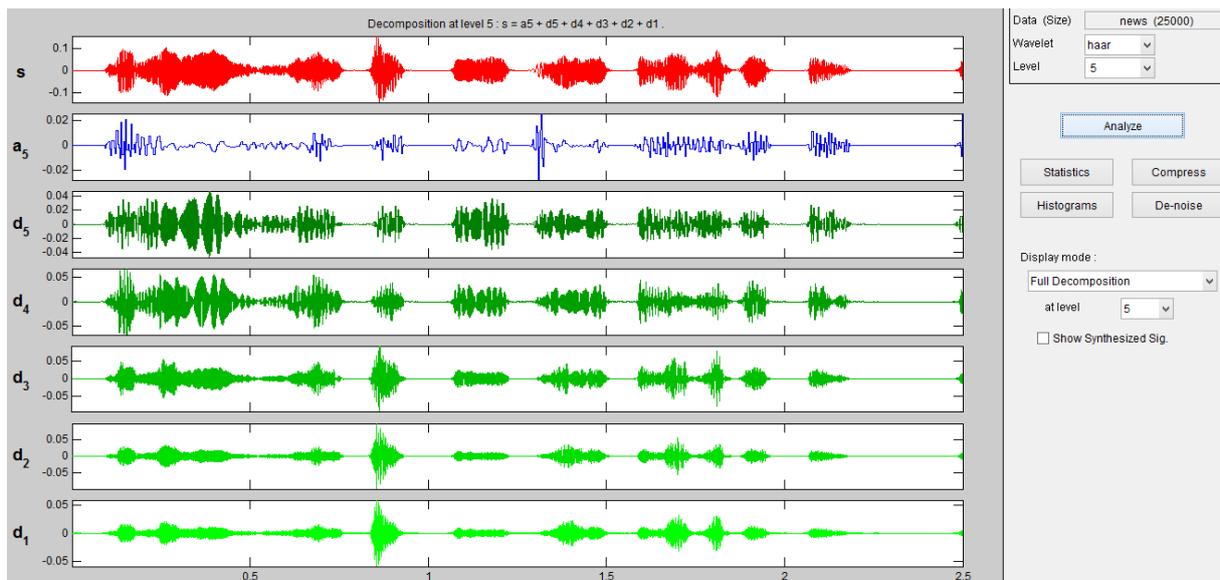


Figure .4.1:Analyse d'un signal de parole avec l'ondelette Haar de profondeur 5

La Figure.4.1 montre l'évolution temporelle des coefficients d'approximation et de détail de la décomposition d'un signal avec l'ondelette Haar de profondeur 5.

Sur (la figure.4.2), apparaissent toutes les informations associées à la décomposition d'un signal de parole « un son français féminin ». L'allure temporelle du signal original est donnée dans le coin supérieur droit (signal en rouge) .

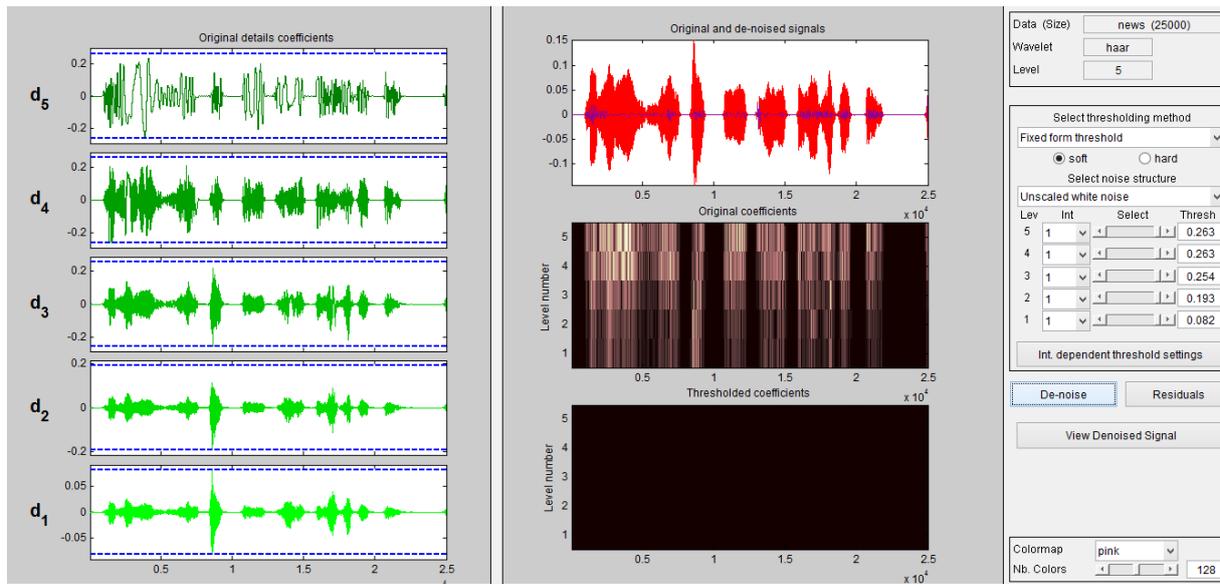


Figure.4.2: Débruitage d'un signal de parole par Ondelette de Haar de profondeur 5

### 4.3.3. visualisation des signaux

En cliquant sur «view denoised signal» une fenêtre de dialogue s'ouvre et nous permet de voir le signal débruité comme représenté sur (la figure.4.3)

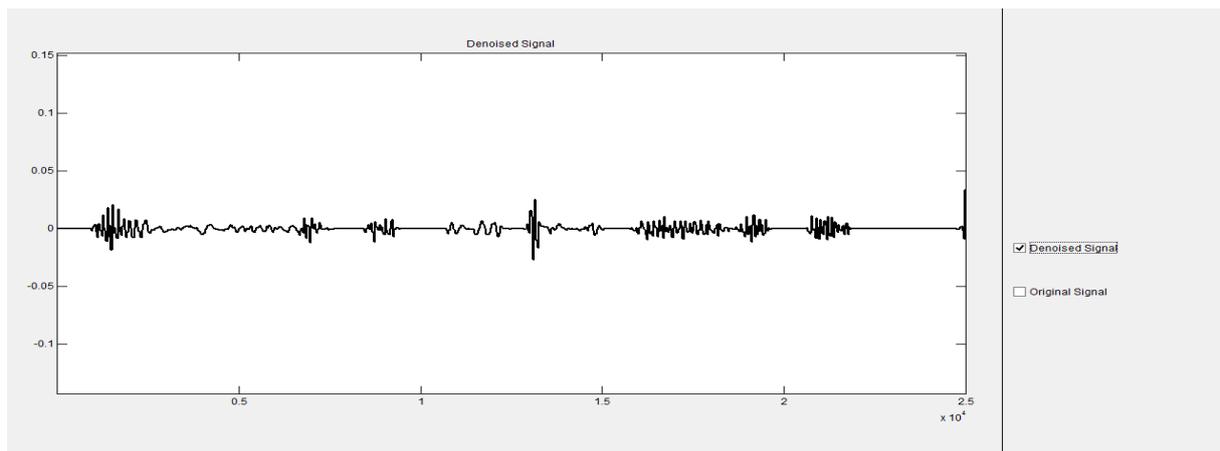


Figure.4.3 : Représentation d'un signal débruité

En cliquant sur « original signal » dans la fenêtre de dialogue précédente, nous aurons le signal original et débruité superposés comme le montre (la Figure.4.4) .

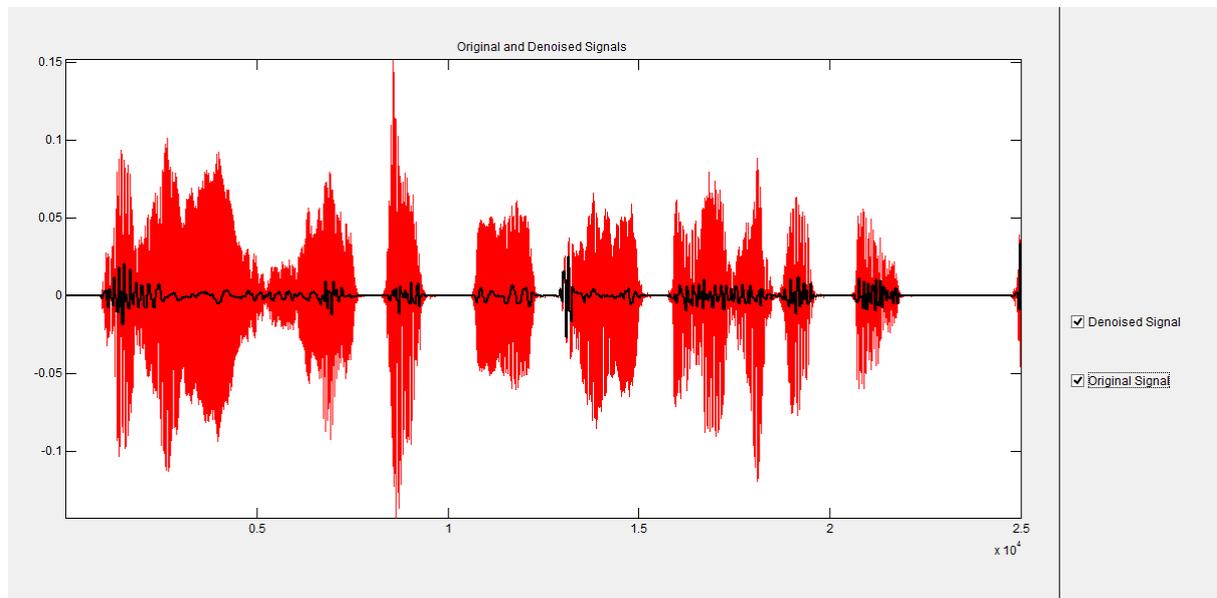


Figure.4.4 : Audiogramme du signal original avec bruit

### 4.4. Génération de la parole bruitée avec divers niveaux de RSB

La première étape consiste à ajouter un bruit de différents niveaux, à un signal de parole original de test (démuni de bruit).

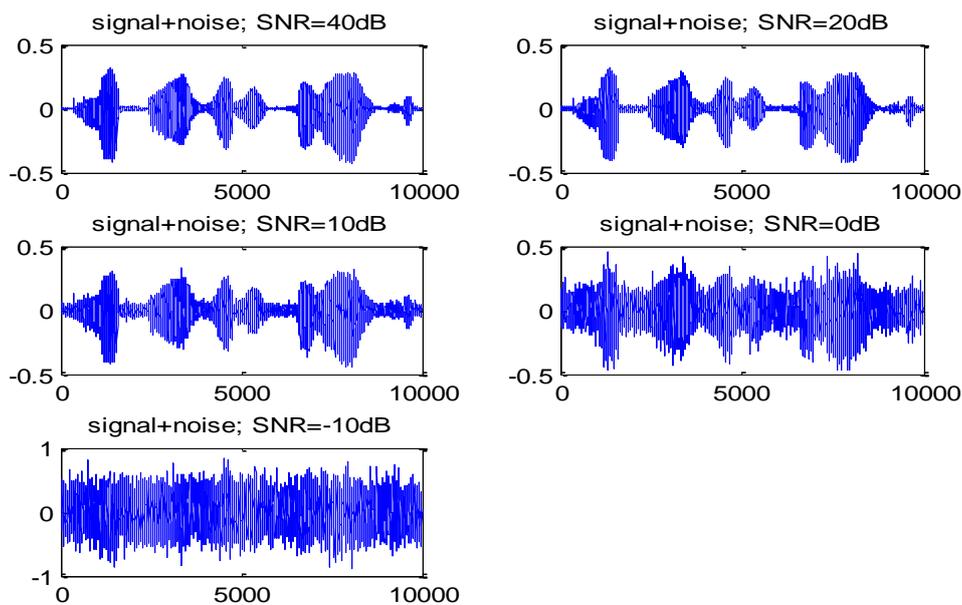


Figure.4.5: signal de parole bruité avec différents SNR

## 4.5. Décomposition d'un signal de parole par ondelette de Haar

### 4.5.1. Analyse d'un signal ayant un SNR=40dB

Dans cette première expérience nous allons analyser une partie d'un signal vocal qui est « la bise et le soleil » donnée par l'audiogramme figure 4.6, dont nous avons ajouté un bruit donnant un SNR de 40 dB. L'ondelette utilisée est l'ondelette de Haar de profondeur 3.

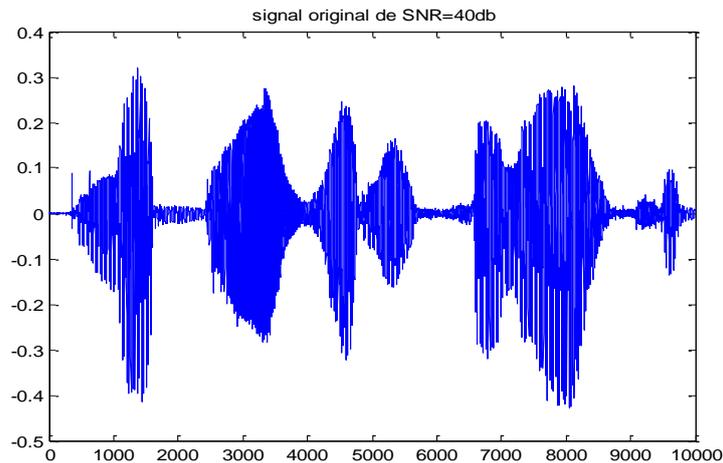


Figure.4.6 : Signal de parole bruité avec SNR =40dB

Cette analyse a donnée les coefficients de détail d'approximation donnée par la (figure 4.7)

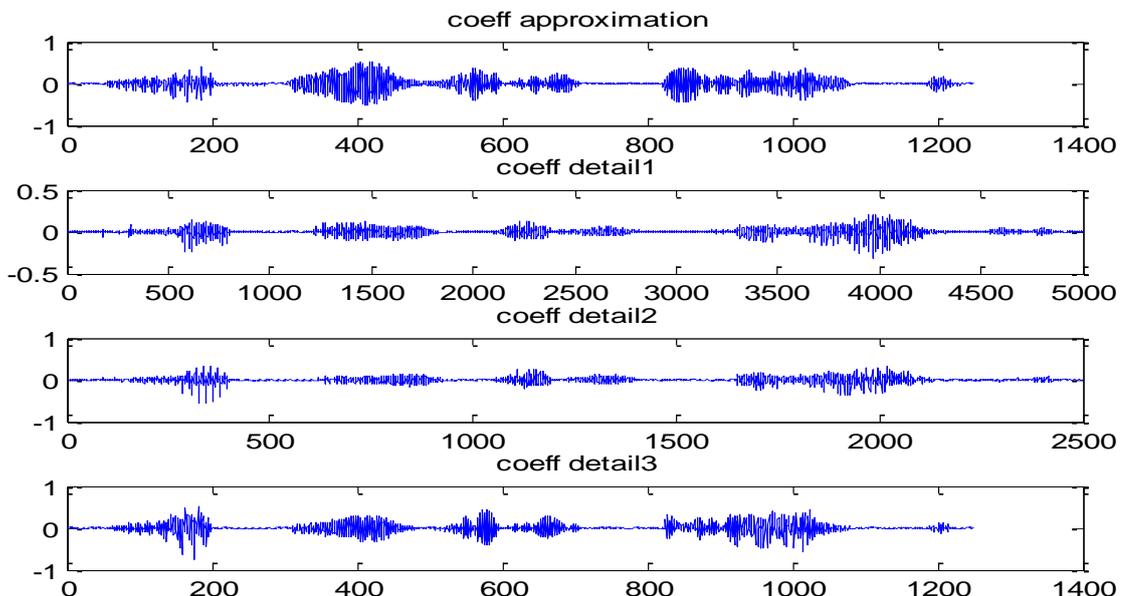


Figure 4.7: Les coefficients de détail et d'approximation de la décomposition par ondelette de Haar niveau 3 du signal

### 4.5. 2. Reconstruction du signal traité

La reconstruction du signal analysé est réalisée par l'ondelette de Haar inverse. Cette procédure utilise l'approximation du dernier niveau et tous les détails. Le signal résultant est donné par la figure 4.8.

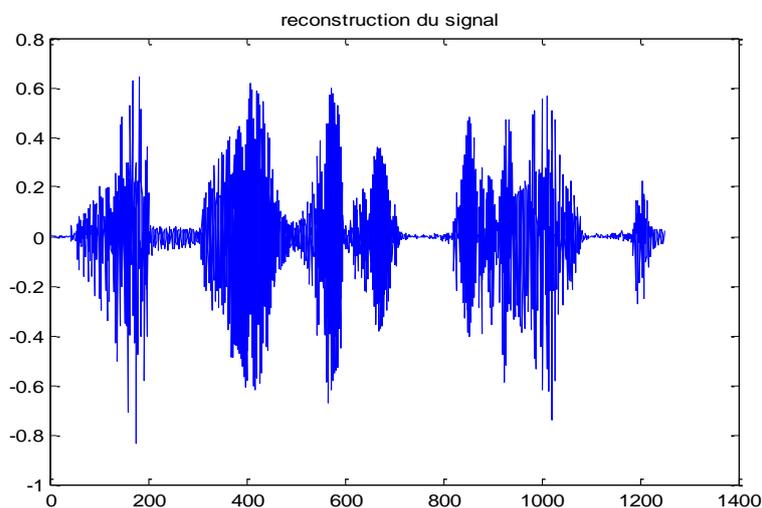


Figure.4.8: Signal de parole bruité reconstruit.

Nous remarquons que jusqu'à présent le signal obtenu n'est pas débruité et nous allons faire ceci par la suite.

### 4.6. Le débruitage d'un signal par les ondelettes

#### 4.6.1. Débruitage d'un signal sinusoïdal par ondelette de Haar

Dans cette partie nous avons simulé l'analyse et la reconstruction d'un signal sinusoïdal pour maîtriser la technique de débruitage par ondelette de Haar.

Pour cela, nous avons utilisé un signal sinusoïdal d'amplitude =1 et de fréquence d'échantillonnage  $f_e=5000$  HZ et de taille=1024 comme visualisé sur la figure 4.9.

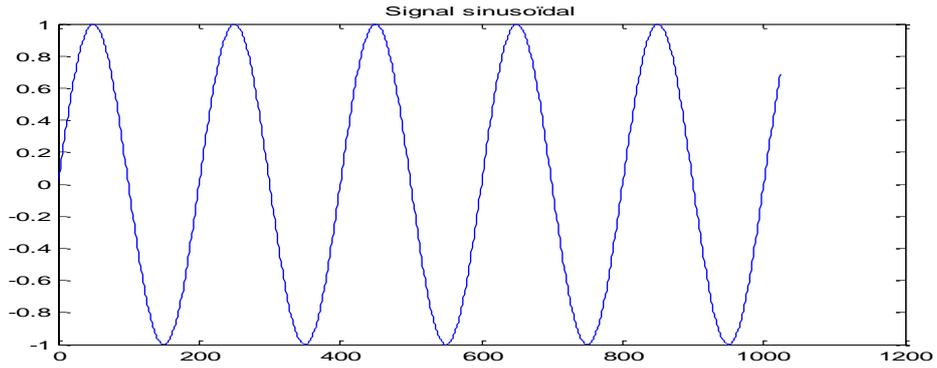


Figure.4.9:Signal sinusoïdal original utilisé

On ajoute au signal précédent un bruit aléatoire de moyenne nulle et de variance égale à l'unité en utilisant la commande *rand* de Matlab. La version bruitée du signal est illustrée par la figure 4.11.

Les Signaux des tous les détails sont représentés sur la figure.4.10, comme suit :

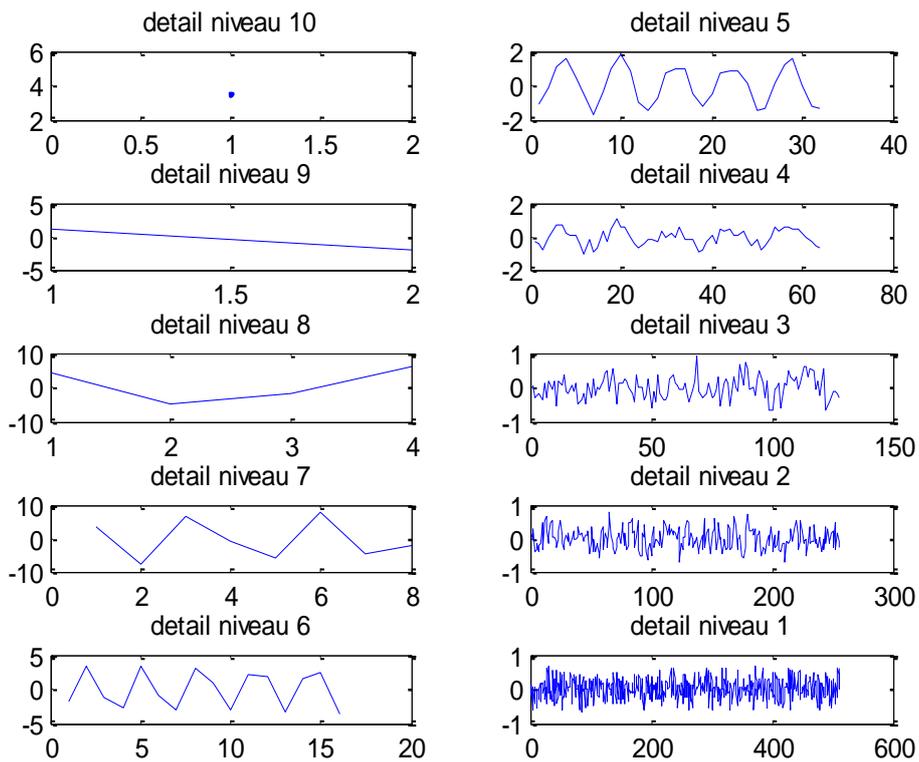


Figure.4.10: Représentation de tous les détails du signal analysé

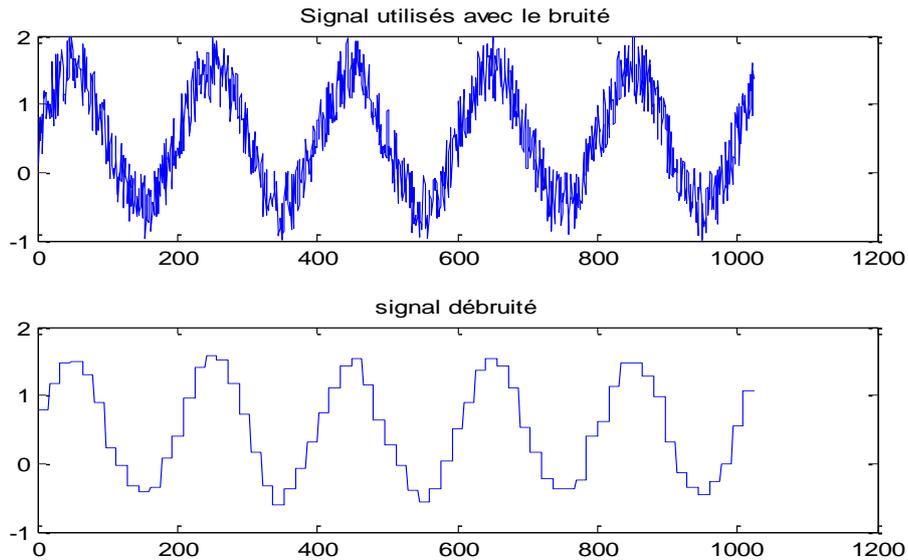


Figure.4.11: Reconstruction du signal

Pour reconstruire le signal et le débruité, nous allons modifier les coefficients de détails qui semblent être un bruit (composantes hautes fréquences du signal analysé), tout en conservant l'approximation, avant la synthèse. Après avoir débruité le signal sinusoïdal, les résultats obtenus sont donnés par la figure. 4.11.

### 4.6.2. Débruitage d'un signal de parole par ondelette de Haar

#### 4.6.2.1. Signal de parole bruitée avec $\text{snr} = 40\text{dB}$

Pour cette expérience nous avons ajouté un bruit tel que le SNR est égal à **40dB**. Ce signal est une phrase en langue française « la bise et le soleil ». La décomposition avec l'ondelette de Haar a donné les coefficients de détails et d'approximation illustrés par les figures 4.12 et 4.13 respectivement.

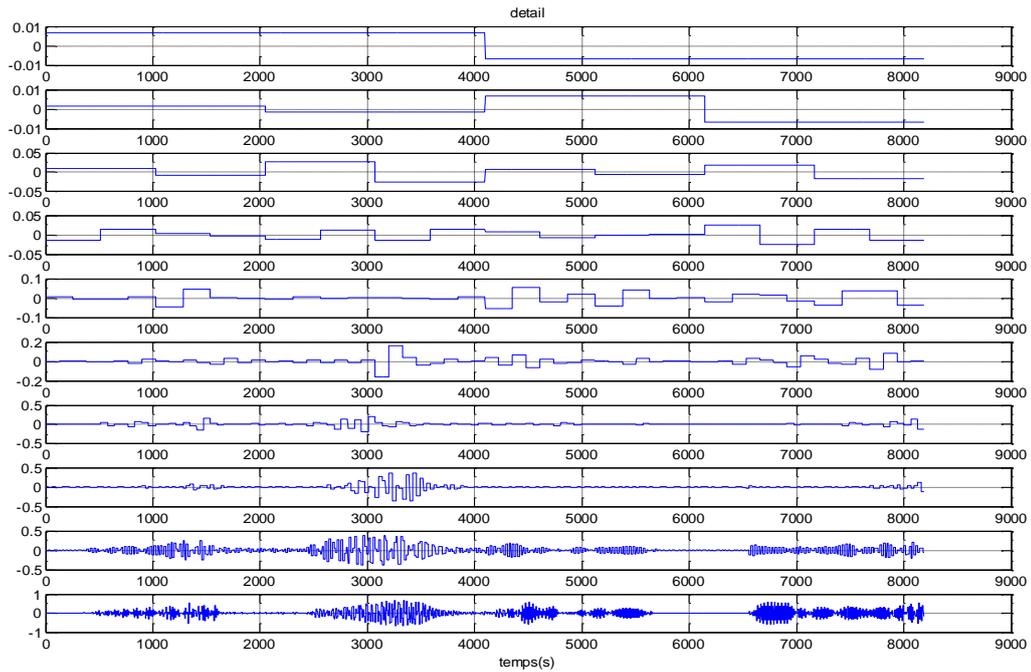


Figure.4.12 : coefficients de détails du signal bruité avec 40 dB.

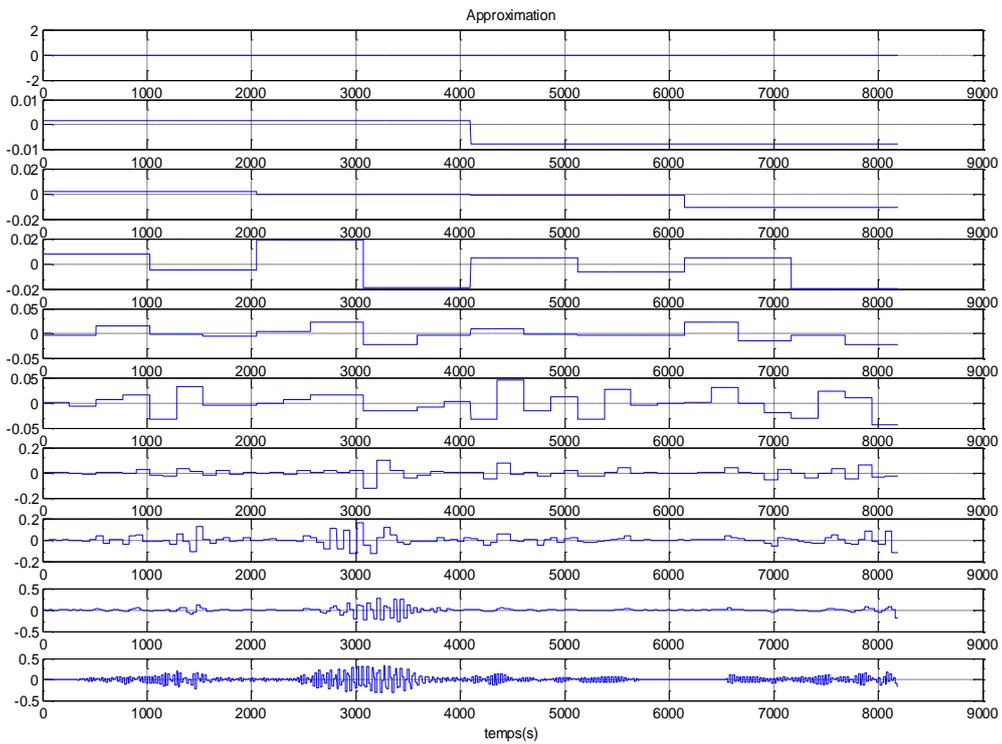


Figure.4.13 : coefficients d'approximation du signal bruité avec 40 dB

Nous remarquons que les détails contiennent les hautes fréquences du signal et qui engendrent le bruit à éliminer. Pour débruiter un signal par les ondelettes, il faut agir sur les coefficients de détails.

L'opération de débruitage a été faite par seuillage des coefficients de détails supprimés (d8 et d9). Un seuil de 0.0015, nous a permis d'obtenir un  $SNR_d = 42.3795$  et une  $eqm = 0.0215$ . En écoutant le signal débruité, ne n'avons perçu aucune différence avec le signal original. Le signal de référence (original), sa version bruitée et débruitée sont donnés sur la figure 4.14 ci-dessous :

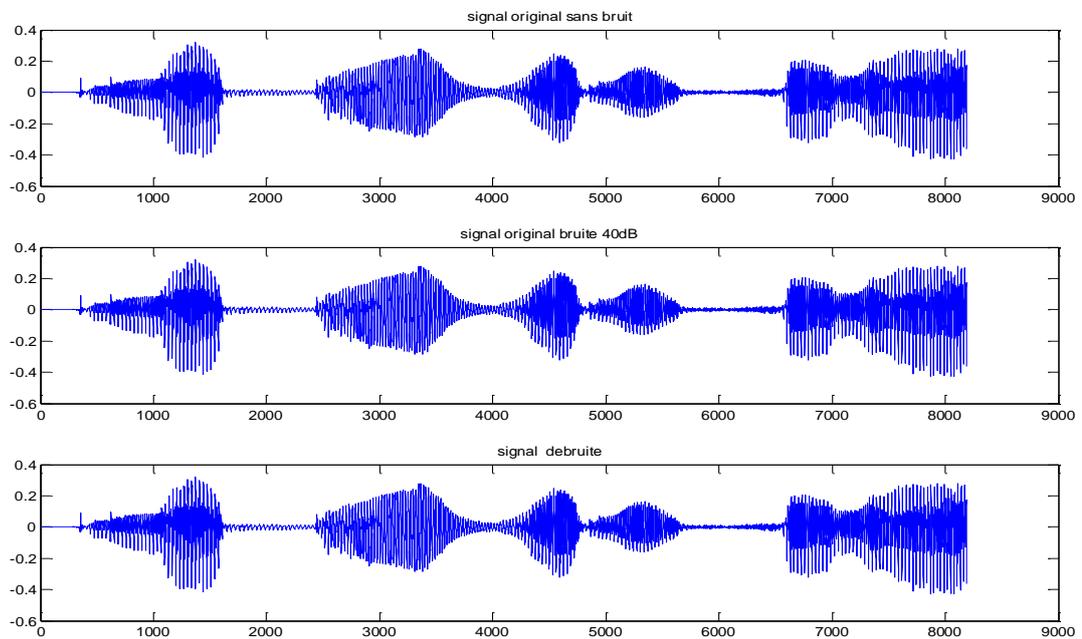


Figure.4.14 : Signal original, sa version bruitée avec 40 dB et le signal débruité

#### 4.6.2.2. Signal de parole avec $SNR = 20$ dB

Pour ce signal, un seuil de  $thr=0.003$  appliqué aux coefficients de détails, nous a permis de récupérer 10 dB. Le SNR trouvé est de  $29.8882$  dB et l' $eqm$  est égale à  $0.0452$ .

La décomposition avec l'ondelette de Haar a donné les coefficients de détails et d'approximation illustrés par les figures 4.15 et 4.16 respectivement.

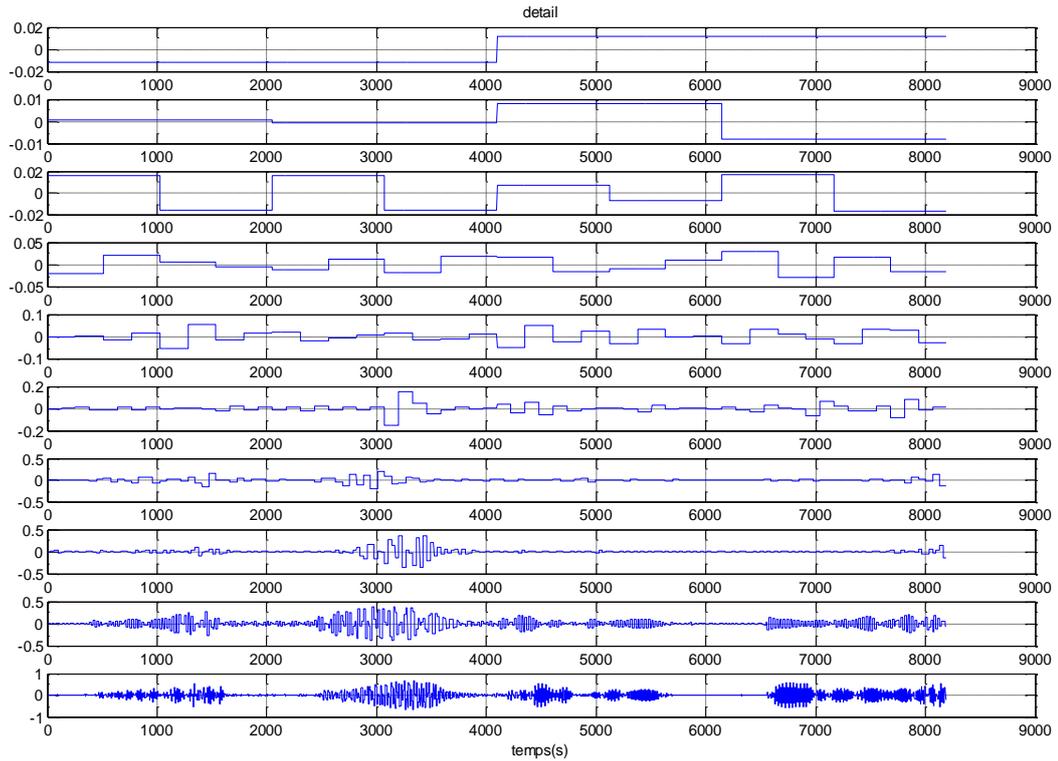


Figure .4.15 : coefficients de détails du signal bruité avec 20 dB.

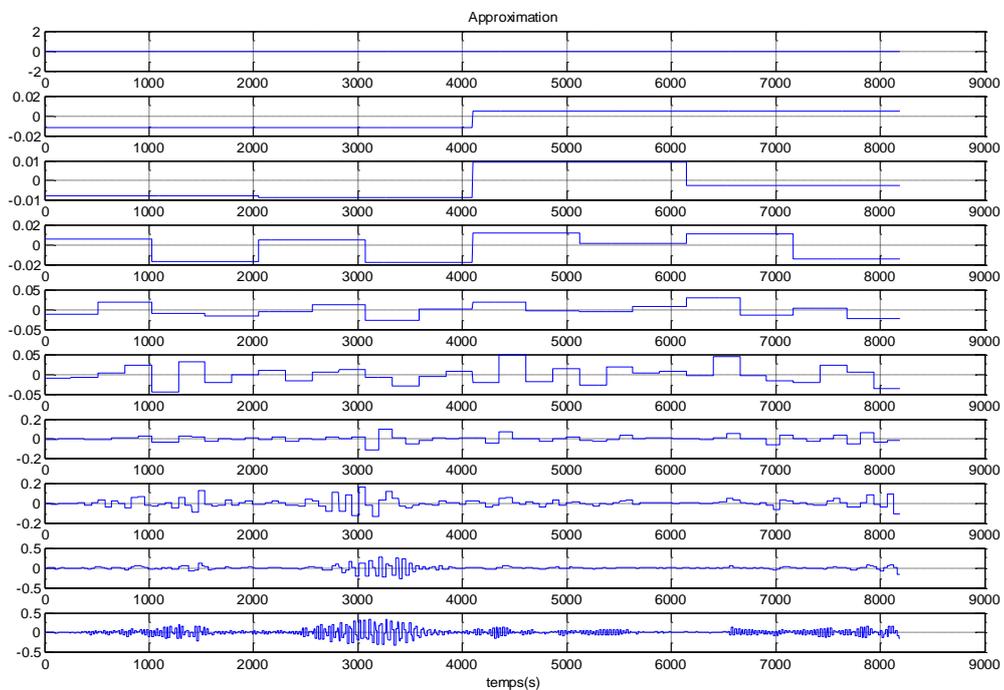


Figure.4.16 : coefficients d'approximation du signal bruité avec 20 dB

La procedure de débruitage est la même que précédemment, C-à-d nous éliminons les détails considerés comme du bruit. le signal original, sa version bruitée et le signal debruite sont representés sur la figure 4.17.

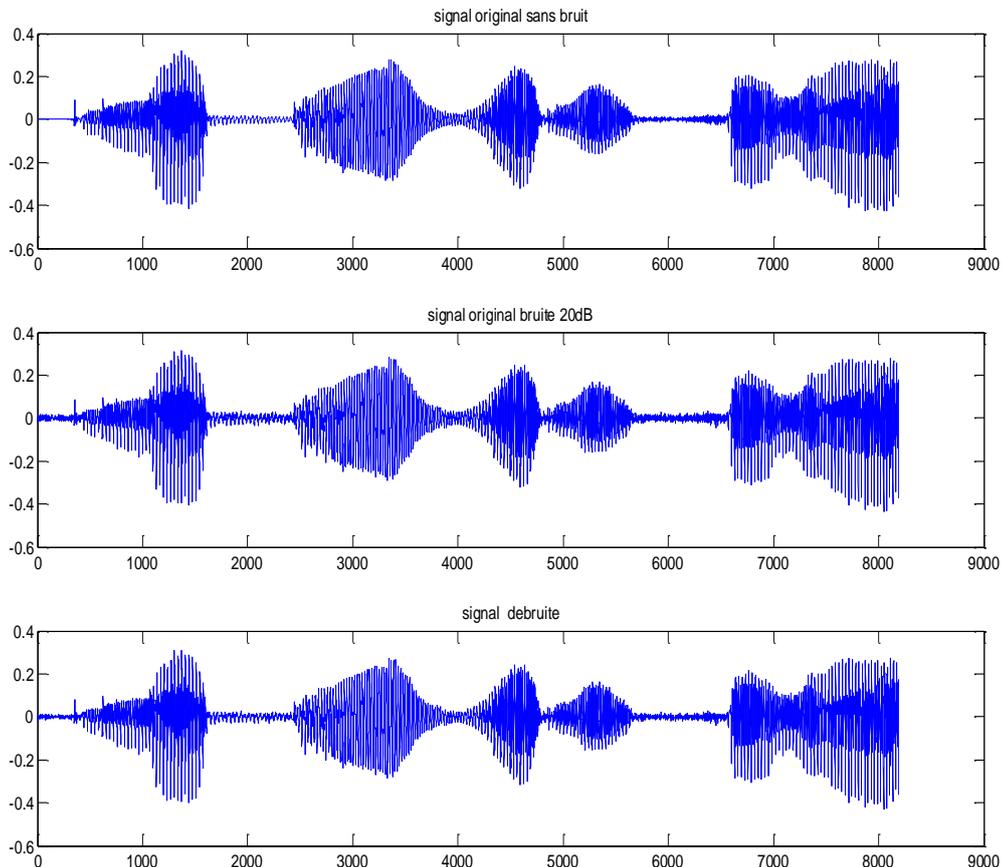


Figure.4.17 : Signal original, sa version bruitée avec 20 dB et le signal debruite

#### 4.6.2.3. Signal de parole SNR = 10db

Dans ce cas on a trouvé un snrd = 18.4091 dB après un débruitage par seuillage des coefficients de détails. Le seuil qui a donné ce SNR est thr=0.0118 et l'eqm est égale à eqm = 0.0873 .

La décomposition avec l'ondelette de Haar a donné les coefficients de détails et d'approximation illustrés par les figures 4.18 et 4.19 respectivement

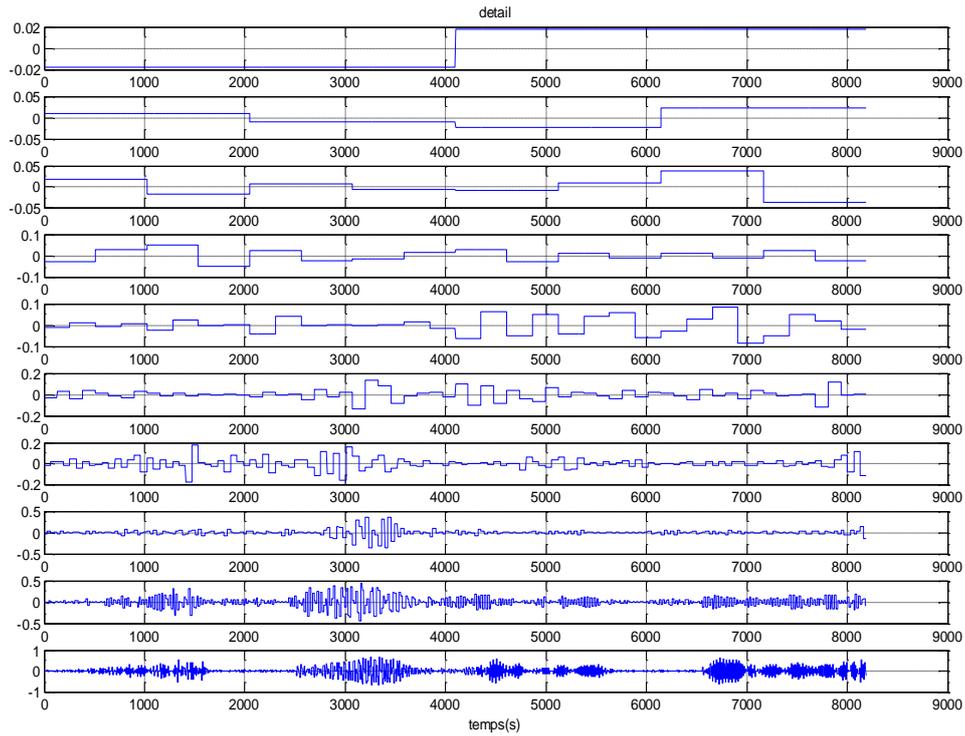


Figure.4.18 : coefficients de détails du signal bruité avec 10 dB.

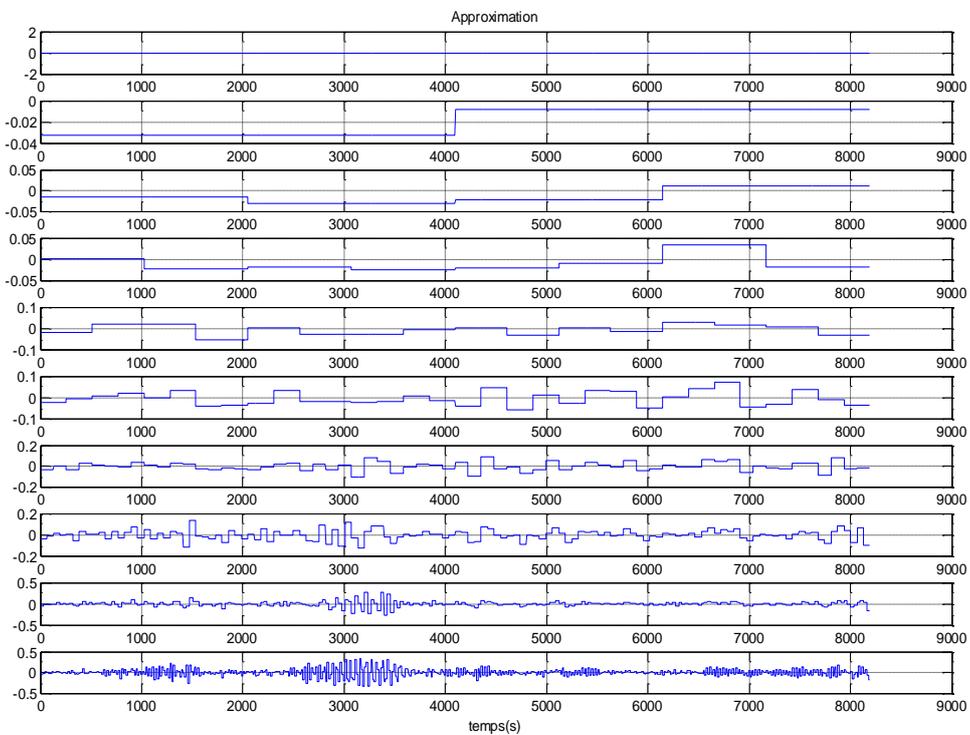


Figure.4.19 : coefficients d'approximation du signal bruité avec 10 dB

Pour cette simulation les résultats trouvés sont comme suit:

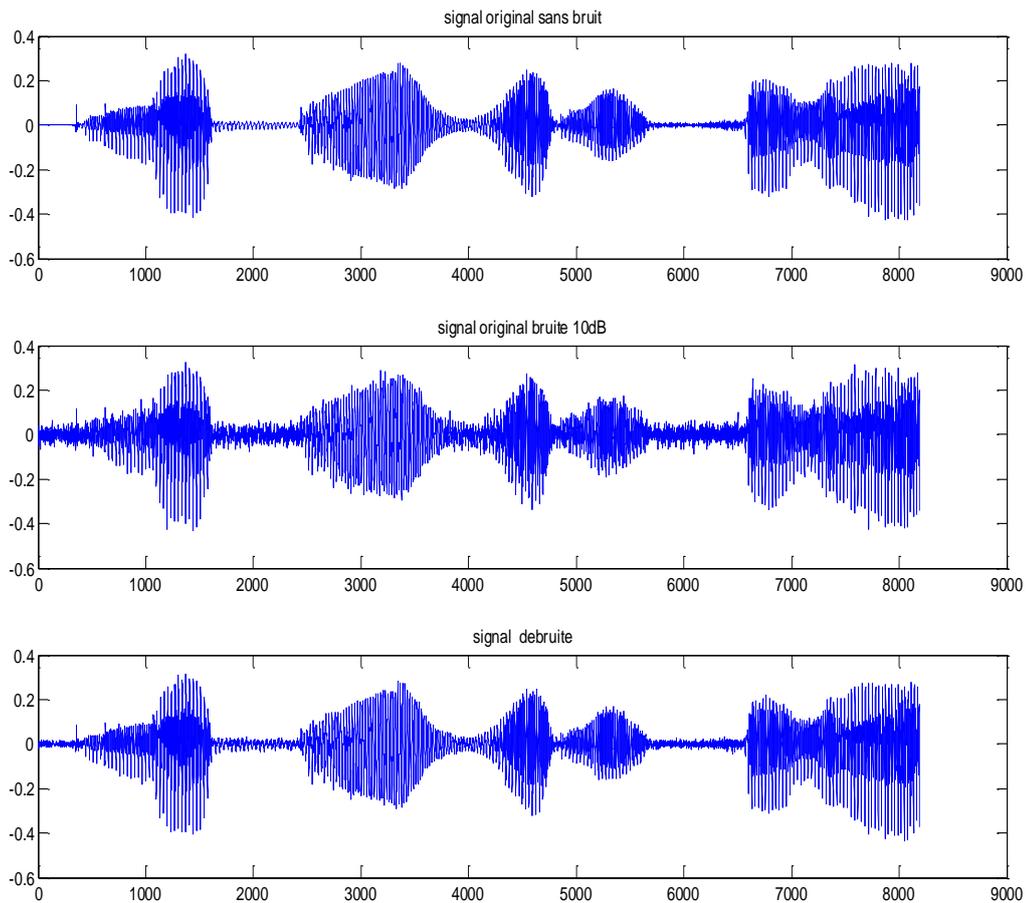


Figure.4.20 : Signal original, sa version bruitée avec 10 dB et le signal debruite

#### 4.7. Résultats récapitulatifs et évaluation de la qualité du débruitage

Les expériences faites dans cette partie de notre travail a aboutit aux résultats donnés au tableau ci-dessous :

	SNR(dB) Parole bruitée	SNR(dB) Parole bruitée	L'erreur quadratique moyenne	Valeur du seuil utilisé	Note du MOS Qualité du son
<b>Haar</b>	40	42.3795	0.0215	0.0015	Très bonne
	20	29.8882	0.0452	0.003	Bonne
	10	18.4091	0.0873	0.0118	Moyenne
<b>Coiffman</b>	40	42.4042	0.0217	0.0015	Très bonne
	20	22.3887	0.0678	0.003	Bonne
	10	12.31	0.1222	0.0118	Moyenne

Tableau 4.1 : Comparaison des résultats obtenus par deux ondelettes (Haar et coiffman)

Du tableau 4.1, nous pouvons faire les remarques suivantes :

- Le débruitage est facile et meilleur lorsque la puissance du bruit accompagnant le signal est faible.
- La qualité du signal debruité (SNR) décroît avec la puissance du bruit.
- L'erreur quadratique moyenne augmente avec le taux du bruit appliqué au signal.

Le tableau ci-dessus permet aussi de comparer l'ondelette de Haar avec ondelette de Coiffman en débruitage d'un signal 1D. On remarque d'après les résultats trouvés, que l'ondelette Haar est meilleure.

### 4.8. Débruitage du bruit quantification

Durant les expériences précédentes, nous avons réduit le bruit additif qui était un bruit gaussien blanc, et maintenant nous allons essayer de réduire un autre type de bruit qui est le bruit de quantification.

Les résultats de débruitage d'un signal de parole quantifié à 8 bits et à 6 bits sont montrés par les figures 4.21 et 4.22 respectivement.

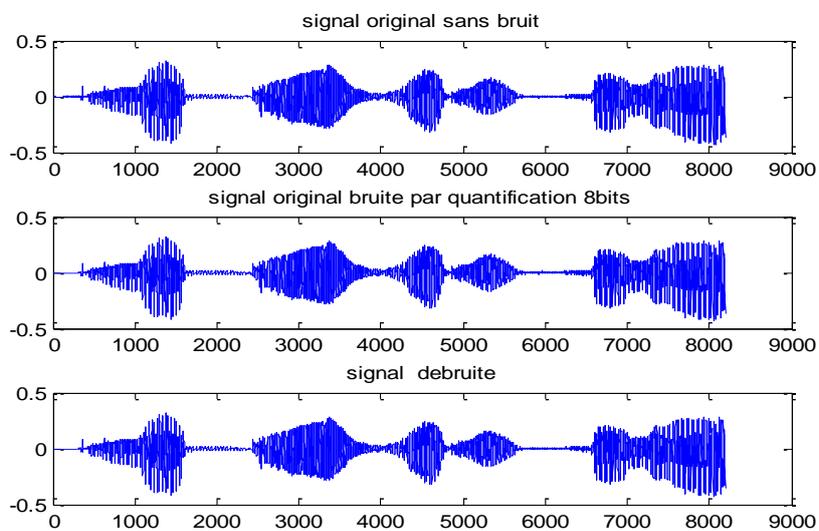


Figure.4.21 : débruitage d'un signal de parole quantifié à 8 bits

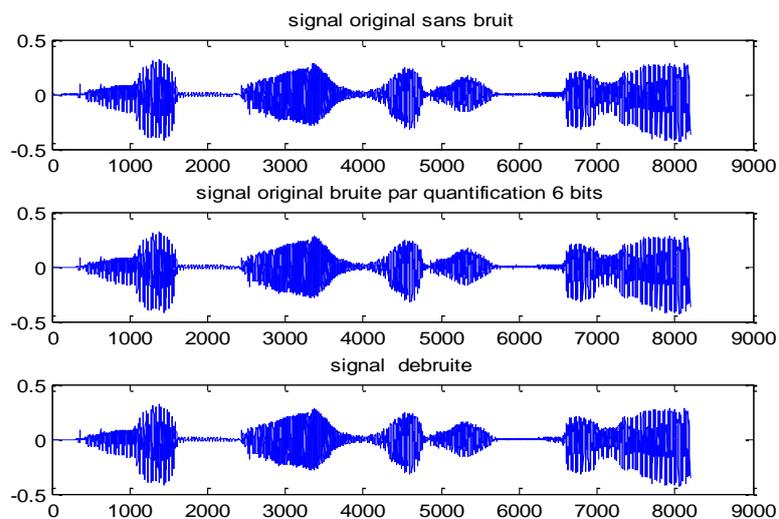


Figure.4.22 : débruitage d'un signal de parole quantifié à 6 bits

Cette simulation a donné les résultats mentionnés dans le tableau ci-dessous.

Quantification à 8bits	SNR_Q8	39.58
	SNR_Q8d	42.6773
	EQM	0.0211
	SEUIL	0.007
Quantification à 6bits	SNR_Q8	27.68
	SNR_Q8d	41.5158
	EQM	0.0226
	SEUIL	0.0008

Tableau 4.2 : débruitage d'un signal quantification à 8 bits et 6 bits

On voit que le bruit de quantification a été notamment réduit par un choix convenable des seuils utilisés dans le seuillage des coefficients de details.

#### 4.9. Conclusion

Dans ce chapitre nous avons simulé différentes expériences pour voir les performances d'ondelettes de Haar en matière de débruitage. Premièrement, nous sommes servis de la boîte outil Matlab pour la décomposition d'un signal à l'aide d'une famille de fonctions appelées ondelettes. Ensuite, nous avons générer des signaux de parole bruités avec différents niveaux du SNR.

Dans toutes les expériences nous avons utilisé un signal de parole original de test (démuni de bruit).

Nous avons commencé par simuler l'analyse et la reconstruction d'un signal sinusoïdal bruité dont on éliminé le bruit par ondelette de Haar. Finalement, nous avons debruité un signal de parole bruité qui est « une phrase en langue française : « la bise et le soleil » avec un SNR=40dB, 20dB et 10dB.

Nous avons terminé cette partie par la réduction du bruit de quantification.

## Conclusion générale

---

### Conclusion générale

Dans ce mémoire nous avons utilisé la transformation en Ondelette pour débruiter un signal de parole. Plusieurs expériences ont été menées afin de montrer la robustesse des ondelettes en matière de débruitage d'un signal monodimensionnel.

Premièrement, nous avons présenté les caractéristiques essentielles du signal vocal par une brève description.

Deuxièmement, nous avons exposé des notions sur les techniques de débruitage de la parole. Quelques types de techniques fondamentales ont été mentionnés. Le débruitage par transformée a attiré plus notre attention car notre travail consiste à débruiter un signal de parole par une transformation orthogonale.

La notion de décorrélation des données était la tâche principale de la transformation utilisée. Une formulation théorique de cette dernière a été détaillée dans le troisième chapitre. Aussi, nous avons présenté les critères d'évaluation de la qualité de compression tels que : l'EQM, SNR et MOS.

La partie expérimentale abordée dans le chapitre 4, a regroupée plusieurs essais permettant d'éclaircir les performances du débruitage par transformée en ondelette. Nous avons simulé l'algorithme de débruitage avec un signal de parole en « langue française », selon les étapes ci-dessous :

- Utilisation de la boîte outil matlab pour débruiter un signal de parole
- Appliquer la décomposition en ondelette de Haar sur une phrase de parole.
- prendre que les coefficients de détails qui sont supérieurs à un certain seuil et rejeter les autres coefficients.
- reconstruire le signal avec les coefficients retenus.
- évaluer la qualité par trois critères SNR, EQM et MOS.

Une comparaison entre ondelette de Haar et Coiffman a été faite, elle a montré la robustesse des ondelettes par rapport à la Coiffman.

Les résultats trouvés sont satisfaisants et promettant pour étendre la technique des ondelettes à d'autres de types de signaux (signaux audio, des signaux 3D, image satellitaire, ...ect).



## *Bibliographie*

- [1] Story, B. H. , Titze, I. R. et Hoffman, E. A. (1996), Vocal tract area functions from magnetic resonance imaging : J. Acous. Soc. Am., 100(1):537–554.
- [2] Cyril Plapous. Traitements pour la réduction de bruit. Application \_a la communication parlée. Signal and Image processing. Université Rennes 1, 2005. French. <tel-00655991>
- [3] “A Comparison of Noise Reduction Techniques for Speech Recognition in Telecommunications Environments” A.G. MAHER, BSc, BE, Professional Assistant, R.W. KING, BEng, PhD, SMIREE Associate Professor, J.G. RATHMELL, BSc, BE, PhD, SMIREE, Senior Lecturer, Department of Electrical Engineering, The Institution of Engineers Australia Communications Conference Sydney, October 20-22 1992
- [4] From: Intro. to Digital Audio Coding and Standards, Bosi & Goldberg
- [5] American Physical Society, 1940
- [6] ABDELKRIM LALLOUANI « débruitage d'un signal de la parole corrompu par un bruit coloré en utilisant la transformée en ondelettes et implantation sur un processeur de traitement numérique des signaux » MONTRÉAL, LE 9 SEPTEMBRE 2004.
- [7] Milind U. Nemade and Satish K. Shah « Speech Enhancement Techniques: Quality vs. Intelligibility” ,International Journal of Future Computer and Communication, Vol. 3, No. 3, June 2014
- [8] Urmila Shrawankar, Dr. Vilas Thakare: « Noise Estimation and Noise Removal Techniques for Speech Recognition in Adverse Environment” Submitted on 11 Aug 2014.
- [9] Asmaa Amehraye : « Débruitage perceptuel de la parole» thèse Soutenue le 15 mai 2009.
- [10] Tarik ALANI, « Introduction aux ondelettes (wavelets) » , 06/11/2008,A2SI-ESIEE – Paris
- [11] Mallat, S. (1998), A wavelet tour of signal processing. New York: Academie Press.
- [12] MATLAB.R2010a.X86.Portable
- [13] Catherine Charles « Introduction aux ondelettes »,université de Liège, 3-4 décembre 2009