

REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE  
MINISTERE DE L'ENSEIGNEMENT SUPERIEUR ET DE LA RECHERCHE SCIENTIFIQUE  
UNIVERSITE MOHAMED BOUDIAF - M'SILA

FACULTE DE TECHNOLOGIE  
DEPARTEMENT D'ELECTRONIQUE  
N°: ... 22 ... /STLC/ 2020



DOMAINE: SCIENCES ET TECHNOLOGIE  
FILIERE: TELECOMMUNICATION  
OPTION: SYS. DES TELECOMMUNICATIONS

**Mémoire présenté pour l'obtention  
du diplôme de Master Académique**

**Par : CHENNI Ahmed Badis et CHARIF Abderrazak**

**Intitulé**

**Analyse et synthèse d'un signal de parole par la  
Matrice de Pencil en vue d'une discrimination de  
locuteurs.**

**Soutenu publiquement le : 30 /09 / 2020 devant le jury composé de:**

Dr. CHALABI Izzeddine	Université M'sila	Président
Dr. KHENNOUF Salah	Université M'sila	Encadreur
Dr. LAIB Abderrazak	Université M'sila	Co-Encadreur
Dr. KENANE El Hadi	Université M'sila	Examineur

**Année universitaire : 2019 /2020**

# Remerciement

---

Tout d'abord, nous remercions ALLAH le tout-puissant, de nous avoir donné la volonté, le courage, la patience et la santé durant toutes ces années d'études et qui grâce à lui ce modeste travail est réalisé.

Par la suite, on souhaite exprimer notre plus profonde gratitude à nos encadreurs, **Dr. KHENNOUF Salah** et **Dr. LAIB Abderrazak**, pour leurs précieux conseils, encouragements enthousiastes, critiques utiles de ce travail et pour ce domaine de recherche si intéressant.

Nos remerciements vont également aux membres du jury, chacun par son nom, pour avoir accepté de faire partie du jury d'évaluation de ce mémoire.

Enfin, j'adresse mes sincères remerciements à tous ceux qui ont contribué, de près ou de loin, à la réalisation de ce travail.

## *Dédicace*

---

### *Je dédie ce travail :*

Aux deux personnes les plus nobles, précieux et les plus chères au monde Ma mère, Mon père que DIEU les gardes.

**A mon cher père** qui n'a jamais cessé de m'encourager et de me donner les conseils fructueux, qui a fait de son mieux pour assurer la continuité de mes études.

*A ma très chère mère, Mère exemplaire pour mes frères et pour moi-même ; tu as su donner l'éducation qu'il nous faut pour affronter les épreuves de la vie.*

*Tu nous as comblés de ton amour maternel et tu répondais présente à chacune de nos sollicitations. Puisse le Tout Puissant t'accorder longue vie afin de profiter des fruits de ce labeur.*

*A mes frères, pour leurs aides, disponibilités et précieux conseils, Que la vie vous apporte toute la joie et le bonheur.*

A tous mes amis sans exception

À tous les personnes qui m'ont souhaité le succès, la joie et le bonheur, et tous ceux qui m'aiment

**CHENNI ahmed Badis**

## *Dédicace*

---

### *Je dédie ce travail :*

Aux deux personnes les plus nobles, précieux et les plus chères au monde Ma mère, Mon père que DIEU les gardes.

**A mon cher père** qui n'a jamais cessé de m'encourager et de me donner les conseils fructueux, qui a fait de son mieux pour assurer la continuité de mes études.

*A ma très chère mère, qui a su donner l'éducation qu'il nous faut pour affronter les épreuves de la vie. Puisse le Tout Puissant t'accorder longue vie afin de profiter des fruits de ce labeur.*

*A mes frères et sœurs*, pour leurs aides, disponibilités et conseils,  
Que la vie vous apporte toute la joie et le bonheur

À tous les personnes qui m'ont souhaité le succès, la joie et le bonheur, et tous ceux qui m'aiment

CHARIF Abderrazak

# Sommaire

---

Remerciements .....	<i>i</i>
Dédicace .....	<i>ii</i>
Liste des figures .....	<i>iii</i>
Liste des tableaux .....	<i>iv</i>
Liste des abréviations .....	<i>v</i>
<b>Introduction générale</b> .....	<b>2</b>
Chapitre-1 : Généralités sur le Traitement Automatique de la Parole .....	<b>5</b>
Introduction général .....	2
1.1 Introduction .....	5
1.2 Traitement Automatique de la voix.....	5
1.2.1 Définition de la parole.....	5
1.2.2 Production de la parole.....	6
1.2.3 Traitement de la parole.....	7
1.2.4 Analyse numérique d'un signal de parole .....	9
1.3 Généralités sur les systèmes de Reconnaissance Automatique de Locuteur.....	10
1.3.1 Principe de fonctionnement.....	10
1.3.2 Classification des systèmes RAL .....	13
1.3.3 Evaluation d'un système de RAL.....	14
1.4 Conclusion.....	15
<b>Chapitre-2 : Méthode proposée pour l'analyse d'un signal de parole</b> .....	<b>17</b>
2.1 Introduction .....	17
2.2 Généralités sur les signaux [7] .....	17
2.2.1 Définition d'un signal.....	17
2.2.2 Définition d'un bruit.....	17
2.2.3 Le traitement de signal .....	18
2.2.4 Fréquence fondamentale dans un signal de parole .....	18
2.3 Classification des signaux [8].....	19
2.3.1 Classification phénoménologique .....	19
2.3.1.1 Signaux déterministes.....	19
2.3.1.2 Signaux aléatoires.....	19
2.3.2 Classification morphologique.....	19
2.4 Analyse spectrale des signaux.....	20
2.4.1 Principe de l'analyse spectrale [9].....	20
2.4.2 Le Spectrogramme .....	22

2.4.3 Limites d'exploitation de la transformée de Fourier [10].....	22
2.5 Estimation paramétrique des signaux.....	23
2.5.1 Différents modèles utilisés dans la modélisation paramétrique .....	23
2.5.2 Passage temps-fréquence.....	24
2.5.3 Méthode de Pencil .....	24
2.5.3.1 Fondements théoriques dans le cas où les données sont non bruitées.....	25
2.5.3.2 Extraction des pôles dominants.....	27
2.5.3.3 Méthode d'analyse proposée basée sur le principe de matrice de Pencil .....	28
2.6 Conclusion.....	30
<b>Chapitre-3 : Expérimentations et résultats obtenus</b> .....	<b>32</b>
3.1 Introduction .....	32
3.2 Base de données utilisée.....	33
3.3 Méthodologie et protocole expérimental.....	33
3.3.1 Influence de la taille des segments et du nombre de pôle sur la résolution.....	34
3.3.2 Comparaison des résultats en variant les paramètres de la matrice de Pencil .....	43
3.4 Conclusion.....	48
Conclusion Générale ET Perspectives.....	50
Références bibliographiques .....	52
Résumé	

## Liste des figures

---

<b>Figure-1.1</b> : Modèle physiologique de la production de la parole.....	6
<b>Figure-1.2</b> : Modélisation de système de production de la parole.....	7
<b>Figure-1.3</b> : Traitement de la parole.....	7
<b>Figure-1.4</b> : Représentation d'un signal de parole, de son spectrogramme et de son énergie...	10
<b>Figure-1.5</b> : Schéma de principe de la phase d'enrôlement d'un système RAL.....	11
<b>Figure-1.6</b> : Schéma de principe de la phase de test d'un système RAL.....	12
<b>Figure-1.7</b> : Evolution des taux de Fausses Acceptations et des Faux Rejets.....	15
<b>Figure-2.1</b> : Chaîne de transmission d'un signal analogique.....	19
<b>Figure-2.2</b> : Classification morphologique des signaux.....	21
<b>Figure-2.3</b> : Signal périodique.....	22
<b>Figure-2.4</b> : Spectre du signal de la figure .....	22
<b>Figure-2.5</b> : Spectrogramme d'un enregistrement de voix chantée.....	23
<b>Figure-2.6</b> : Principe de la méthode basée sur la matrice de Pencil.....	30
<b>Figure-2.7</b> : Méthode d'analyse proposée.....	30
<b>Figure-3.1</b> : Signal de parole original et reconstitué par la matrice de Pencil (Cas d'application : Signal saint d'un homme).....	36
<b>Figure-3.2</b> : Signal de parole original et reconstitué par la matrice de Pencil (Cas d'application : Signal saint d'une femme).....	37
<b>Figure -3.3</b> : Signal de parole original et reconstitué par la matrice de Pencil (Cas d'application : Signal saint d'une femme).....	38
<b>Figure-3.4</b> : Signal de parole original et reconstitué par la matrice de Pencil (Cas d'application : Signal saint d'une femme).....	39
<b>Figure-3.5</b> : Signal de parole original et reconstitué par la matrice de Pencil (Cas d'application : Signal bruité d'un homme).....	40
<b>Figure-3.6</b> : Signal de parole original et reconstitué par la matrice de Pencil (Cas d'application : Signal bruité d'un homme).....	41
<b>Figure-3.7</b> : Signal de parole original et reconstitué par la matrice de Pencil (Cas d'application : Signal bruité d'une femme).....	42
<b>Figure-3.8</b> : Signal de parole original et reconstitué par la matrice de Pencil (Cas d'application : Signal bruité d'une femme).....	43
<b>Figure-3.9</b> : Comparaison des résultats obtenus en utilisant la fréquence fondamentale ( $f_0$ ) (Cas d'application : Signal de parole (saint et bruité) d'une femme).....	44

## Liste des figures

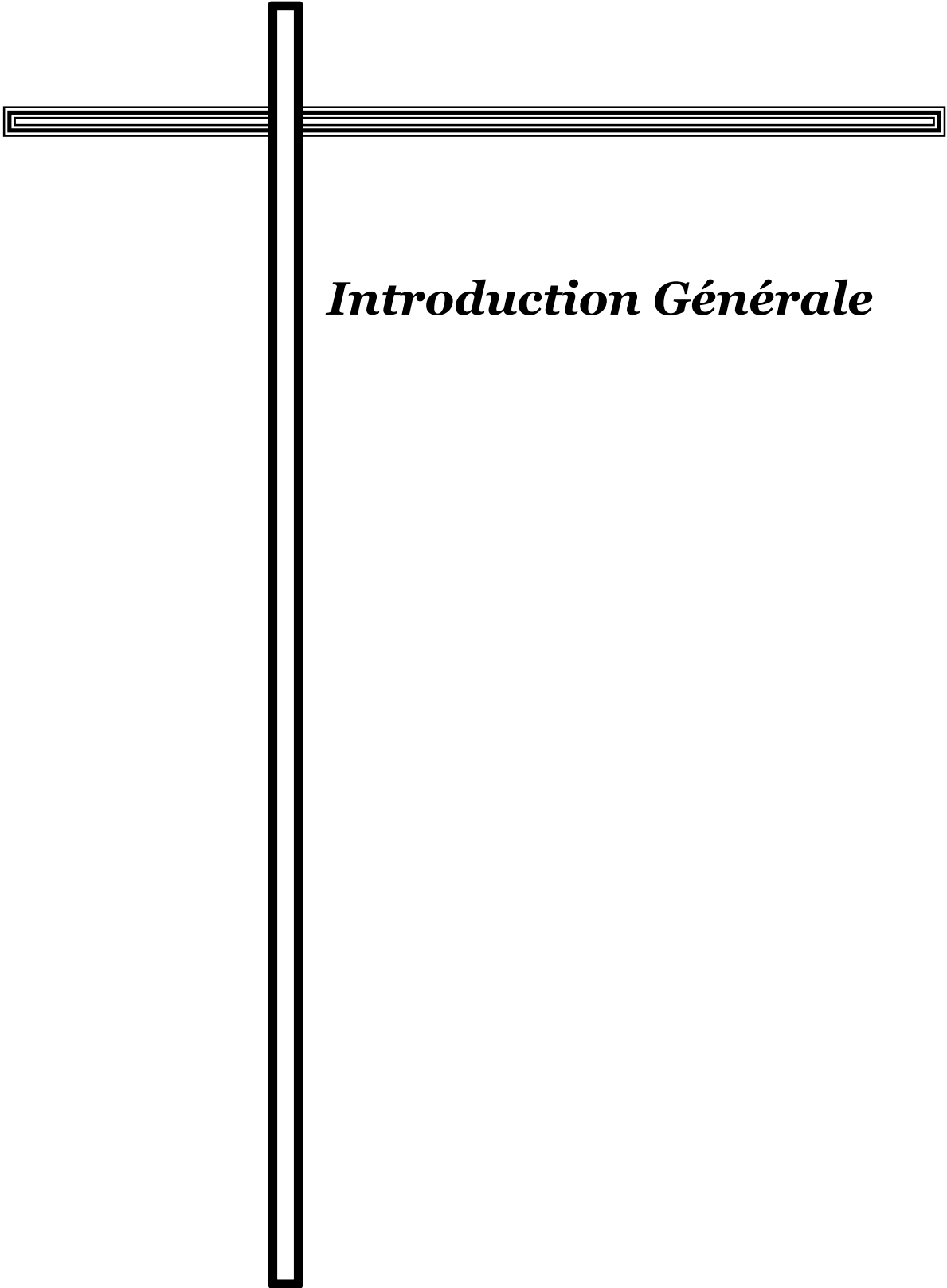
---

<b>Figure-3.10</b> : Comparaison des résultats obtenus en utilisant la fréquence fondamentale ( $f_0$ ) (Cas d'application : Signal de parole (saint et bruité) d'un homme).....	45
<b>Figure-3.11</b> : Comparaison des résultats obtenus en utilisant en utilisant la phase ( $\Phi$ ) (Cas d'application : Signal de parole (saint et bruité) d'une femme).....	45
<b>Figure-3.12</b> : Comparaison des résultats obtenus en utilisant en utilisant la phase ( $\Phi$ ) (Cas d'application : Signal de parole (saint et bruité) d'un homme).....	46
<b>Figure-3.13</b> : Comparaison des résultats obtenus en utilisant le coefficient d'atténuation ( $\alpha$ ) (Cas d'application : Signal de parole (saint et bruité) d'une femme).....	46
<b>Figure-3.14</b> : Comparaison des résultats obtenus en utilisant le coefficient d'atténuation ( $\alpha$ ) (Cas d'application : Signal de parole (saint et bruité) d'un homme).....	47
<b>Figure-3.15</b> : Comparaison des résultats obtenus en utilisant l'amplitude ( $R_0$ ) (Cas d'application : Signal de parole (saint et bruité) d'une femme).....	47
<b>Figure-3.16</b> : Comparaison des résultats obtenus en utilisant l'amplitude ( $R_0$ ) (Cas d'application : Signal de parole (saint et bruité) d'un homme).....	48

## Liste des tableaux

---

<b>Tableau-3.1</b> : Résultats obtenus pour un signal de parole saint d'un homme.....	35
<b>Tableau-3.2</b> : Résultats obtenus pour un signal de parole saint d'une femme.....	38
<b>Tableau-3.3</b> : Résultats obtenus pour un signal de parole bruité d'un homme.....	40
<b>Tableau-3.4</b> : Résultats obtenus pour un signal de parole bruité d'une femme.....	42



## ***Introduction Générale***

## Introduction générale

La communication par la voix est l'un des enjeux majeurs du dialogue Homme Machine, puisque la voix véhicule à la fois un contenu linguistique explicite que l'on peut représenter sous forme écrite et un contenu non linguistique comme le type du locuteur, son attitude, ses gestes, etc. Cela rend le Traitement de la Parole une composante fondamentale des sciences de l'ingénieur et un domaine de recherche actif, au croisement du traitement du signal numérique.

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. Le traitement du signal vocal s'inscrit dans une succession de procédures, que ce soit pour la reconnaissance automatique ou pour la synthèse de la parole Représenté par l'Analyse et synthèse, les deux activités duales, l'analyse fournissant une description du signal acoustique que la synthèse utilise pour le reproduire.

- L'Analyse acoustique est une partie importante dans le traitement que subit le signal sonore pour pouvoir réaliser un système de haute qualité de synthèse, de compréhension, ou de reconnaissance de la parole. Cette opération consiste à tirer à partir du signal vocal un ensemble de paramètres pertinents, discriminants et robustes susceptibles de le représenter il y a Plusieurs techniques d'analyse sont utilisées parmi lesquelles on peut prendre l'analyse par les spectrogrammes.
- La synthèse de parole présente plusieurs avantages, elle est d'une part plus naturelle pour le grand public, elle est plus rapide et efficace qu'un message écrit court et le champ de vision reste libre pour effectuer une autre tâche de lecture. Les deux principaux critères exigés par la synthèse de la voix sont l'intelligibilité et l'aspect naturel. Si de nos jours, le premier critère est atteint, le deuxième est encore au stade de développement. En effet, si les synthétiseurs reproduisent une voix tout à fait intelligible, les intonations et l'expressivité ne sont pas encore au point.

## Introduction générale

Tout système de synthèse de parole à partir du texte, est amené à répondre de manière plus ou moins précise et développée selon la qualité et la finalité du système, représenté par trois problèmes de natures différentes :

Il s'agit dans un premier temps d'analyser et de structurer le texte afin de déterminer un mode de prononciation cohérent ; Par la suite, le texte analysé doit être transformé en une suite de sons de parole accompagnée d'indications concernant leur agencement ; Enfin, il faut générer un signal acoustique, cette suite de sons tout en possédant les caractéristiques apparentes de la parole naturelle.

Le but de notre travail qui s'inscrit dans le domaine de Traitement Automatique de la Parole, en particulier la synthèse de la parole est d'élaborer un système de synthèse de la parole et d'effectuer des modifications prosodiques de signal vocal en utilisant la technique Matrix Pencil.

L'algorithme Matrix Pencil est une technique d'estimation non aléatoire qui utilise un instantané de données pour calculer ses paramètres. L'estimation de la direction d'arrivée était la principale application qui utilisait l'algorithme MP. La communauté du traitement de la parole n'a pas encore identifié tous les avantages de cet algorithme. L'algorithme MP a la capacité de fournir une estimation très précise du spectre du signal. Grâce à cette fonctionnalité, la communauté vocale peut développer diverses façons d'améliorer les applications de traitement de la parole. L'algorithme Matrix Pencil consiste à concaténer, à l'aide d'un lissage, des unités de parole pré- stockées en modifiant le pitch et la durée des segments. Cette technique est associée à la méthode de synthèse par concaténation.

Donc dans ce travail, nous étudierons le programme par la matrice de Pencil l'algorithme Matrix Pencil (MP) pour Analyse et synthèse d'un signal de parole en vue d'une discrimination de locuteurs. Nous avons structure notre travail en 3 chapitres:

- Le chapitre un portera sur généralités sur le traitement automatique de la parole
- Le second chapitre sera consacré à l'analyse et synthèse d'un signal de parole par la matrice de Pencil.
- Dans le troisième et dernier chapitre, les étapes de la Matrice de Pencil seront détaillées et leurs implémentations sous MATLAB données.
- Ce travail sera clôturé par une conclusion générale dans laquelle sera donné un résumé sur les résultats obtenus et les perspectives de développement de ce qui a été réalisé jusqu'ici.

# *Chapitre-1*

## *Généralités sur le traitement automatique de la parole*

## Chapitre-1

### Généralités sur le Traitement Automatique de la Parole

#### 1.1 Introduction

La parole est le moyen de communication directe entre les humains qui est le plus sophistiqué. Les subtiles variations du langage sont capables de susciter chez l'auditeur non seulement une palette forte variée d'émotions et de sentiments, mais aussi une attention complète de son cerveau. De plus en plus, les entreprises de télécommunications et de services (banques, assurances), désireuses d'améliorer leur service à la clientèle, tentent d'introduire des applications basées sur les technologies de la parole.

La palette de ces technologies est fort riche, partant de systèmes de reconnaissance de la parole entraînés pour un seul locuteur à des systèmes capables de reconnaître des centaines de milliers de mots. Dans un autre registre, un grand nombre de services demandent une reconnaissance de l'identité du locuteur (accès aux boîtes vocales, à des services par abonnements, consultation de comptes en banques, etc...).

#### 1.2 Traitement Automatique de la voix

La voix, qui est utilisée de façon naturelle par les êtres humains pour reconnaître un individu, est une modalité comportementale qui peut subir les influences d'une pathologie, d'un stress ou d'un changement émotionnel. Elle peut également être modifiée selon la volonté du locuteur. Elle garde cependant des caractéristiques pertinentes qui permettent d'identifier le locuteur dans le cas où il contrefait sa voix [1].

##### 1.2.1 Définition de la parole

La parole est un signal continu, d'énergie finie, non stationnaire. Sa structure, complexe et variable dans le temps, peut être considérée :

- ❖ Périodique (plus exactement pseudopériodique) pour les sons voisés,
- ❖ Aléatoire pour les sons fricatifs,
- ❖ Impulsionnelle dans les phases explosives des sons occlusifs.

### 1.2.2 Production de la parole

La production de la parole fait intervenir plusieurs organes. La source de la parole provient des poumons qui émettent un flux d'air, qui va traverser le larynx pour faire vibrer ou non les cordes vocales. Il va ensuite traverser le conduit vocal (cavité nasale et buccale) et les articulateurs tels que les lèvres et la langue (figure-1.1).

Cet ensemble agit comme un filtre, considéré comme linéaire, dont la réponse impulsionnelle comporte des fréquences de résonance caractérisées par des pics, appelés formants, dans le spectre du signal de sortie [1].

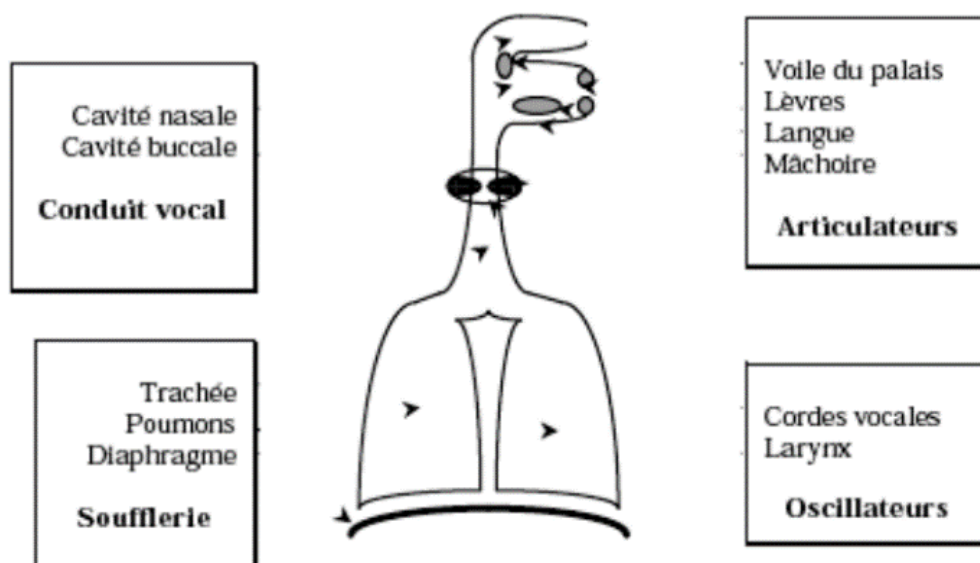


Figure-1.1 Modèle physiologique de la production de la parole.

Le signal résultant est globalement non stationnaire mais peut être considéré comme stationnaire sur de très courtes périodes, de l'ordre de 20ms (signal pseudo-stationnaire). Sur un segment de parole de cette longueur la voix est habituellement et schématiquement séparée en deux classes distinctes : Voisée lorsqu'il y a vibration des cordes vocales, le signal est alors quasi-périodique, non voisée dans le cas d'un simple soufflement, le signal est alors considéré comme aléatoire.

Dans le premier cas, la source d'excitation est modélisée par un train d'impulsions périodiques, de fréquence dite de voisement  $F_0$ , qui correspond à la fréquence de vibration des cordes vocales, la fréquence fondamentale ou pitch ; dans le second cas, la source est modélisée par un bruit blanc. Cette représentation binaire de la production de la parole est représentée dans la figure-1.2.

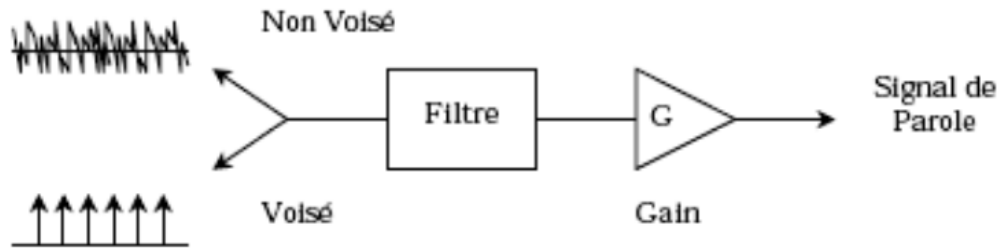


Figure-1.2. Modélisation de système de production de la parole [1].

### 1.2.3 Traitement de la parole

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 (figure 1.3) :

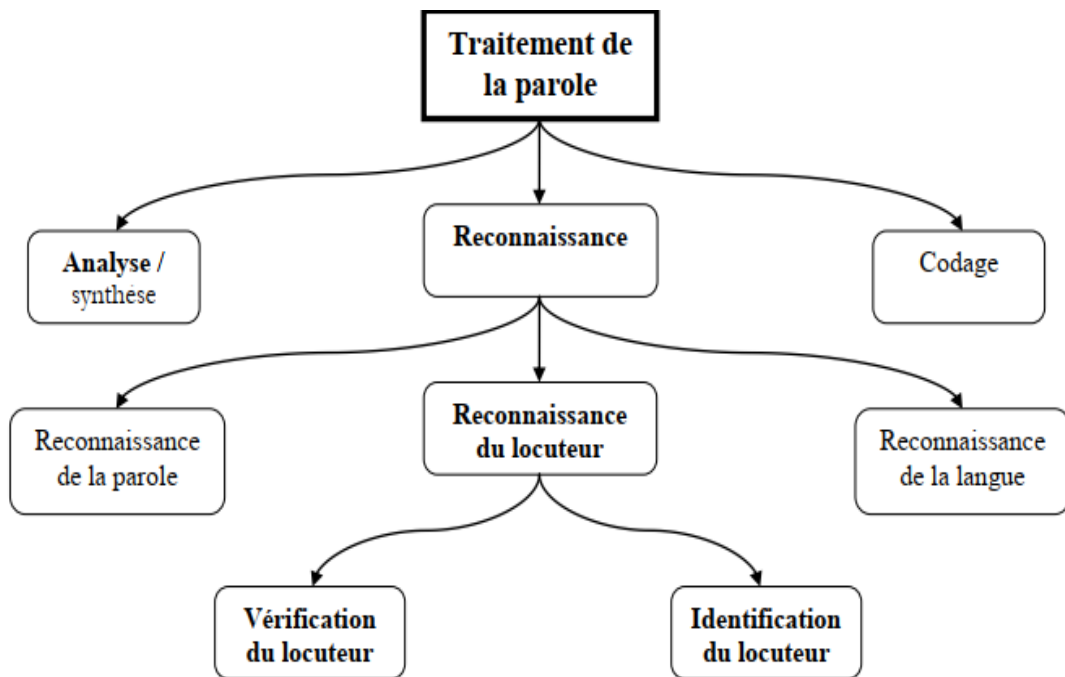


Figure-1.3. Traitement de la parole [2].

#### 1.2.3.1 Codage

Le rôle des codeurs est de permettre la transmission ou le stockage de parole avec un débit réduit, ce qui passe tout naturellement par une prise en compte judicieuse des propriétés de production et de perception de la parole. On comprend aisément que, pour obtenir de bons résultats dans chacune de ces tâches, il faut tenir compte des caractéristiques du signal étudié.

Vu la complexité de ce signal de parole, due en grande partie au couplage étroit entre production, perception, et compréhension, il n'est pas étonnant que les recherches menées par les spécialistes soient directement liées aux progrès obtenus dans de nombreuses autres disciplines scientifiques.

### 1.2.3.2 Reconnaissance

Les reconnaisseurs ont pour mission de décoder l'information portée par le signal de parole à partir des données fournies par l'analyse. On distingue fondamentalement deux types de reconnaissance : la reconnaissance du locuteur, dont l'objectif est de reconnaître la personne qui parle, et la reconnaissance de la parole, où l'on s'attache plutôt à reconnaître ce qui est dit.

En reconnaissance du locuteur, on fait la différence entre l'identification et la vérification du locuteur, selon que le problème est de vérifier que la voix analysée correspond bien à la personne qui est sensée la produire, ou qu'il s'agit de déterminer qui, parmi un nombre fini et préétabli de locuteurs, a produit le signal analysé. On sépare reconnaissance du locuteur dépendante du texte (la phrase à prononcer est fixée dès la conception du système) et reconnaissance indépendante du texte (la phrase à prononcer n'est pas précisée). On distingue enfin reconnaisseurs de parole mono-locuteur, multi-locuteur, ou indépendant du locuteur, ainsi que reconnaisseurs de mots isolés, de mots connectés, et de parole continue.

### 1.2.3.3 Analyse et synthèse

Les analyseurs de parole cherchent à mettre en évidence les caractéristiques du signal de parole tel qu'il est produit, ou parfois tel qu'il est perçu (on parle alors d'analyseur perceptuel). Les analyseurs sont utilisés soit comme composant de base de systèmes de codage, de reconnaissance ou de synthèse (voir ci-dessous), soit en tant que tels pour des applications spécialisées, comme l'aide au diagnostic médical (pour les pathologies du larynx, par analyse du signal de parole) ou l'étude des langues.

Les synthétiseurs sonores permettent de créer de la parole artificielle à partir d'un texte. Pour obtenir ce résultat, on doit utiliser à la fois des techniques de traitement linguistique et des techniques de traitement du signal pour transformer cette version phonétique en son numérisé écoutable sur un haut-parleur.

On distingue deux types de synthétiseurs :

- Synthétiseurs de parole à partir d'une représentation numérique, dont la mission est de produire de la parole à partir des caractéristiques numériques d'un signal vocal telles qu'obtenues par analyse,
- Synthétiseurs de parole à partir d'une représentation symbolique (texte), capables de prononcer n'importe quelle phrase sans qu'il soit nécessaire de la faire prononcer par un locuteur humain au préalable [3].

#### 1.2.4 Analyse numérique d'un signal de parole

Les techniques d'analyse d'un signal de parole offrent une base solide aux techniques d'authentification. Ainsi la reconnaissance du locuteur a bénéficié des nombreux efforts de recherche en traitement du signal, originellement destinés au codage, en synthèse ou en reconnaissance de la parole. Les mécanismes de production ainsi que les paramètres caractéristiques du signal présentés ici en sont directement issus [4].

Les traitements effectués sur la parole sont aujourd'hui réalisés dans le domaine numérique. Au-dessus de 8kHz l'information vocale est négligeable, la bande de fréquence généralement utilisée est [0-8000Hz]. Un échantillonnage du signal de parole à 16kHz convient pour conserver la quasi-totalité de l'information. L'amplitude est alors quantifiée généralement sur 16bits afin d'obtenir une bonne qualité. Pour un codage bas-débit l'échantillonnage est réalisé à 8kHz, ce qui permet de conserver la bande téléphonique (300-3400Hz). Le signal est représenté dans le domaine fréquentiel par l'utilisation des transformées de Fourier, ou encore sous une forme pouvant regrouper les informations de temps et de fréquence : le spectrogramme (Figure-1.4).

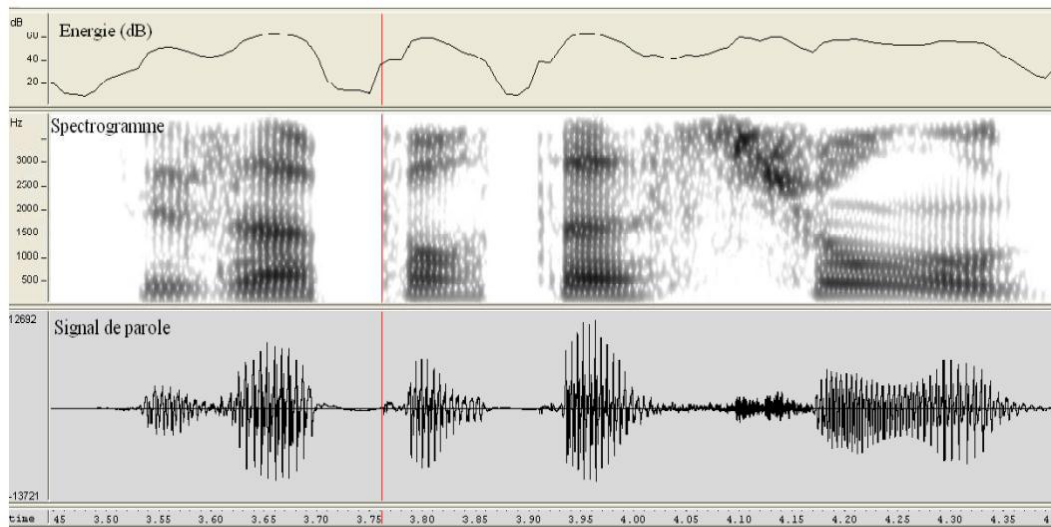


Figure-1.4. Représentation d'un signal de parole, de son spectrogramme et de son énergie [4].

### 1.3 Généralités sur les systèmes de Reconnaissance Automatique de Locuteur

#### 1.3.1 Principe de fonctionnement

La tâche de reconnaissance de personnes se décompose en deux phases. La première étape, appelée enrôlement, consiste à obtenir une représentation du locuteur. Cette étape joue un rôle essentiel dans le processus de reconnaissance. Lors de cette phase, le système construit sa représentation de l'individu, qui sera utilisée par la suite dans son authentification. Cette représentation de l'utilisateur doit être unique et permanente.

La deuxième étape, appelée est étape de test, consiste à soumettre au système les données provenant d'un locuteur souhaitant être authentifié. Cet utilisateur annonce, de plus, une identité connue du système. Le test consiste à mesurer la ressemblance entre les données fournies par l'utilisateur et le modèle existant correspondant à l'identité annoncée. Les étapes d'enrôlement et de test sont décrites dans la suite, sous la forme d'une succession de modules fonctionnels. Nous présentons enfin les éléments permettant d'apprécier les performances des systèmes de d'authentification.

##### 1.3.1.1 Structure de la phase d'enrôlement

La phase d'enrôlement peut être décrite comme l'utilisation de deux modules : module de paramétrisation et module de modélisation, comme représentés (figure-1.5).

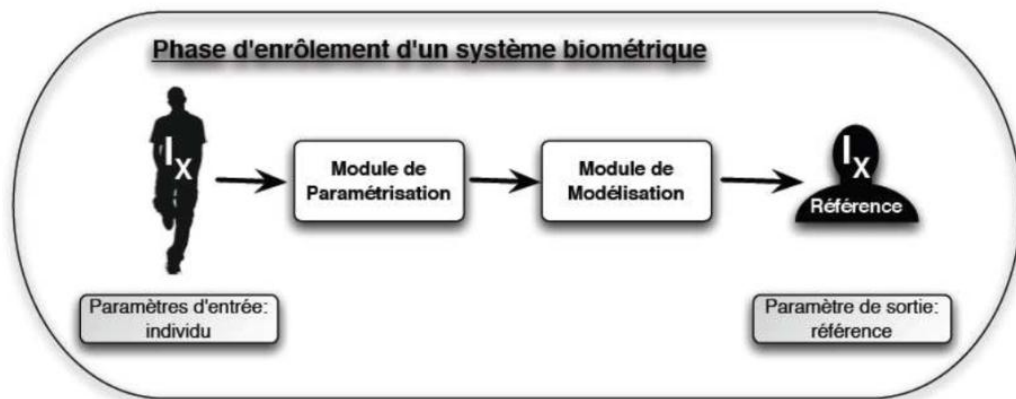


Figure-1.5. Schéma de principe de la phase d'enrôlement d'un système RAL [5].

### A. Module de paramétrisation

La variation de la nature du signal acoustique rend le traitement des données brutes issues de ce dernier très difficile. En effet, ces données contiennent des informations complexes, souvent redondantes et mélangées à du bruit. Le module de paramétrisation, qui traite le signal acoustique reçu, doit remplir plusieurs objectifs :

- Séparer le signal du bruit ;
- Extraire l'information utile à la reconnaissance ;
- Convertir les données brutes à un format directement exploitable par le système.

Chacune de ces tâches pose des problèmes complexes et influe fortement sur les résultats des systèmes automatiques de reconnaissance.

### B. Module de modélisation

Le module de modélisation exploite les données fournies par le module de paramétrisation afin de créer la représentation d'un individu qui servira, par la suite, à l'authentifier. Le modèle utilisé est généralement une représentation statistique des données acquises lors de la phase d'enrôlement. Quelle que soit la nature du modèle créé, il doit cependant respecter certaines contraintes :

- Il doit être le plus précis possible afin de limiter l'ambiguïté inter-individus.
- Il doit prendre en compte la variabilité intra-individus afin de représenter l'individu au cours du temps.

### 1.3.1.2 Structure de la phase de test

La structure de la phase de test est illustrée dans la figure-1.6 suivante et comprend les modules suivant [6].

- o Un module de paramétrisation ;
- o Un module de reconnaissance ;
- o Un module de décision.

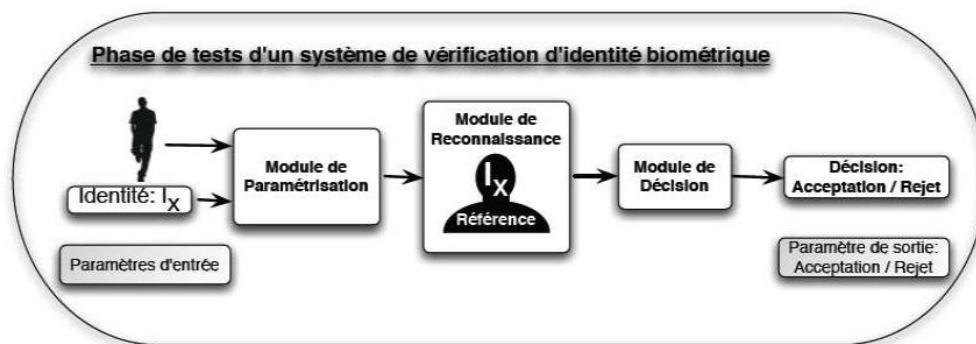


Figure-1.6. Schéma de principe de la phase de test d'un système RAL.

#### A. Le module de paramétrisation

Le module de paramétrisation est le même que celui utilisé durant la phase d'enrôlement. Il remplit, lors de la phase de test, les mêmes tâches que durant la phase de paramétrisation : séparation signal/bruit, extraction de l'information utile et éventuellement, normalisation des données.

Il est important de conserver la même paramétrisation lors des phases d'enrôlement et de test afin de fournir au système automatique des informations comparables et de même nature.

#### B. Le module de reconnaissance

Le module de reconnaissance a un rôle central dans le système RAL. Il compare les paramètres extraits du signal brut à un modèle d'individu calculé lors de la phase d'enrôlement. Lors de cette comparaison, le module de reconnaissance calcule une mesure de similarité entre les données d'entrée et le modèle testé. C'est une valeur numérique, aussi appelée score. Cette mesure de similarité est ensuite transmise au module de décision.

## C. Le module de décision

Comme le décrit la partie précédente, un module de reconnaissance fournit en sortie un score. La nature de ce score varie selon les modules de reconnaissance utilisés. Il s'agit la plupart du temps d'une distance, d'une probabilité ou d'une vraisemblance. Le module de décision doit, à partir de ce score, fournir une décision qui constituera la réponse finale du système RAL.

### 1.3.2 Classification des systèmes RAL

#### 1.3.2.1 Classification par tâche

La reconnaissance du locuteur est un terme générique qui répond à plusieurs définitions selon le scénario applicatif envisagé. Les scénarios applicatifs sont regroupés en cinq catégories principales :

- ⊗ L'identification de locuteurs,
- ⊗ La vérification du locuteur,
- ⊗ L'indexation des documents audio en locuteurs,
- ⊗ La segmentation en locuteur,
- ⊗ Le suivi de locuteurs.

Chacune de ces catégories propose son protocole de reconnaissance selon que l'identité du locuteur à reconnaître soit proclamée, ou que les locuteurs à reconnaître soient connus ou non du système de RAL. Le système de RAL peut valider une identité pour la vérification du locuteur, proposer une identité à partir d'un ensemble de locuteurs, déterminer les durées de parole d'un locuteur, compter le nombre de locuteurs présents dans un signal audio [5].

#### 1.3.2.2 Classification par dépendance au texte

Ces différentes tâches de reconnaissance du locuteur permettent de mettre en place de nombreux scénarios applicatifs distincts. La détermination du scénario s'effectue en fonction des besoins et contraintes exprimés : la tâche (vérifier une identité, segmenter les tours de parole des locuteurs, trouver une identité, etc...), le mode de dépendance au texte (prompté, fixé au préalable, libre), les références connues des locuteurs (taille de la population, sexe de la population, durée des références, etc...), le signal audio disponible pour la tâche (taille,

un locuteur présent ou plusieurs, qualité d'enregistrement, etc...), le matériel disponible (puissance de calcul et de stockage, traitement différé ou en temps réel, etc...). Une telle définition du scénario est nécessaire pour considérer certains principes généraux de RAL tels que : (Vérifier la source pour éviter le plagiat)

- Selon le niveau de dépendance au texte les méthodes mises en œuvre sont différentes,
- La qualité et la quantité des enregistrements d'apprentissage et de test sont déterminantes,
- L'augmentation du nombre de locuteurs à reconnaître diminue les performances en identification du locuteur.

Les méthodes de RAL mises en œuvre peuvent être très différentes selon le scénario ainsi défini.

### 1.3.3 Evaluation d'un système de RAL

Un système de RAL peut être confronté à deux types de tests :

- ❖ **Test clients** : Lors duquel l'échantillon audio présenté au système correspond à l'identité clamée ;
- ❖ **Test imposteur** : Lors duquel l'échantillon biométrique présenté au système provient d'un individu inconnu du système.

Le système automatique doit répondre à chaque tentative d'authentification auquel il fait face par une décision binaire. Il peut donc engendrer deux types d'erreurs :

- ❖ **Faux rejet (FR)** : Erreur commise lorsque le système rejette, à tort, un client légitime (i.e. erreur commise lors d'un test client) ;
- ❖ **Fausse acceptation (FA)** : Erreur commise lorsqu'un imposteur est accepté par erreur en tant qu'utilisateur légitime (i.e. erreur commise lors d'un test imposteur).

Ces deux types d'erreurs n'ont pas toujours la même incidence en terme de sécurité et de qualité de service. La fausse acceptation peut être très pénalisante dans le cas d'une application requérant un niveau de sécurité élevé. Il n'est pas tolérable par exemple que n'importe qui puisse accéder à des informations personnelles, bancaires ou même de type secret défense. Le faux rejet peut également pénaliser des applications où l'utilisateur ne peut se permettre de perdre du temps en tentant de s'authentifier à plusieurs reprises. C'est

le cas, par exemple, pour des services de secours d'urgence. Un utilisateur du système doit pouvoir être reconnu par le système dans les meilleurs délais [5].

Nous verrons dans la suite que les taux de fausses acceptations et de faux rejets sont liés et que le réglage d'un système de vérification d'identité doit tenir compte du coût de chaque type d'erreurs dans le cadre de l'application visée. En fonction du type d'application souhaitée, le seuil de vérification peut être choisi pour minimiser le taux de fausses acceptations : application de sécurité, ou minimiser le taux de faux rejets pour augmenter l'ergonomie d'utilisation. Il n'est pas possible de minimiser conjointement ces deux taux (Figure-1.7) .

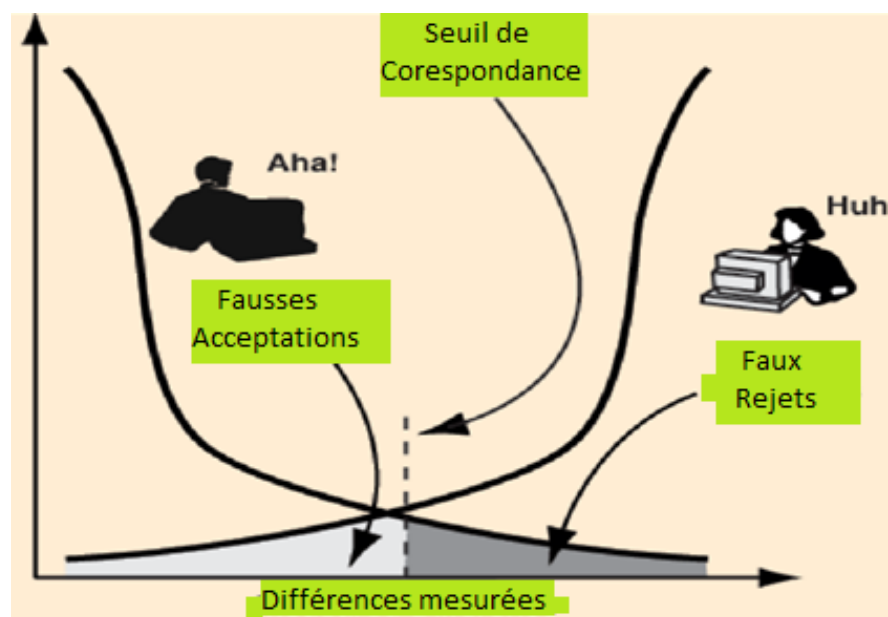


Figure-1.7 Evolution des taux de Fausses Acceptations et des Faux Rejets [5].

#### 1.4 Conclusion

Dans ce chapitre, on a présenté des généralités sur le traitement automatique de la parole. On a défini la parole et le processus de sa production et on a présenté la structure générale d'un système en RAL et ses composants modulaires. Dans le prochain chapitre, nous allons présenter les méthodes proposées pour l'analyse d'un signal de parole.

## ***Chapitre-2***

### ***Méthode proposée pour l'analyse d'un signal de parole***

## Chapitre-2

### Méthode proposée pour l'analyse d'un signal de parole

#### 2.1 Introduction

Le traitement du signal est une discipline indispensable de nos jours. Il a pour objet l'élaboration ou l'interprétation des signaux porteurs d'informations. Son but est donc de réussir à extraire un maximum d'information utile sur un signal perturbé par du bruit en s'appuyant sur les ressources de l'électronique et de l'informatique, et dans le contexte de l'analyse du spectre, les méthodes d'estimation à haute résolution présentent des alternatives intéressantes à la transformée de Fourier classique.

Ces méthodes sont basées sur l'hypothèse que le signal analysé est composé de sinusoïdes amorties, et donnent des estimations de leurs paramètres : fréquences, facteurs d'amortissement, amplitudes d'origine et phases. Cet article étudie l'utilisation d'une méthode d'estimation spécifique à haute résolution, la méthode matrice-crayon, pour l'analyse du spectre des signaux musicaux. Malgré son coût de calcul, il est démontré que cette technique donne de très bons résultats dans le cas de signaux fortement amortis, d'enregistrements de données courts ou de fréquences proches.

#### 2.2 Généralités sur les signaux [7]

##### 2.2.1 Définition d'un signal

Un signal est la représentation physique de l'information, qu'il convoie de sa source à son destinataire. La description mathématique des signaux est l'objectif de la théorie du signal. Elle offre les moyens d'analyser, de concevoir et de caractériser des systèmes de traitement de l'information.

##### 2.2.2 Définition d'un bruit

Le bruit est un phénomène physique qui vient perturber ou gêner la transmission ou l'interprétation d'un signal.

### 2.2.3 Le traitement de signal

C'est la discipline technique qui, s'appuyant sur les ressources de l'électronique, de l'informatique et de la physique appliquée, a pour objet l'élaboration ou l'interprétation des signaux porteurs de l'information. Son application se situe dans les domaines concernés par la transmission ou l'exploitation des informations transportées par ces signaux.

Un exemple de système de mesure a d'une façon générale la structure de la figure-2.1 ci-dessous, le phénomène physique que l'on veut étudier est présenté à un capteur qui le transforme en un signal électrique tension ou courant, à ce niveau un bruit s'ajoute. Le signal transmit à travers le canal de transmission atteint le récepteur, puis il subit un traitement pour extraire l'information utile et exclure le signal de bruit.

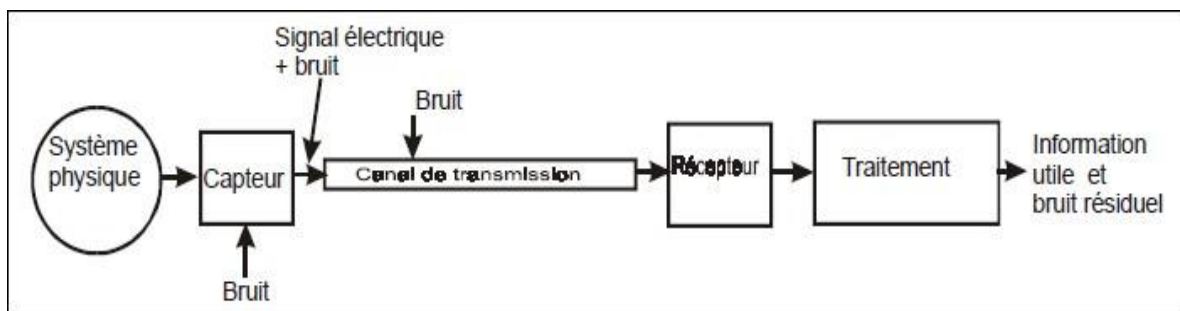


Figure-2.1. Chaîne de transmission d'un signal analogique.

### 2.2.4 Fréquence fondamentale dans un signal de parole

La fréquence fondamentale  $F_0$  (ou pitch), qui est la période des sons voisés, joue un rôle important dans un signal de parole, car c'est elle qui véhicule une grande partie de l'information prosodique. L'intensité de la voix et les durées successives des syllabes complètent ces informations. La prosodie, qui peut être considérée comme l'effet des différentes variations de la fréquence fondamentale  $F_0$ , de l'intensité et de la durée, peut faire ressortir bien des caractéristiques du locuteur (son genre, son âge, ses émotions, etc.).

La plage de variation moyenne de la fréquence fondamentale varie d'un locuteur à un autre en fonction de son âge et de son sexe. Elle s'étend approximativement de 80 à 200 Hz chez les hommes, de 150 à 450 Hz chez les femmes, et de 200 à 600 Hz chez les enfants. La détection de  $F_0$  joue un rôle essentiel dans le domaine de traitement de la parole et doit être, si possible, pour les applications modernes, calculée en temps réel.

## 2.3 Classification des signaux [8]

En traitement de signal, on distingue deux types de classification ; classification phénoménologique et classification morphologique.

### 2.3.1 Classification phénoménologique

Dans ce type de classification, on considère la nature de l'évolution du signal en fonction du temps. Il apparaît deux types de signaux : Les signaux déterministes (ou certains) et les signaux aléatoires (probabilistes).

#### 2.3.1.1 Signaux déterministes

Les signaux déterministes (ou certains) sont des signaux dont leur évolution en fonction du temps peut être parfaitement décrite par un modèle mathématique. Ces signaux, qui peuvent être reproduits, proviennent de phénomènes pour lesquels on connaît les lois physiques correspondantes et les conditions initiales, permettant ainsi de prévoir le résultat. On retrouve dans cette classe les signaux périodiques, les signaux transitoires, les signaux pseudo-aléatoires, etc...

#### 2.3.1.2 Signaux aléatoires

Les signaux aléatoires (probabilistes) sont des signaux dont leur comportement temporel est imprévisible. Pour la description de ces signaux, on fait appel à leurs propriétés statistiques. Lorsque, leurs propriétés statistiques sont invariantes dans le temps, on dit qu'ils sont stationnaires.

### 2.3.2 Classification morphologique

Le temps est un paramètre important de classification. Le traitement numérique des signaux conduit à faire la distinction entre les signaux dits à temps continus (signaux continus) et les signaux dits à temps discrets (signaux discrets ou échantillonnés). Un autre paramètre des signaux traités est à prendre en compte, c'est l'amplitude qui peut aussi être continue ou discrète (quantifiée). Dans ce type de classification, on considère les signaux à évolution temporelle continue et les signaux à évolution temporelle discrète ainsi que ceux dont l'amplitude est continue ou discrète.

Dans cette catégorie de classification, on obtient quatre type de signaux : les signaux analogiques, les signaux quantifiés, les signaux échantillonnés et les signaux numériques (figure-2.2).

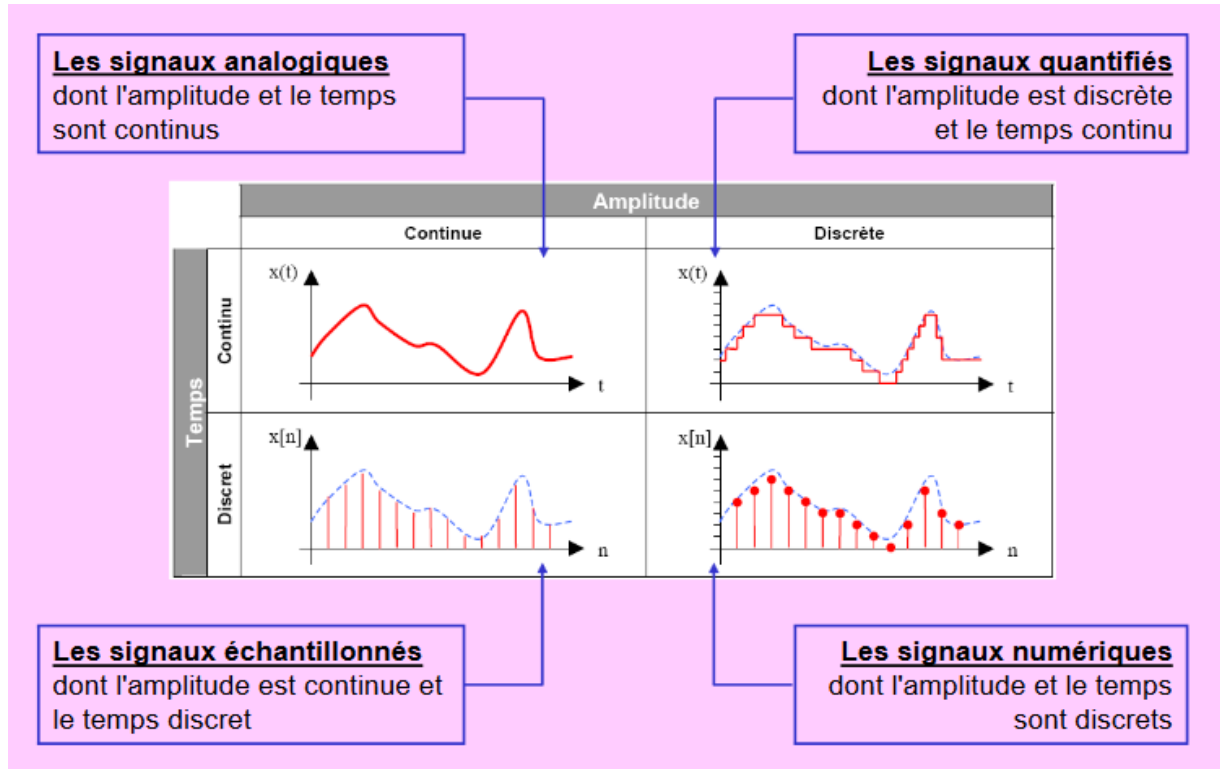


Figure-2.2. Classification morphologique des signaux [9]

## 2.4 Analyse spectrale des signaux[10]

### 2.4.1 Principe de l'analyse spectrale

Le but de l'analyse spectrale est de décomposer un signal en des signaux sinusoïdaux élémentaires. Un signal se représente naturellement par sa variation par rapport au temps. Le comportement d'un certain nombre de dispositifs se décrit facilement pour un signal sinusoïdal permanent (régime harmonique). Il convient alors de décomposer les signaux en signaux sinusoïdaux élémentaires. Selon la forme du signal, cette décomposition est plus ou moins difficile. Pour simplifier le problème, nous envisagerons des signaux périodiques en prenant comme exemple le cas du signal de la figure-2.3 ci-dessous.

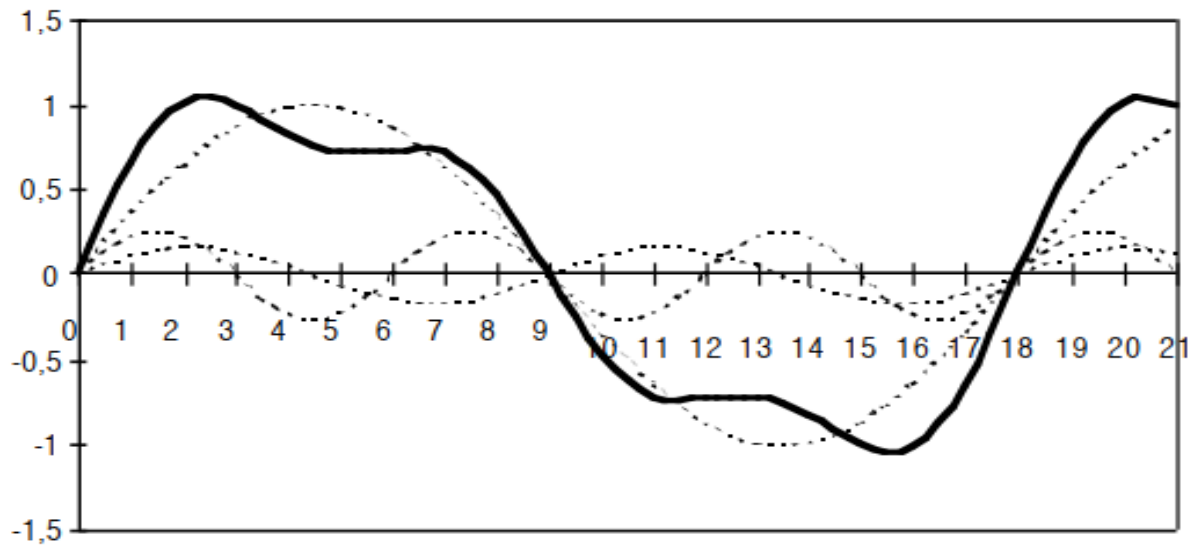


Figure-2.3. Signal périodique

Ce signal est constitué de trois composantes sinusoïdales élémentaires. On appelle spectre, la représentation des composantes suivant l'axe des fréquences (figure-2.4).

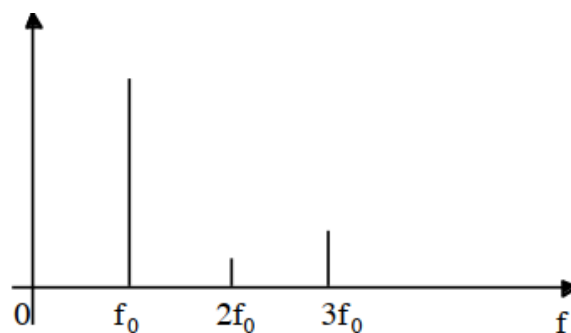


Figure-2.4. Spectre du signal de la figure 2.3

On représente en ordonnées l'amplitude des différentes composantes du signal. Pour que l'information soit complète, il faudrait associer à chaque raie sa phase. L'appareil qui permet de mesurer l'amplitude des différentes composantes d'un signal est un l'analyseur de spectre. Généralement, l'analyse mathématique des signaux est réalisée par la transformée de Fourier ou par la transformée en Z. Dans le cas où la transformée de Fourier permet des passages temps fréquence et inversement, les chercheurs, utilisant des calculateurs numériques, ne peuvent traiter directement l'information en continu, elle doit être discrétisée, échantillonnée pour être exploitable.

### 2.4.2 Le Spectrogramme

Le spectrogramme est un outil de visualisation utilisant la technique de la transformée de Fourier et donc du calcul de spectres. Largement utilisé en 1947, à l'apparition du sonographe, le spectrogramme est devenu l'outil incontournable des études en phonétique pendant plusieurs années. L'apparition de l'informatique et les écrans graphiques de haute qualité a permis d'abandonner le sonographe mais la technique du spectrogramme est toujours utilisée du fait de sa simplicité de mise en œuvre et du grand nombre d'études qui ont déjà été réalisées. Le spectrogramme permet de mettre en évidence les différentes composantes fréquentielles du signal à un instant donné, une transformée de Fourier rapide étant régulièrement calculée à des intervalles de temps rapprochés (figure-2.5).

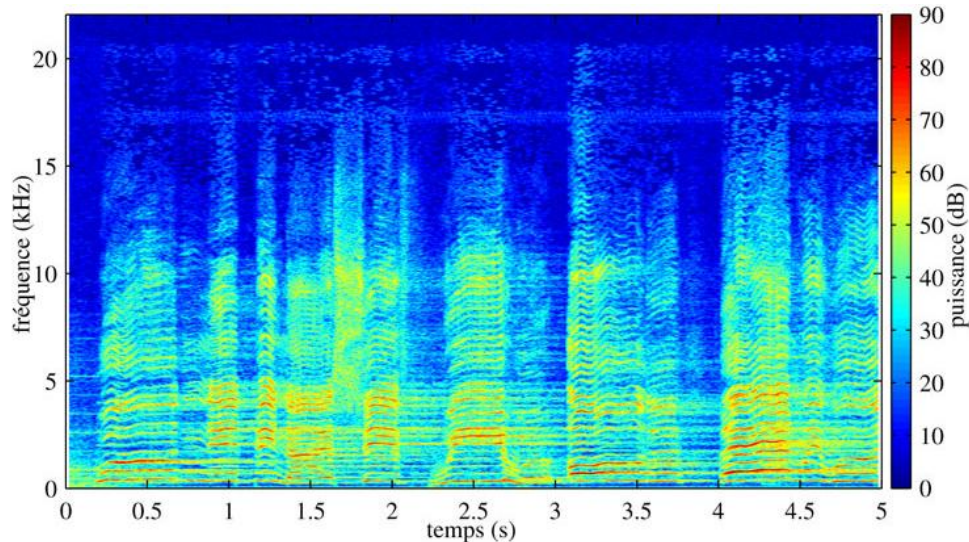


Figure-2.5. Spectrogramme d'un enregistrement de voix chantée

### 2.4.3 Limites d'exploitation de la transformée de Fourier [10]

La transformée de Fourier et l'implantation algorithmique efficace qui y a été associée, la Transformée de Fourier Rapide (FFT), présente de nombreux avantages en tant que méthode d'analyse temps-fréquence. La rapidité de sa mise en œuvre l'a projeté au rang d'élément incontournable des systèmes de traitement de signal. Mais, après l'apparition de la notion de représentation temps-fréquence, qui fait suite à l'utilisation de représentations spectrographiques, les études théoriques du domaine ont permis de mettre à jour quelques désavantages qui sont impossibles à éliminer et qui constituent ainsi les limites d'exploitation de la transformée de Fourier.

La transformée de Fourier et les autres transformées existantes ne permettent donc pas de représenter correctement un signal temporel discret, qui est déjà une approximation de la réalité. Ainsi, bien que la transformée de Fourier permette d'extraire d'un signal des connaissances a priori inaccessibles, l'information obtenue ne peut pas, théoriquement, être correcte. Mais ces limites théoriques relatives aux représentations temps-fréquence ne sont pas les seuls problèmes existants.

Le défaut majeur de la transformée de Fourier pour l'étude de la parole vient de l'inévitable intermodulation source/conduit présente dans le spectre qui ne permet pas de connaître précisément la hauteur du fondamental. Cette intermodulation est due à la convolution qui est réalisée par le conduit vocal sur la fréquence fondamentale produite par les cordes vocales.

## 2.5 Estimation paramétrique des signaux

L'estimation de caractéristique spectrale de signaux discrets déterministes ou stochastique est généralement basée sur l'utilisation de la transformée de Fourier cette approche de l'analyse spectrale se révèle intéressante du point de vue calculabilité et donne des résultats satisfaisants pour une large classe de signaux, cependant plusieurs inconvénients inhérents à la méthode en limitent ses performances.

L'utilisation de méthodes d'estimation spectrale plus performantes sur de courts échantillons (telles les méthodes paramétriques) représente alors une alternative intéressante. Les techniques de modélisation paramétrique permettent de déterminer les paramètres d'un modèle mathématique décrivant un signal, un système ou un processus.

### 2.5.1 Différents modèles utilisés dans la modélisation paramétrique

Les différents modèles utilisés dans la modélisation paramétrique sont [11] :

- ❖ **Modèles rationnel linéaires** : Le signal à modéliser  $x(k)$  est considéré comme un processus aléatoire approximé par un modèle rationnel linéaire, appelé dans le cas général modèle ARMA (Auto Regressive Moving Average), et défini par l'équation de récurrence suivante :

$$x(k) = - \sum_{i=1}^p a_i x(k-i) + \sum_{i=1}^q b_i u(k-i) \quad (2.1)$$

- ❖ **Modèles exponentiels (Prony) :** Le modèle considéré est une somme de fonctions exponentielles amortie ou pas.
- ❖ **Modèles géométriques :** Une approche différente repose sur une exploration des propriétés géométriques des sous-espaces de signaux, générés par des modèles par exemple exponentiels, et des sous-espaces de bruit, générés par des bruits additifs.

### 2.5.2 Passage temps-fréquence

Les méthodes d'estimations paramétriques de Prony et de Pencil sont basées sur un modèle de signal (une somme d'exponentielles amorties) caractérisé notamment par son ordre et des couples de résidus et de pôles. Cette technique d'estimation permet de déterminer les pôles du modèle. Il est intéressant de remarquer que le formalisme utilisé par ces deux méthodes est très proche de celui de la transformée en Z. En effet, le modèle proposé correspond à une description classique par pôles ( $p_i$  ou  $z_i$ ) et résidus ( $R_i$ ) qui permettent de passer facilement d'un espace à l'autre par les relations suivantes :

$$T_Z\{x(t)\} = X(z) = \sum_{k=0}^{+\infty} x(k) z^{-k} = \sum_{k=0}^{+\infty} \left( \sum_{i=0}^{+\infty} R_i z_i^k \right) z^{-k} = \sum_{i=1}^M \frac{R_i}{1 - z_i z^{-1}} \quad (2.2)$$

$$TL[x(t)] = X(p) = \int_0^{+\infty} x(t) e^{-pt} dt = \int_0^{+\infty} \left( \sum_{i=1}^M R_i e^{z_i t} \right) e^{-pt} dt = \sum_{i=1}^M \frac{R_i}{p - p_i} \quad (2.3)$$

### 2.5.3 Méthode de Pencil

La méthode de Pencil [12], [13] développée par Hua et Sarkar en 1989, utilise le même modèle que celui de Prony et propose une autre solution pour déterminer les pôles du modèle. Cette approche se révèle plus efficace dans le sens où l'estimation des pôles n'est effectuée qu'en une seule étape au lieu de deux étapes. Elle utilise le principe de la décomposition en valeurs singulières et approximation de la matrice de données et introduit une redondance supplémentaire, en s'appuyant sur les propriétés d'une matrices (appelé matrice Pencil) qui est défini à l'aide de deux matrices de données décalées.

#### 2.5.3.1 Fondements théoriques dans le cas où les données sont non bruitées

Le fonctionnement théorique de la méthode de Pencil repose sur les propriétés d'une matrice construite à partir de deux matrices de données  $Y_1$  et  $Y_2$  de dimension  $(N - L) \times L$ .

$$[Y_1] = \begin{bmatrix} x(0) & x(1) & \dots & x(L-1) \\ x(1) & x(2) & \dots & x(L) \\ \vdots & \vdots & \ddots & \vdots \\ x(N-L-1) & x(N-L) & \dots & x(N-2) \end{bmatrix} (N-L) * L \quad (2.4)$$

$$[Y_2] = \begin{bmatrix} x(1) & x(2) & \dots & x(L-1) \\ x(2) & x(3) & \dots & x(L) \\ \vdots & \vdots & \ddots & \vdots \\ x(N-L) & x(N-L+1) & \dots & x(N-1) \end{bmatrix} (N-L) * L \quad (2.5)$$

La matrice Pencil est définie par la formule suivante :

$$P(\lambda) = [Y_2] - \lambda[Y_1] \quad (2.6)$$

Le principe de la méthode repose alors sur la propriété suivante : Si le paramètre Pencil L satisfait la condition  $M \leq L \leq N - M$ , alors les pôles  $z_i$ ,  $i = 1, \dots, M$ , sont réducteurs de rang pour la matrice  $P(\lambda)$ , c'est-à-dire :

$$\begin{cases} \text{rang}([Y_2] - \lambda[Y_1]) = M, \text{ si } \lambda \neq z_i \\ \text{rang}([Y_2] - \lambda[Y_1]) = M - 1, \text{ si } \lambda = z_i \end{cases} \quad (2.7)$$

La démonstration de cette propriété s'appuie sur la relation :

$$P(\lambda) = [Y_2] - \lambda[Y_1] = [Z_1][R][Z_0][Z_2] - \lambda[Z_1][R][Z_2] \quad (2.8)$$

$$P(\lambda) = [Z_1][R]\{[Z_2] - \lambda[I]\}[Z_2]; \forall \lambda \quad (2.9)$$

Avec :

I : Matrice Identité carrée ( $M \times M$ ) et en introduisant les matrices  $Z_1$ ,  $Z_2$ ,  $Z_0$  et R.

Tels que:

$$[Y_2] = [Z_1][R][Z_0][Z_2] \quad (2.10)$$

$$[Y_2] = [Z_1][R][Z_2] \quad (2.11)$$

Et définie de la façon suivante :

$$[Z_1] = \begin{bmatrix} 1 & x(1) & \dots & 1 \\ z_1 & z_2 & \dots & z_M \\ \vdots & \vdots & \ddots & \vdots \\ z_1^{N-L-1} & z_2^{N-L-1} & \dots & z_M^{N-L-1} \end{bmatrix} (N-L) * M \quad (2.12)$$

$$[Z_2] = \begin{bmatrix} 1 & z_1 & \dots & z_1^{L-1} \\ 1 & z_2 & \dots & z_2^{L-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & z_M & \dots & z_M^{L-1} \end{bmatrix} L * M \quad (2.13)$$

$$[Z_0] = \text{diag}[Z_1, Z_2, \dots, Z_M] \quad (2.14)$$

$$[R] = \text{diag}[R_1, R_2, \dots, R_M] \quad (2.15)$$

Où la notation « diag » représente une matrice carrée diagonale.

Lorsque  $\lambda = zi$ , alors le terme diagonal d'ordre  $i$  dans la matrice  $(Z_0 - \lambda I)$  devient nul ce qui revient à annuler une colonne dans  $Z_1$  et une ligne dans  $Z_2$ . On peut alors distinguer deux cas :

- Si  $M \leq L \leq N - M$  alors  $Z_1$  et  $Z_2$  sont de rang  $M$ , et la matrice  $[Y_2] - z_i [Y_1]$  sera de rang  $(M - 1)$ , pour  $i = 1, \dots, M$  ; la valeur de  $z_i$  est réductrice de rang pour la matrice Pencil.
- Si  $M > L$  ou  $M > N - L$  alors, respectivement  $\text{rang}([Z_2]) = L$  ou  $\text{rang}([Z_1]) = N - L$  et  $\text{rang}([R]\{[Z_0] - \lambda[I]\}[Z_2]) = L$  ou  $\text{rang}([Z_1][R]\{[Z_0] - \lambda[I]\}) = N - L$  ; par suite le rang de la matrice  $P(\lambda)$  ne décroît pas lorsque  $\lambda = z_i$ .

La matrice  $P(\lambda)$  comporte  $(N - L)$  lignes et  $L$  colonnes. On suppose que la condition  $M \leq L \leq N - M$  est remplie. Comme  $P(z_i)$  est de rang  $M-1$  il existe au moins un vecteur non nul  $u_i$  tel que :

$$P(z_i)u_i = ([Y_2] - z_i[Y_1])u_i = 0 \quad (2.16)$$

Le vecteur  $u_i$  est de longueur  $N - L$  et il est vecteur propre de la matrice  $[Y_1] + [Y_2]$  associée à la valeur propre  $z_i$  soit :

$$[Y_1] + [Y_2]u_i = z_i u_i \quad (2.17)$$

Cela signifie que les  $z_i$  peuvent être déterminés à partir des valeurs propres généralisées de la matrice paire  $\{Y_2, Y_1\}$ , ou de façon équivalente à partir des valeurs propres ordinaires de la matrice  $[Y_1] + [Y_2]$ . Les résidus sont alors déterminés comme pour la méthode de Prony par la méthode des moindres carrés. Ainsi dans la méthode de Pencil, les solutions ne sont plus obtenues à partir de la recherche des racines d'un polynôme, mais par la recherche des valeurs propres d'une matrice. Si le signal est non bruité, cette matrice  $Y_1 + Y_2$  admet  $M$  valeurs propres non nulles et  $L - M$  valeurs propres nulles.

En revanche en présence de bruit, toutes les valeurs propres sont en général non nulles. Pour améliorer les performances de la méthode, une décomposition en sous espaces peut être mise en place [14][15].

### 2.5.3.2 Extraction des pôles dominants

Dans ce paragraphe nous proposons une extraction des pôles dominants après utilisation de la méthode de Pencil. La démarche théorique est la suivante :

$$x(t) = \sum_{i=1}^N R_i \exp(p_i t) \quad (2.18)$$

- $x(t) = \text{signal}$ ;
- $M$  : Nombre de poles ;
- $R_i$  : Residus (amplitudes complexes)
- $p_i : \sigma_i + j\omega_i$  ;
- $\sigma_i$  : Coefficients d'amortissement ;
- $\omega_i$  : Frequences angulaire ( $\omega_i = 2\pi f_i$ )

Après l'échantillonnage, la variable de temps,  $t$ , est remplacé par  $k\Delta t$ , avec  $\Delta t$  est le pas d'échantillonnage ;

$$x(t) = \sum_{i=1}^N R_i z_i^k = \sum_{i=1}^N R_i \exp(p_i k \Delta t), \forall k \in [0, N-1], (R_i, z_i) \in \mathbb{C}^2 \quad (2.19)$$

Avec :

$$z_i = \exp(p_i T_s) = \exp((\sigma_i + j\omega_i) \Delta t);$$

L'énergie de chaque composante entre deux instants  $t_1$  et  $t_2$  est donnée par :

$$\begin{aligned} E_i &= \int_{t_1}^{t_2} |R_i \exp\{(\sigma_i + j\omega_i)t\}|^2 dt \\ &= \int_{t_1}^{t_2} |R_i|^2 \exp(2\sigma_i t) dt \\ &= \frac{|R_i|^2}{(2\sigma_i)} \{\exp(2\sigma_i t_2) - \exp(2\sigma_i t_1)\} \end{aligned} \quad (2.20)$$

Ou bien :

$$E_i = |E_i|^2 \sum_{k=0}^{N-1} |z_i|^{2k} \Delta t \quad i = 1 \dots M; \quad (2.21)$$

### 2.5.3.3 Méthode d'analyse proposée basée sur le principe de matrice de Pencil

La première étape de la résolution en utilisant la méthode de Pencil consiste à choisir une valeur pour le paramètre Pencil L, afin de former la matrice de donnée Y. Le paramètre L est également très important pour l'élimination de quelques effets du bruit dans la matrice de données. Les résultats de simulation ont montré que les meilleurs choix de L sont entre N/3 et 2N/3, ce choix permet de minimiser la variance des paramètres détectés. Dans cet intervalle, la précision des résultats obtenus sera quasi-semblable.

En conclusion, la méthode d'analyse d'un signal de parole par la matrice de Pencil que nous avons proposé est constitué de trois étapes comme suit :

- Segmentation,
- Application de la matrice de Pencil pour chaque segment S de parole,
- Calcul du vecteur moyen pour chaque paramètre  $(f), (\Phi), (\alpha), (R_0)$ ,

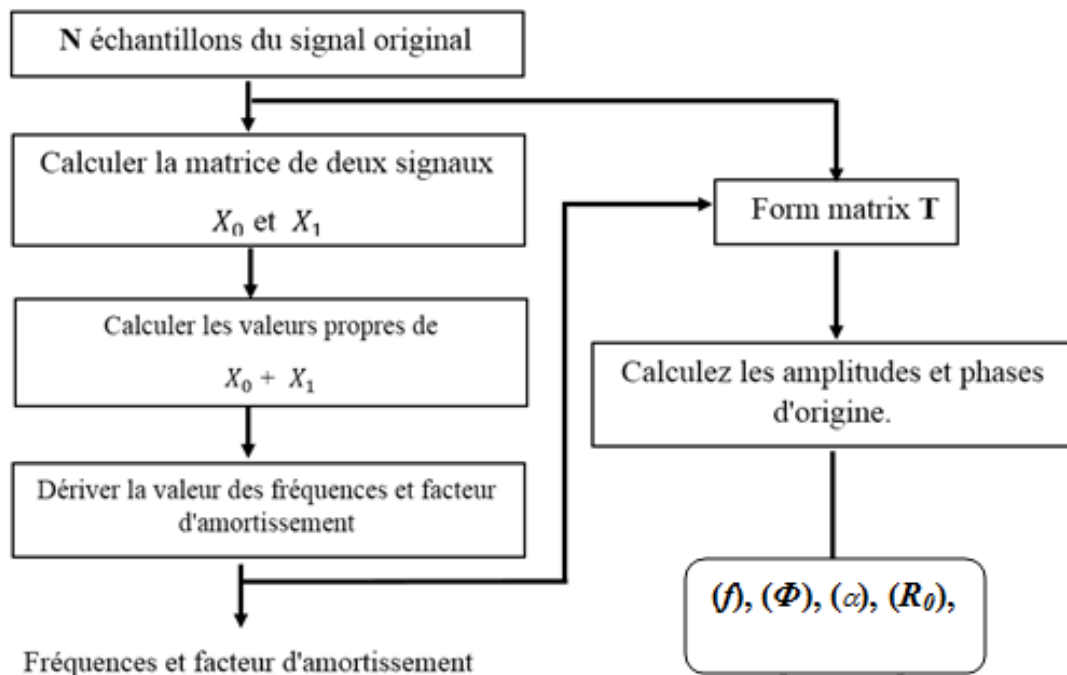


Figure-2.6. Principe de la méthode basée sur la matrice de Pencil.

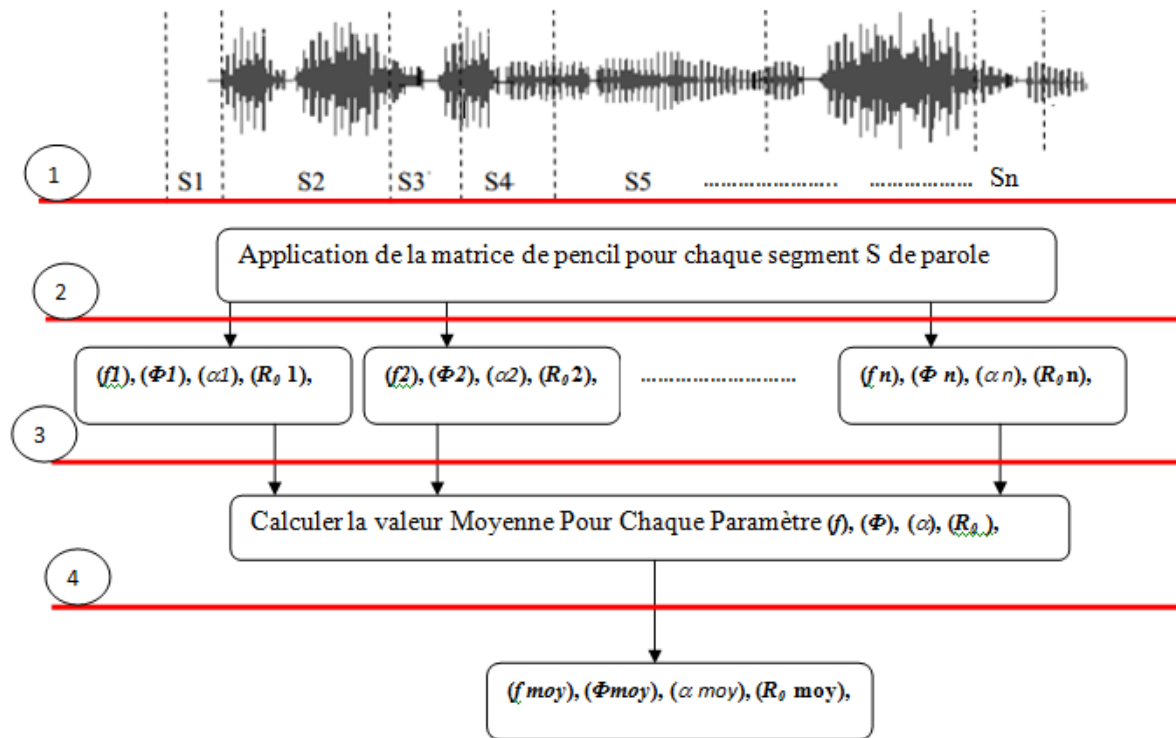


Figure-2.7. Méthode d'analyse proposée

## 2.6 Conclusion

Dans ce chapitre nous avons présenté en détails la méthode proposée pour l'analyse de signal de parole basée sur la matrice de Pencil. D'abord, nous avons commencé par des généralités sur les signaux suivie d'une présentation de la classification des signaux ensuite nous avons discuté les avantages et les inconvénients des différentes techniques rencontrés dans l'analyse spectrale. Par la suite, nous avons expliqué en détails le principe de la méthode de la matrice de Pencil, et enfin nous avons donné les algorithmes utilisés dans nos programmes pour l'extraction de caractéristique et création du modèle référence par cette dernière. Dans le chapitre suivant nous exposerons les séries d'expériences et les résultats obtenus.

## ***Chapitre-3***

***Expérimentations***

***et***

***Résultats obtenus***

## Chapitre-3

### Expérimentations et résultats obtenus

#### 3.1 Introduction

L'information contenue dans un signal de parole comporte une partie dépendante du message et une partie caractéristique du locuteur. La modélisation du locuteur se fait en extrayant un jeu de vecteurs de paramètres pour créer un modèle spécifique à chaque locuteur à partir de ces données. La paramétrisation utilisée pour la reconnaissance de la parole, bien que développée dans le but d'atténuer la variabilité des prononciations (locuteurs, accents, etc...), reste aussi utilisée pour la reconnaissance du locuteur. Plusieurs raisons nous ont motivés à étudier cette paramétrisation en y détectant les indices les plus spécifiques à la reconnaissance du locuteur tels que :

- ❖ Améliorer les performances des systèmes de reconnaissance du locuteur en fonctionnant dans un milieu ouvert (robustesse à l'imposture et au bruit).
- ❖ Rechercher d'un jeu de paramètres caractérisant le locuteur,
- ❖ Etudier l'évaluation intra et interlocuteurs afin de comprendre l'évolution de la voix d'un locuteur au cours du temps.

Pour ce faire, il faut un système qui permet d'analyser la parole, d'en extraire des paramètres que l'on pourra transformer et finalement qui nous permet de restituer ces paramètres pour générer un signal de parole. Nous utiliserons comme algorithme la matrice de Pencil, Afin de déterminer les paramètres distinctifs du locuteur et ceux qui dépendent du message. Pour comprendre et identifier les paramètres importants, nous avons choisi de décomposer la parole en analysant ses caractéristiques fréquentielles. Enfin, nous postulons que les paramètres les plus discriminants sont ceux qui permettent de transformer la voix d'un locuteur pour ressembler au mieux à celle d'un locuteur cible.

Dans ce chapitre, nous présentons une évaluation expérimentale de l'approche basée sur la matrice de Pencil que nous avons proposé pour l'analyse (décomposition et reconstitution) d'un signal de parole afin d'utiliser, éventuellement, le signal reconstitué dans la reconnaissance automatique du locuteur. Nous exprimons en premier lieu l'effet de la taille des segments ( $TS$ ) et le nombre de pôles ( $M$ ) sur la résolution du signal reconstitué pour les différents cas. En second lieu, nous allons comparer les résultats obtenus en variant les paramètres de la matrice de Pencil (fréquence fondamentale ( $f_0$ ), l'amplitude ( $R_0$ ), le coefficient d'atténuation ( $\alpha$ ) et la phase ( $\Phi$ )) afin de montrer la robustesse contre le bruit.

### 3.2 Base de données utilisée

La base de données que nous avons utilisée pour évaluer expérimentalement les approches proposées, est extraite de la base de données "Hub-4 Broadcast-News-96". Cette dernière contient des enregistrements des informations télédiffusées en 1996 et couvre au total, 104 heures recueillies des chaînes de télévision : ABC (*American Broadcasting Company*), CNN (*Cable News Network*), et des chaînes de Radio : NPR (*National Public Radio*) / PRI (*Public Radio International*) .

Elle contient des informations et des interviews enregistrées de la chaîne CNN. Cet enregistrement réunit 15 locuteurs différents (9 males et 6 femelles), dont la plupart d'entre eux sont des journalistes avec de nombreux changements de ces intervenants tout au long de ce fichier, organisés en 259 combinaisons (intra et interlocuteur) pour la phase d'apprentissage et 195 combinaisons (intra et interlocuteur) pour le test.

La base de données Hub-4, étant très diversifiée puisqu'elle contient de la parole, de la musique, du bruit, des silences et d'autres sons, représente bien l'aspect naturel des discussions en mode multi-locuteur, dans une conférence (ou téléconférence) réelle. Il est à noter que les segments de parole utilisés pendant la phase d'apprentissage, ayant des durées moyennes de 4s, sont différents de ceux utilisés pendant la phase de test qui ont des durées moyennes de 3s, et que la fréquence d'échantillonnage,  $F_e$  est égale à 16 kHz.

### 3.3 Méthodologie et protocole expérimental

Ce travail expérimental est organisé en deux séries d'expériences et chaque série comporte quatre expériences (selon le cas d'application). La première série, étudie l'influence de la taille des segments et du nombre de pôles sur la résolution du signal de parole reconstitué. La deuxième série d'expériences étudie l'effet de la variation des différents paramètres de la matrice de Pencil et faire la comparaison entre les signaux reconstitués pour chaque cas d'application.

#### 3.3.1 Influence de la taille des segments et du nombre de pôle sur la résolution

Cette série d'expérience vise à déterminer l'influence de la taille des segments ( $TS$ ) et du nombre de pôles ( $M$ ) sur la résolution du signal de parole reconstitué. Les quatre cas d'applications pour cette série d'expérience sont :

1. Expérience N°1 : Cas d'un signal de parole saint d'un homme
2. Expérience N°2 : Cas d'un signal de parole saint d'une femme
3. Expérience N°3 : Cas d'un signal de parole bruité d'un homme
4. Expérience N°4 : Cas d'un signal de parole bruité d'une femme

Après investigation, nous avons choisis d'utiliser les paramètres suivant pour toutes les expériences de la série :  $M= 7, 13, 30, \text{ et } 50$  et  $TS= 100 \text{ et } 150$ . De plus, on a adopté la notation suivante :

- ❖  $MR$  (Mauvaise Résolution),
- ❖  $BR$  (Bonne Résolution),
- ❖  $RA$  (Résolution Acceptable).

Les résultats obtenus de cette série d'expériences dans les différents cas d'applications sont présentés dans les tableaux et les figures suivants :

## 3.3.1.1 Cas d'application : Signal de parole saint d'un homme

Tableau-3.1 : Résultats obtenus pour un signal de parole saint d'un homme

	$M=7$	$M=13$	$M=30$	$M=50$
$TS=100$	MR	MR	RA	-
$TS=150$	MR	-	RA	BR

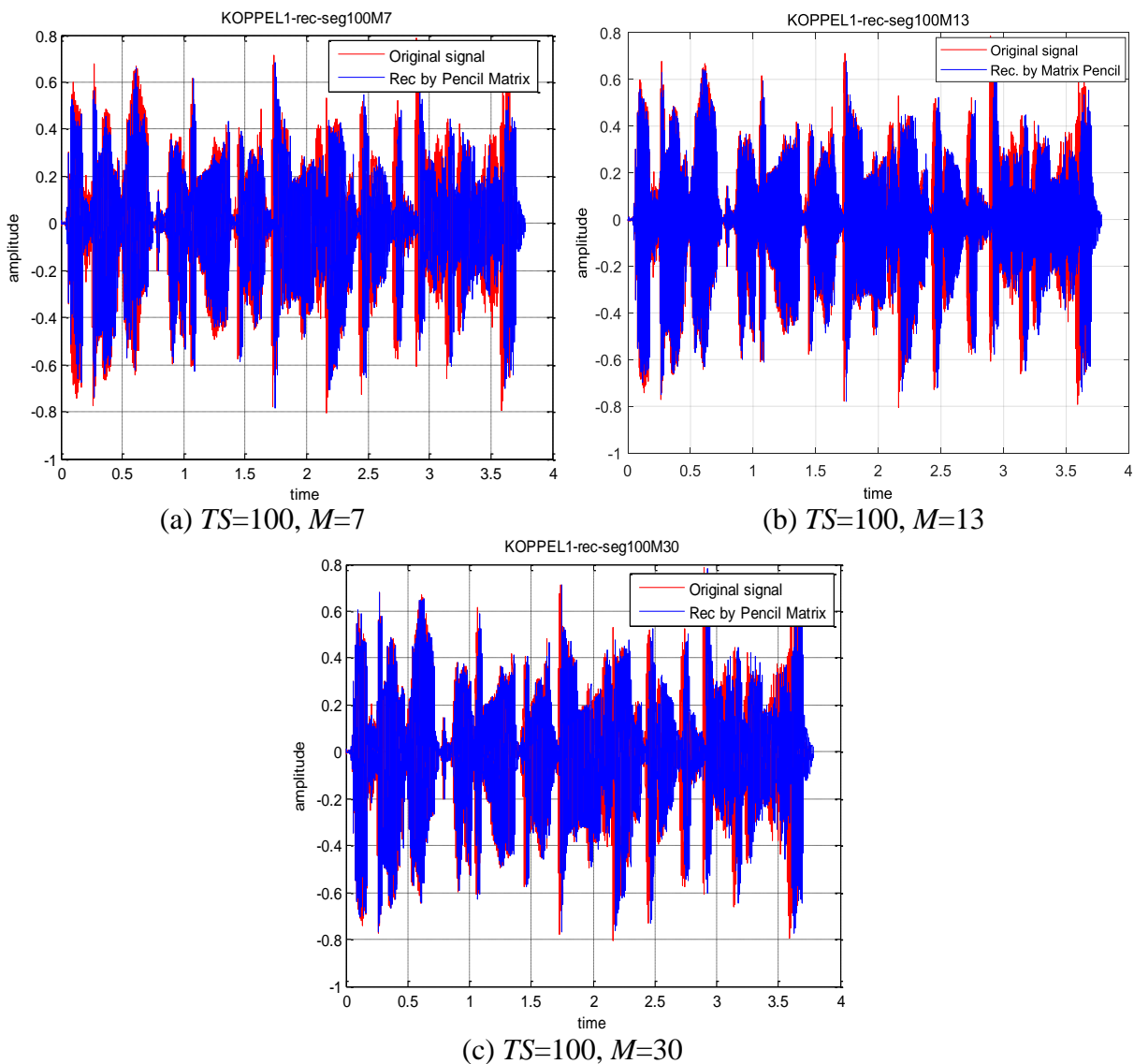


Figure-3.1 : Signal de parole original et reconstitué par la matrice de Pencil  
(Cas d'application : Signal saint d'un homme)

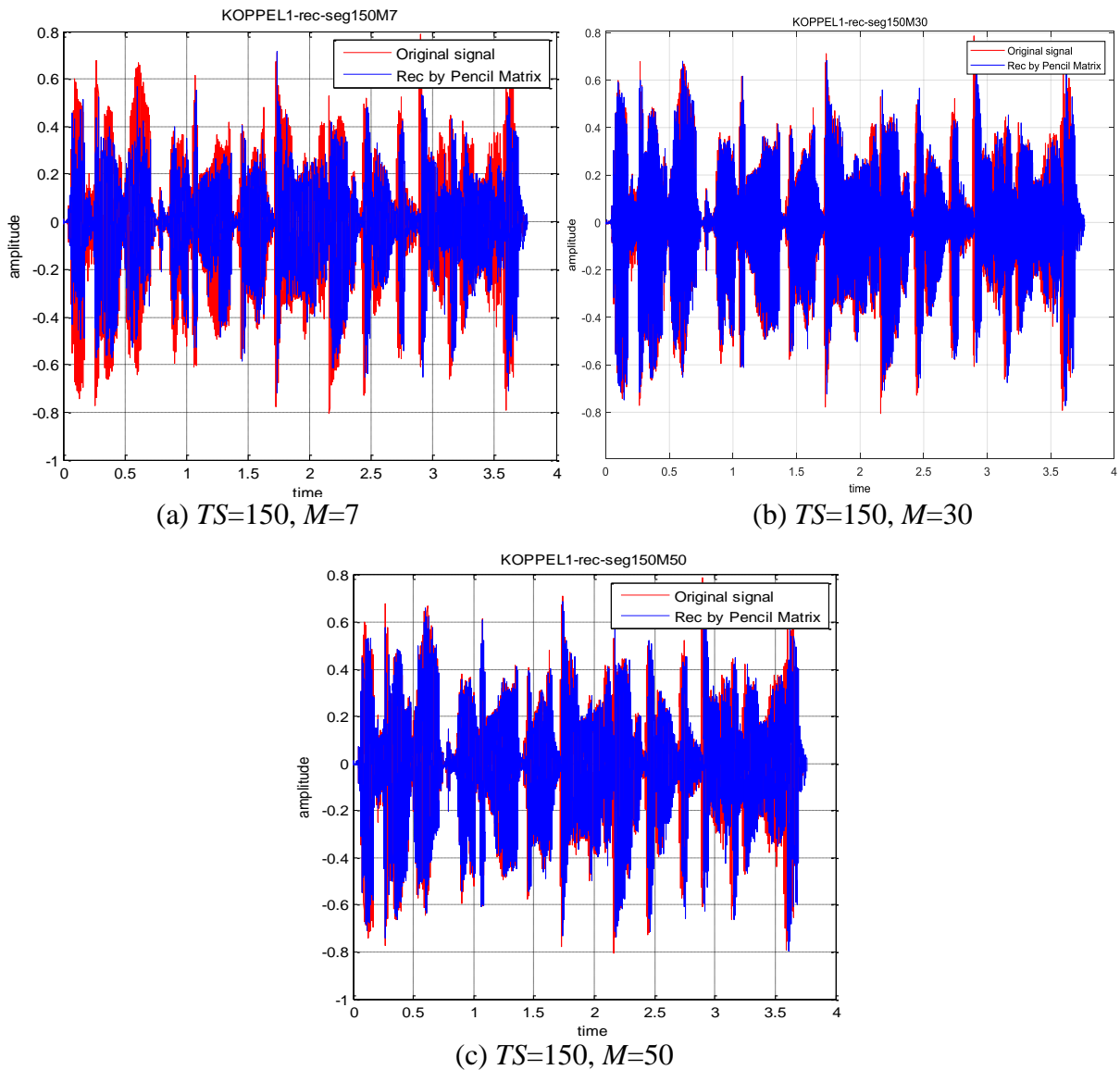


Figure-3.2 : Signal de parole original et reconstitué par la matrice de Pencil  
(Cas d'application : Signal saint d'un homme)

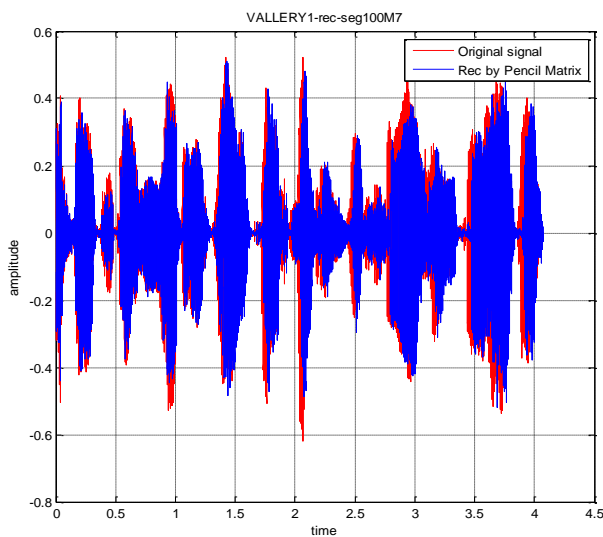
En examinant les résultats précédents, la couleur rouge représente le signal de parole original et la couleur bleu représente le signal de parole reconstitué par la matrice de Pencil, on remarque les observations suivantes :

- Pour les valeurs ;  $TS=100$  (échantillons) et  $M=7$  et  $10$ , on peut voir qu'il y a aucune superposition entre le signal original (couleur rouge) et le signal reconstitué (couleur bleu), cela se traduit par une mauvaise qualité à l'écoute du signal de parole reconstitué (tableau-3.1).
- Même remarque pour les valeurs de  $TS =150$  (échantillons) et  $M=7$ .
- Pour les valeurs ;  $TS=150$  (échantillons) et  $M=30$  on peut constater clairement l'existence d'une superposition quasi-totale entre le signal original et le signal reconstitué, cela se traduit par une bonne qualité à l'écoute du signal de parole reconstitué (tableau-3.1).

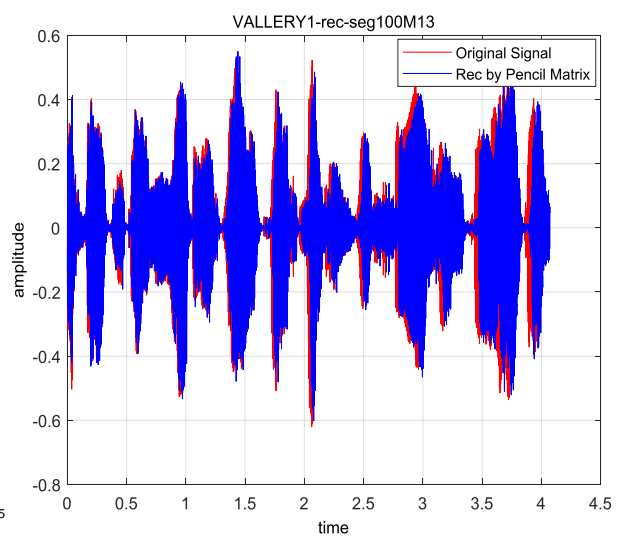
3.3.1.2 Cas d'application : Signal de parole saint d'une femme

Tableau-3.2 : Résultats obtenus pour un signal de parole saint d'une femme

	$M=7$	$M=13$	$M=30$	$M=50$
$TS =100$	$MR$	$MR$	$RA$	-
$TS =150$	$MR$	-	$RA$	$BR$



(a)  $TS=100, M=7$



(b)  $TS=100, M=13$

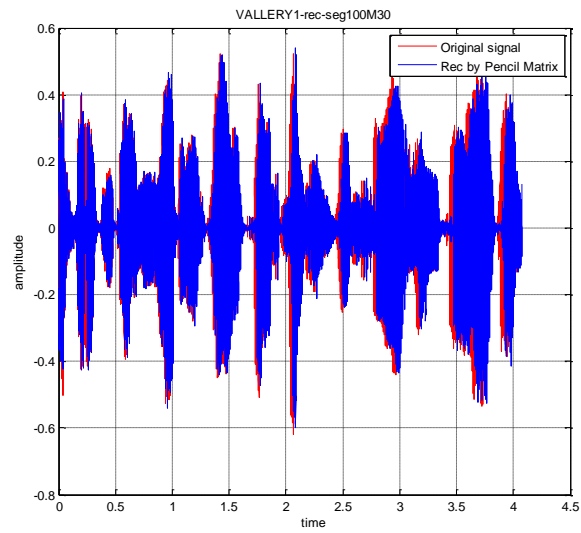
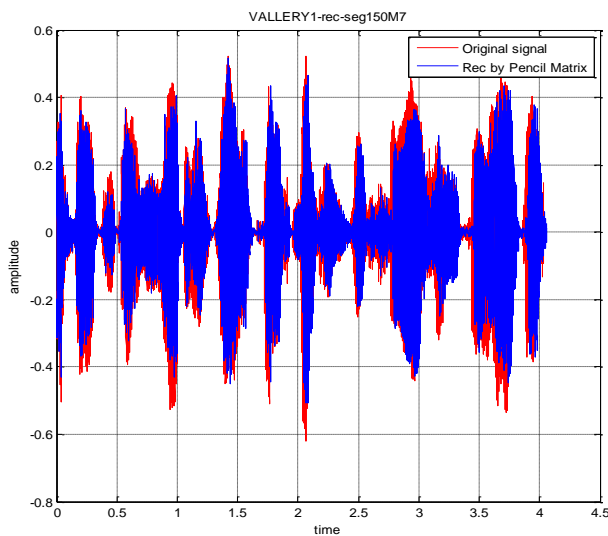
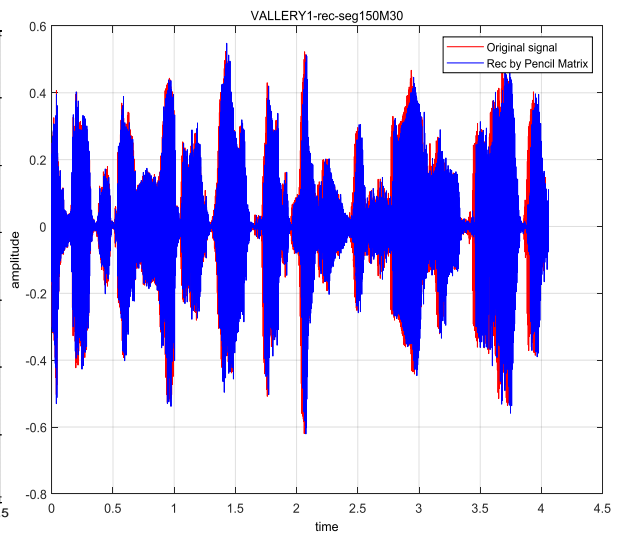
(c)  $TS=100, M=30$ 

Figure-3.3 : Signal de parole original et reconstitué par la matrice de Pencil  
(Cas d'application : Signal saint d'une femme)

(a)  $TS=150, M=7$ (b)  $TS=150, M=30$

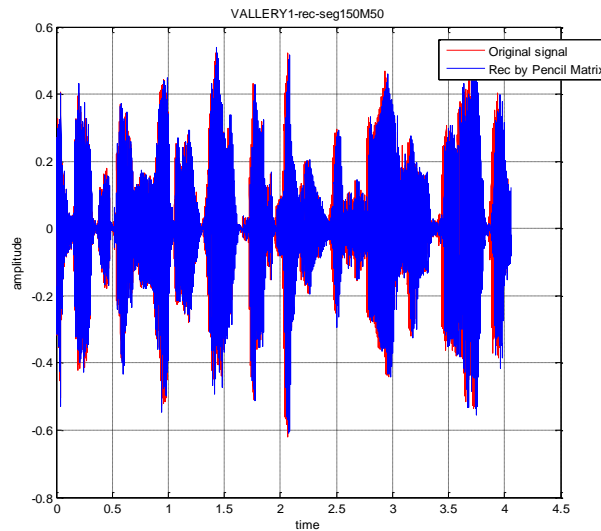
(c)  $TS=150, M=50$ 

Figure-3.4 : Signal de parole original et reconstitué par la matrice de Pencil  
(Cas d'application : Signal saint d'une femme)

Dans le cas d'un signal de parole saint d'une femme et d'après les résultats obtenus, on note les remarques suivantes :

- Pour les valeurs ;  $TS = 100$  (échantillons) et  $M=7$  et  $10$ , on peut voir qu'il y a aucune superposition entre le signal original (couleur rouge) et le signal reconstitué (couleur bleu), cela se traduit par une mauvaise qualité de son pour le signal reconstitué (tableau-3.2).
- Même remarque pour les valeurs de  $TS = 150$  échantillons et  $M=10$ .
- Or, à partir de  $M=30$  ( $TS=150$ ) on peut constater clairement une superposition quasi-totale entre le signal original et le signal reconstitué, cela se traduit par une bonne qualité de son de parole reconstitué (tableau-3.2).

### 3.3.1.3 Cas d'application : Signal de parole bruité d'un homme

Tableau-3.3 : Résultats obtenus pour un signal de parole bruité d'un homme

	$M=7$	$M=13$	$M=30$	$M=50$
$TS = 100$	$MR$	$MR$	$RA$	-
$TS = 150$	$MR$	-	$RA$	$BR$

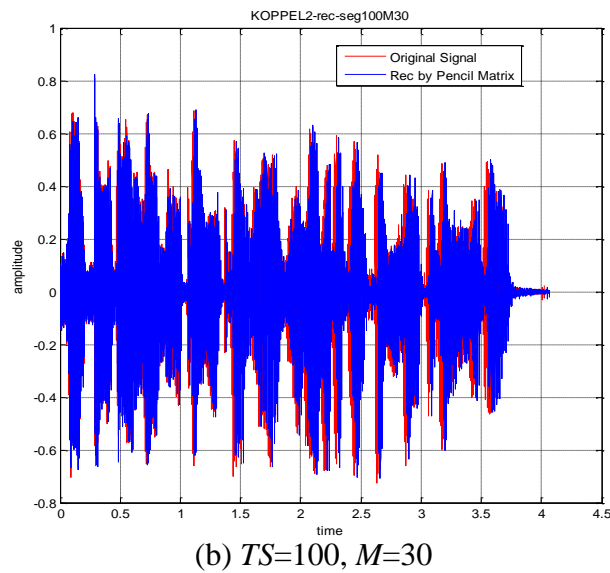
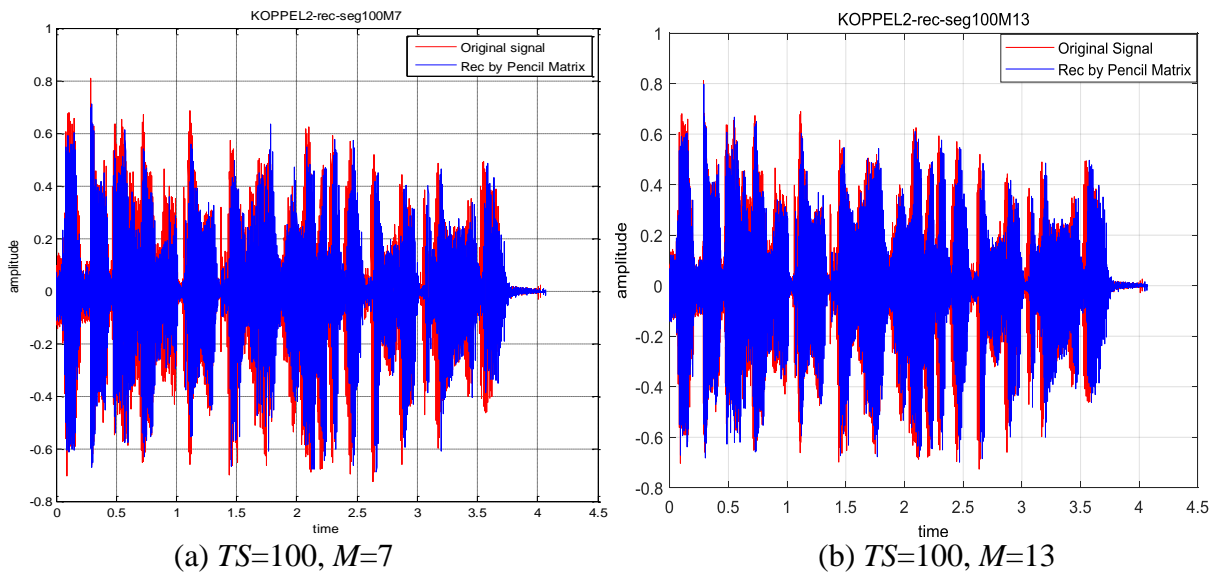


Figure-3.5. Signal de parole original et reconstitué par la matrice de Pencil  
(Cas d'application : Signal bruité d'un homme)

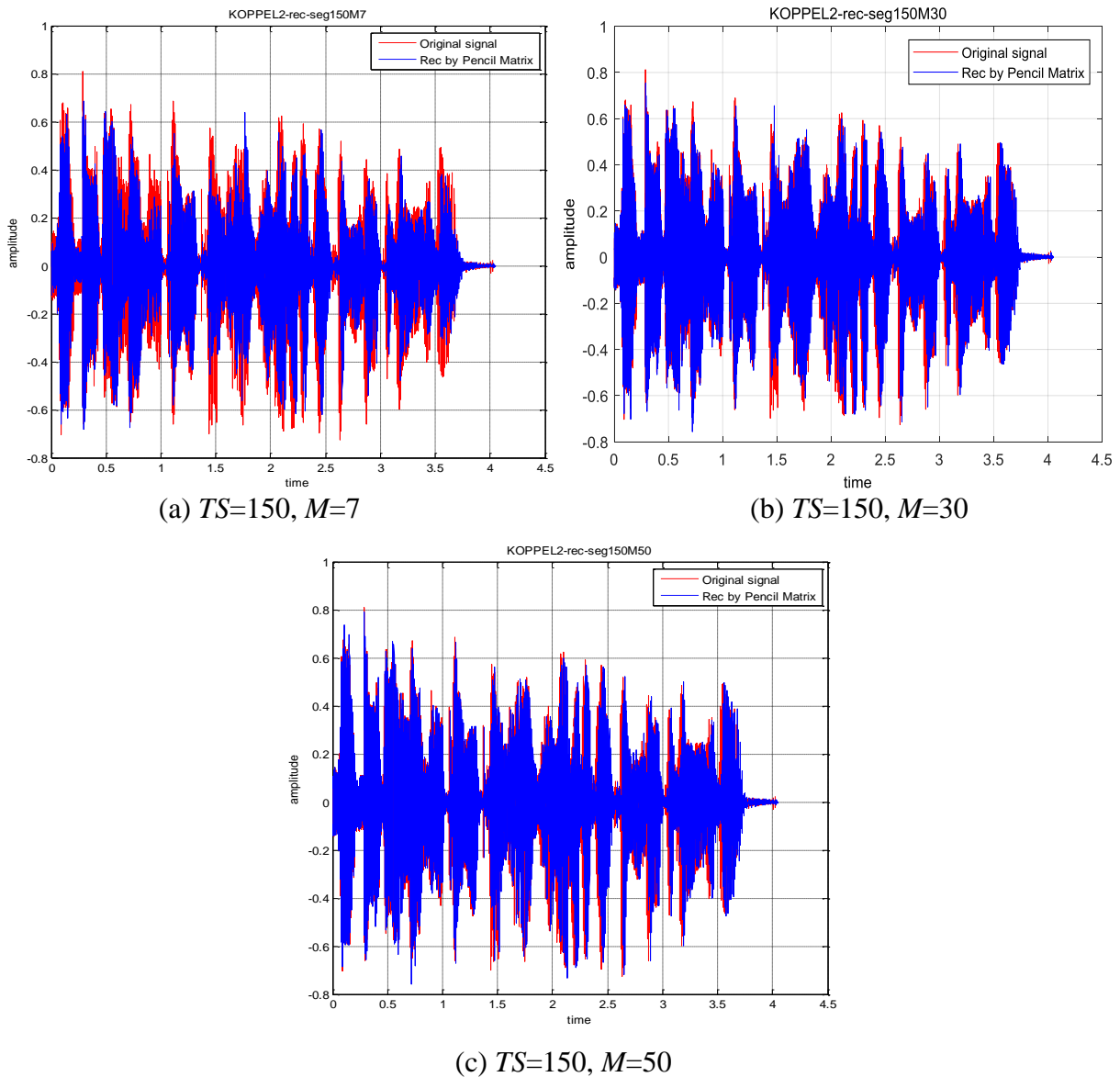


Figure-3.6 : Signal de parole original et reconstitué par la matrice de Pencil  
(Cas d'application : Signal bruité d'un homme)

Dans le cas d'un signal de parole bruité d'un homme et d'après les résultats obtenus, on note les remarques suivantes :

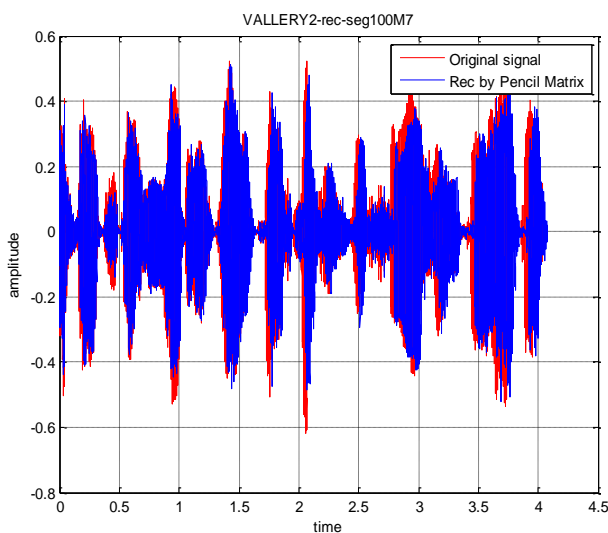
- Pour les valeurs ;  $TS = 100$  (échantillons) et  $M=7$  et  $10$ , on peut voir qu'il y a aucune superposition entre le signal original (couleur rouge) et le signal reconstitué (couleur bleu), cela se traduit par une mauvaise qualité de son pour le signal reconstitué (tableau-3.3).

- Même remarque pour les valeurs de  $TS = 150$  échantillons et  $M = 10$ .
- Or, à partir de  $M = 30$  ( $TS = 150$ ) on peut constater clairement une superposition quasi-totale entre le signal original et le signal reconstitué, cela se traduit par une bonne qualité de son de parole reconstitué (tableau-3.3).

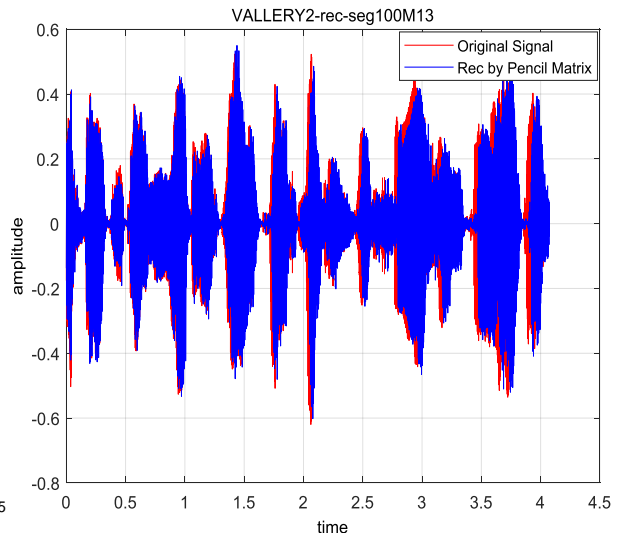
### 3.3.1.4 Expérience N°4 : Signal de parole bruité d'une femme

Tableau-3.4 : Résultats obtenus pour un signal de parole bruité d'une femme

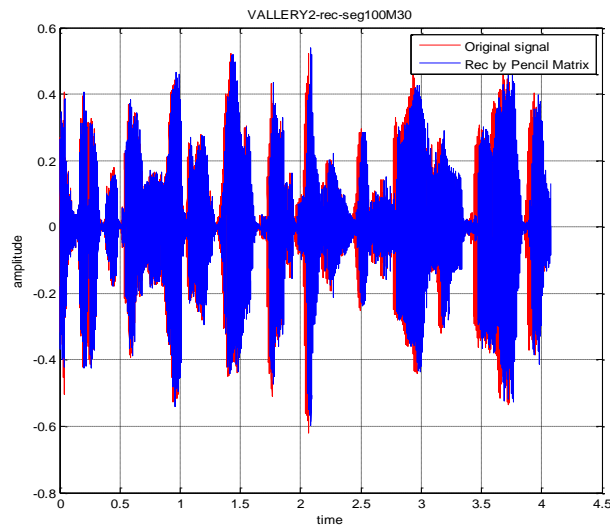
	$M=7$	$M=13$	$M=30$	$M=50$
$TS = 100$	$MR$	$MR$	$RA$	-
$TS = 150$	$MR$	-	$RA$	$BR$



(a)  $TS=100, M=7$

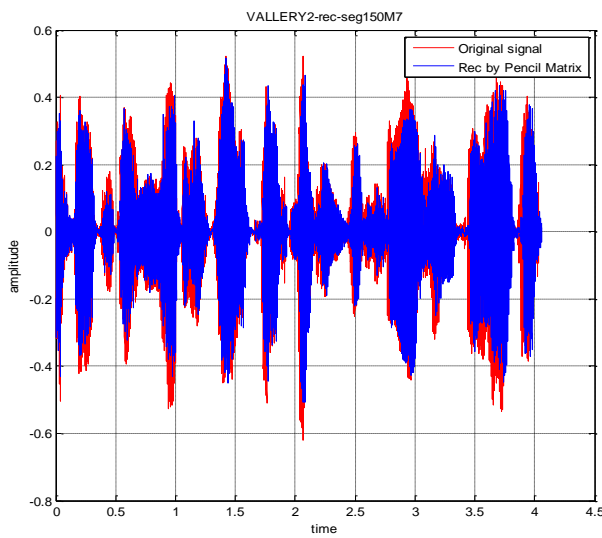


(b)  $TS=100, M=13$

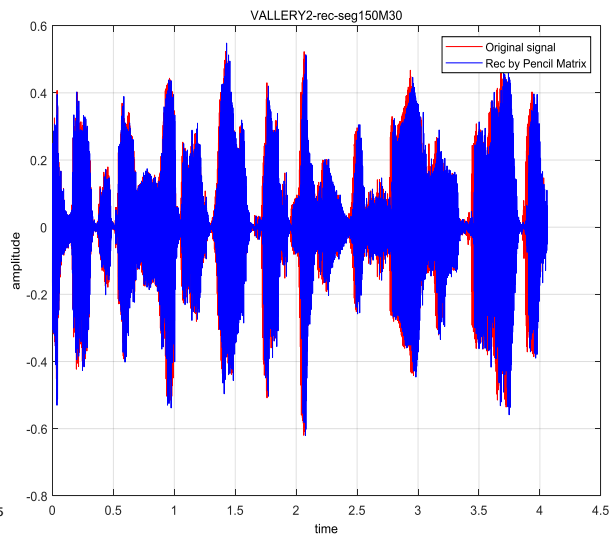


(c)  $TS=100, M=30$

Figure-3.7. Signal de parole original et reconstitué par la matrice de Pencil  
(Cas d'application : Signal bruité d'une femme)



(a)  $TS=150, M=7$



(b)  $TS=150, M=30$

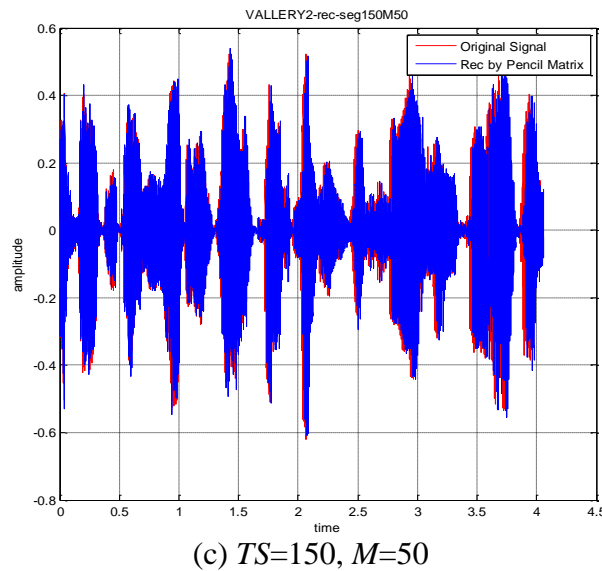


Figure-3.8 : Signal de parole original et reconstitué par la matrice de Pencil  
(Cas d'application : Signal bruité d'une femme)

Dans le cas d'un signal de parole bruité d'une femme et d'après les résultats obtenus, on note les remarques suivantes :

- Pour les valeurs ;  $TS = 100$  (échantillons) et  $M=7$  et  $10$ , on peut voir qu'il y a aucune superposition entre le signal original (couleur rouge) et le signal reconstitué (couleur bleu), cela se traduit par une mauvaise qualité de son pour le signal reconstitué (tableau-3.4).
- Même remarque pour les valeurs de  $TS = 150$  échantillons et  $M=10$ .
- Or, à partir de  $M=30$  ( $TS=150$ ) on peut constater clairement une superposition quasi-totale entre le signal original et le signal reconstitué, cela se traduit par une bonne qualité de son de parole reconstitué (tableau-3.4).

### 3.3.2 Comparaison des résultats en variant les paramètres de la matrice de Pencil

Dans cette deuxième série d'expérience nous allons utiliser deux fichiers de parole pour chaque locuteur homme (saint et bruité) et deux fichiers de parole pour chaque locuteur femme (saint et bruité), et on va les analyser en variant les paramètres de la matrice de Pencil qui sont ; la fréquence fondamentale ( $f_0$ ), la phase ( $\Phi$ ), le coefficient d'atténuation ( $\alpha$ ) et l'amplitude ( $R_0$ ), puis les comparer du point de vue robustesse contre le bruit.

### 3.3.2.1 Expérience N°5 : En utilisant la fréquence fondamentale ( $f_0$ )

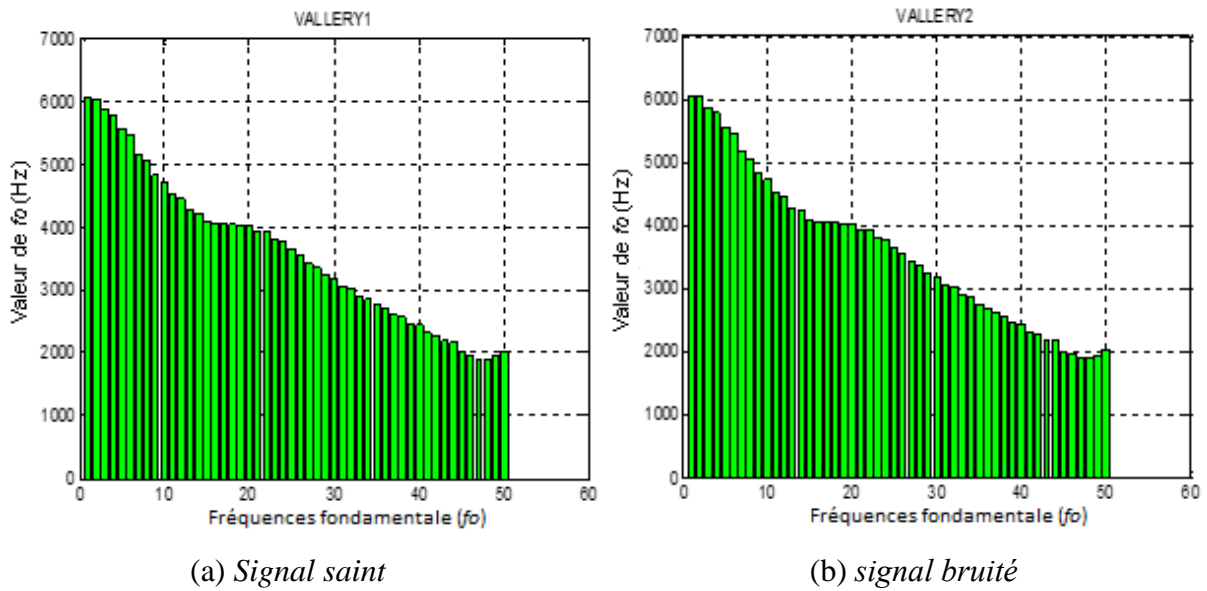


Figure-3.9. Comparaison des résultats obtenus en utilisant la fréquence fondamentale ( $f_0$ )  
(Cas d'application : Signal de parole (saint et bruité) d'une femme)

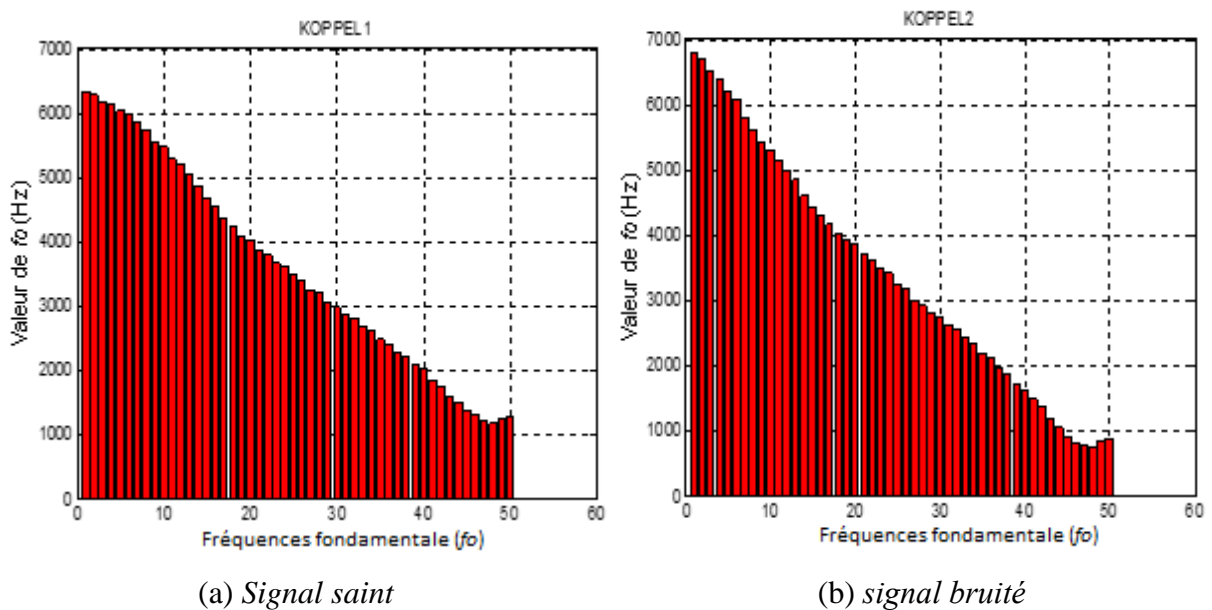


Figure-3.10. Comparaison des résultats obtenus en utilisant la fréquence fondamentale ( $f_0$ )  
(Cas d'application : Signal de parole (saint et bruité) d'un homme)

Les résultats illustrés dans la figure-3.9 et la figure-3.10, nous laisse constater que dans les deux figures-3.9 (a et b) (cas de signaux saint et bruité d'une femme) les deux courbes sont pratiquement identiques. Or pour les deux figures-3.10 (a et b) (cas de signaux saint et bruité pour un homme), il y a une diminution progressive de l'amplitude en termes de fréquence, et on remarque des ressemblances entre les deux formes de même locuteur (femme vs femme) et (homme vs homme). Ceci peut être interpréter par une grande robustesse contre le bruit.

### 3.3.2.2 Expérience N°6 : En utilisant la phase ( $\Phi$ )

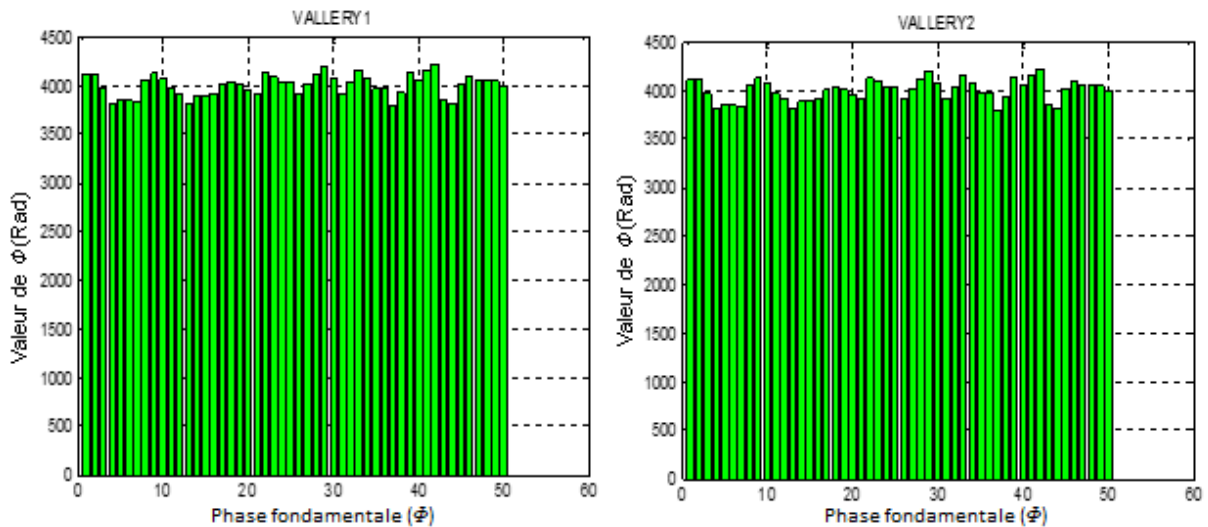


Figure-3.11. Comparaison des résultats obtenus en utilisant en utilisant la phase ( $\Phi$ )  
(Cas d'application : Signal de parole (saint et bruité) d'une femme)

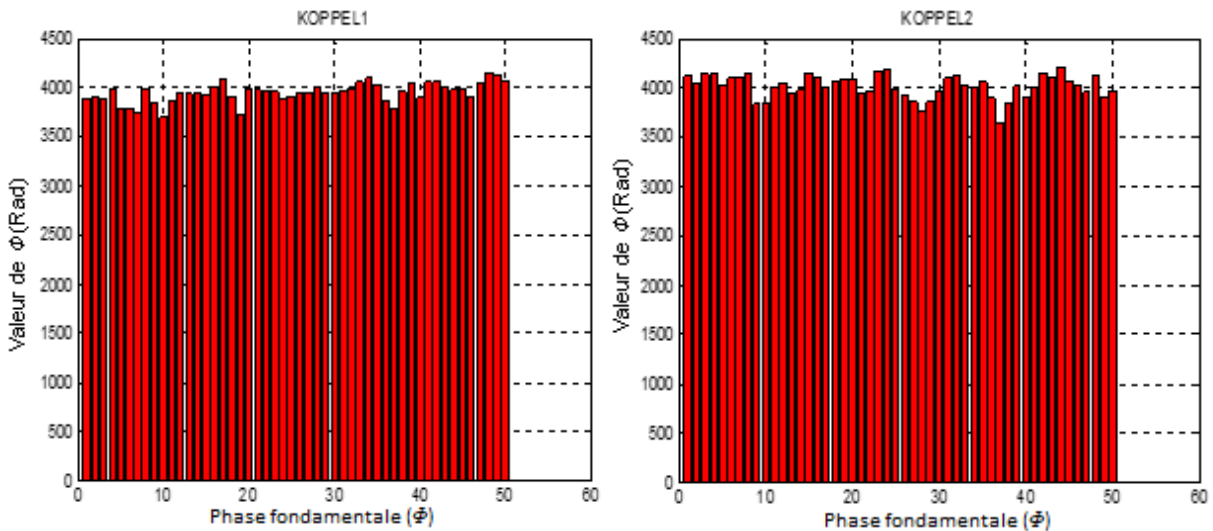


Figure-3.12. Comparaison des résultats obtenus en utilisant en utilisant la phase ( $\Phi$ )  
(Cas d'application : Signal de parole (saint et bruité) d'un homme)

Les résultats illustrés dans la figure-3.11 et la figure-3.12, nous laisse constater que dans les deux figures-3.11 (a et b) (cas de signaux saint et bruité d'une femme) les courbes sont quasiment identiques. Or pour les deux figures-3.12 (a et b) (cas de signaux saint et bruité pour un homme), il y a une légère différence entre le signal saint et le signal bruité, et on remarque des similitudes entre les deux figures de même locuteur (femme vs femme) et (homme vs homme). Ceci peut être interpréter par une grande virilité contre le bruit.

### 3.3.2.3 Expérience N°7 : En utilisant le coefficient d'atténuation ( $\alpha$ )

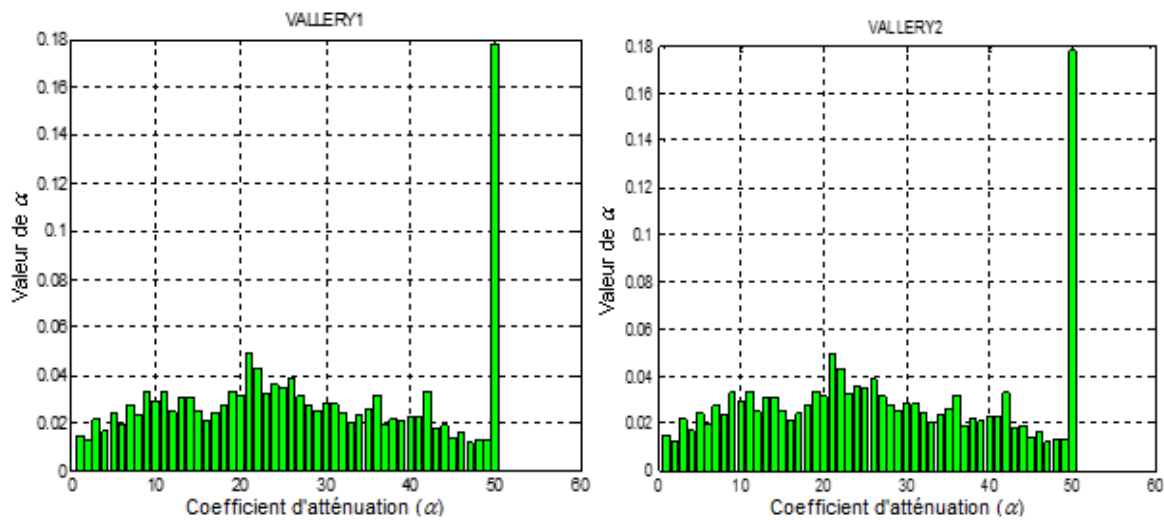


Figure-3.13. Comparaison des résultats obtenus en utilisant le coefficient d'atténuation ( $\alpha$ )  
(Cas d'application : Signal de parole (saint et bruité) d'une femme)

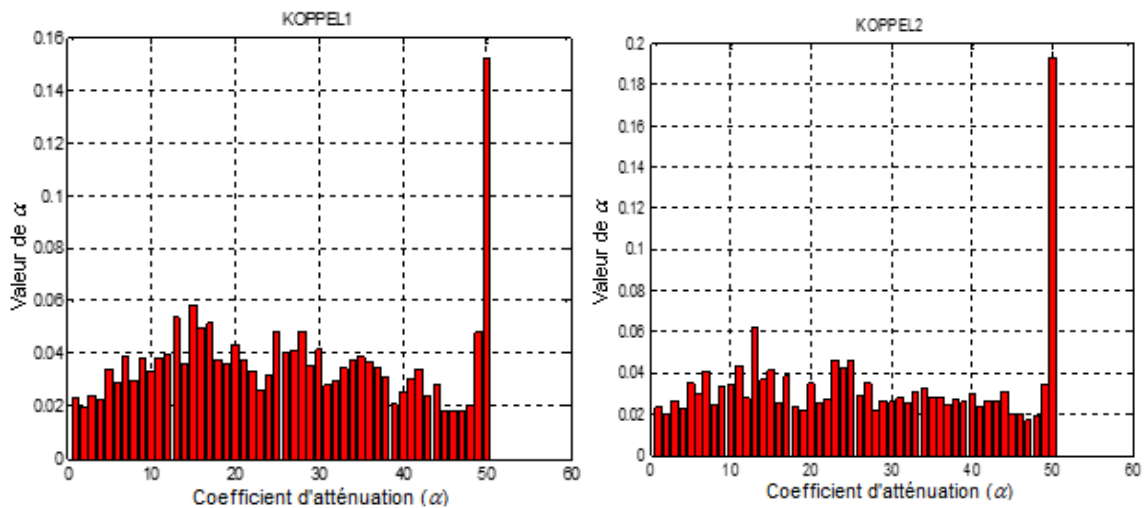


Figure-3.14. Comparaison des résultats obtenus en utilisant le coefficient d'atténuation ( $\alpha$ )  
(Cas d'application : Signal de parole (saint et bruité) d'un homme)

D'après les résultats montrés dans la figure-3.15 et la figure-3.16, on remarque les mêmes observations précédentes, que dans les deux figures-3.15 (a et b) (cas de signaux saint et bruité d'une femme) les courbes sont presque semblables. Or pour les deux figures-3.16 (a et b) (cas de signaux saint et bruité pour un homme), il y a une légère dissimilitude entre le signal saint et bruité, et on constate des similitudes entre les figures du même locuteur (femme vs femme) et (homme vs homme). Ceci peut être interpréter par une grande virilité contre le bruit.

### 3.3.2.4 Expérience N°8 : En utilisant l'amplitude ( $R_o$ )

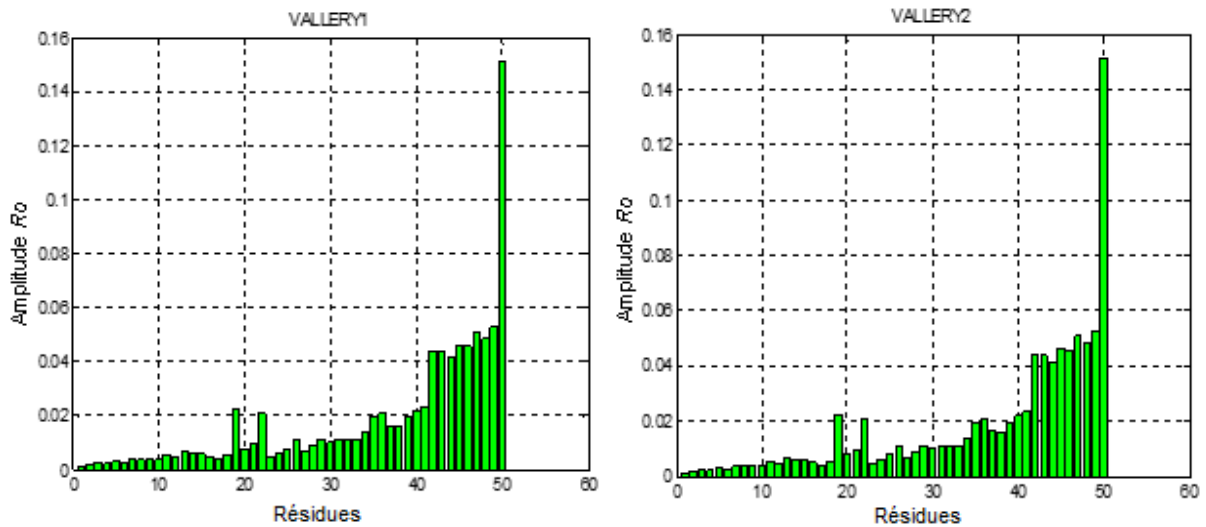


Figure-3.15. Comparaison des résultats obtenus en utilisant l'amplitude ( $R_o$ )  
(Cas d'application : Signal de parole (saint et bruité) d'une femme)

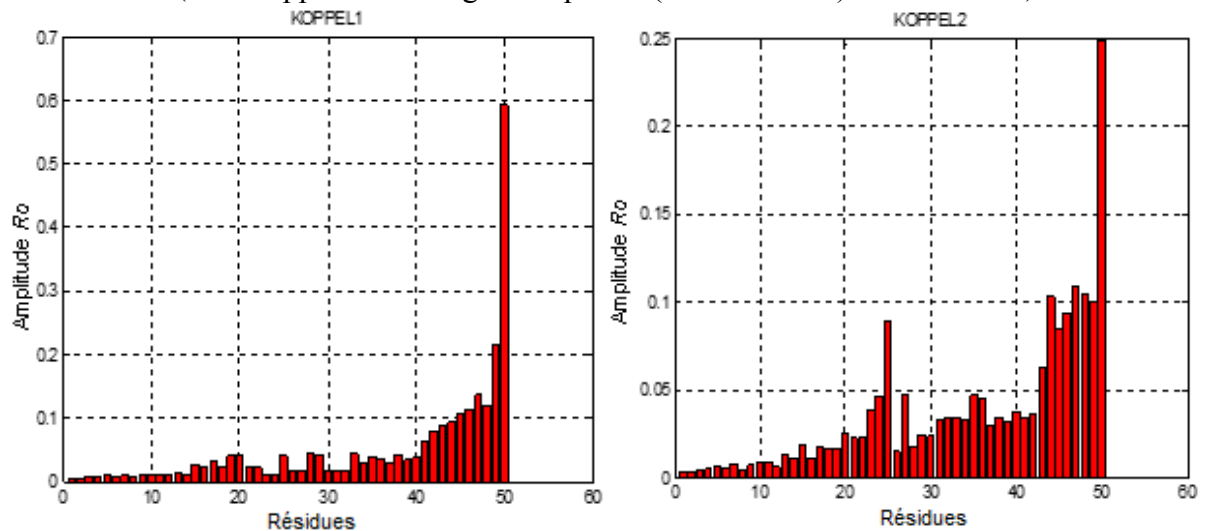
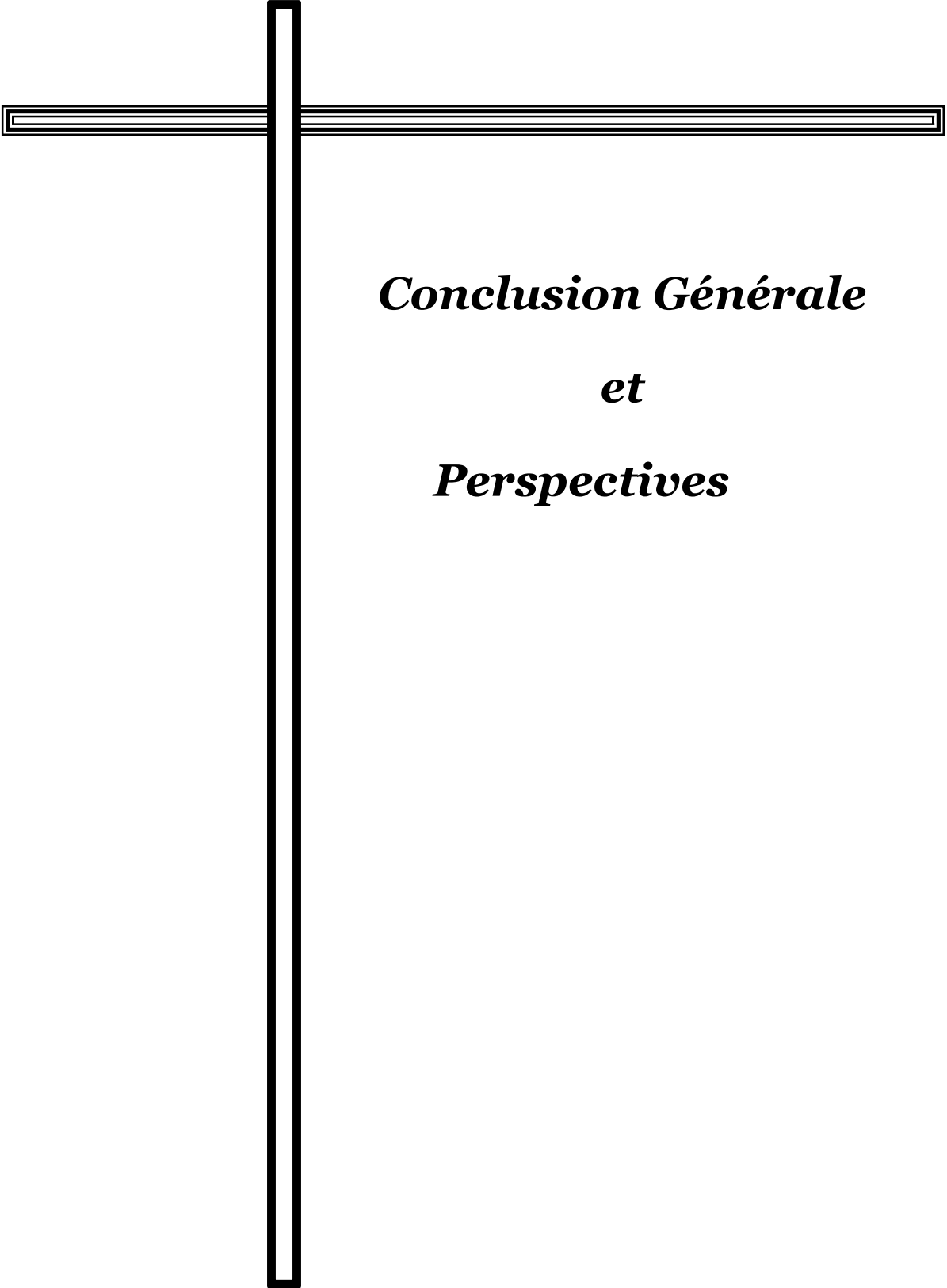


Figure-3.16. Comparaison des résultats obtenus en utilisant l'amplitude ( $R_o$ )  
(Cas d'application : Signal de parole (saint et bruité) d'un homme)

D'après les figures (3.15 et 3.16), on voit clairement que dans le premier cas (Figure-3.15 a et b) (cas des signaux saint et bruité d'une femme) les courbes sont très proche l'une de l'autre. Cependant, dans le deuxième cas (Figure-3.16 a et b) (cas des signaux saint et bruité pour un homme), faible différence est enregistrée entre le signal saint et le signal bruité. De même que les explications données auparavant.

### 3.4 Conclusion

Dans ce chapitre, nous avons exposés les différentes expériences de discrimination menées sur notre base de données, et nous avons détaillés et discuter les différents résultats de obtenus. Deux séries d'expériences ont été étudiées en utilisant les paramètres de la matrice de Pencil. Le but de ces expériences est d'analyser un signal de parole par la matrice de Pencil pour extraire les différentes caractéristiques acoustique, les résultats obtenus ont montré l'efficacité de cette approche.



***Conclusion Générale***  
***et***  
***Perspectives***

### **Conclusion générale et perspective**

L'algorithme MP est un algorithme préexistant utile pour la décomposition des signaux ESM. Il présente des avantages à la fois dans les applications de recherche directionnelle et, comme dans cette thèse, dans les applications de traitement de la parole. Il a été montré que l'algorithme MP peut être appliqué au traitement de la parole, à l'amélioration de la parole et à l'estimation de la hauteur tonale.

Dans cette thèse, la théorie mathématique de l'algorithme MP a été décrite (Au chapitre 2), il a été montré et la manière dont l'approche MP est appliquée à un signal de parole. L'estimation de l'erreur de reconstruction de la parole à l'aide de l'algorithme MP a également été introduite. Au chapitre 3, Le programme de l'algorithme MP ainsi modifié a été appliqué En deux parties à savoir :

- Influence de la taille des segments et du nombre de pôle sur la résolution.
- Comparaison des résultats en variant les paramètres de la matrice de Pencil.

La méthode matrice-crayon a été appliquée avec succès à l'analyse des sons et à l'identification des courbes d'admittance. En comparaison avec d'autres techniques, la méthode matrice-crayon apparaît plus robuste, plus facile à utiliser et permet d'obtenir les mêmes performances en termes de résolution spectrale. L'algorithme a été implémenté sous MATLAB en raison des caractéristiques uniques et puissantes de ce dernier et qui ne sont disponibles dans aucun autre langage de programmation. À travers les résultats de nos expériences la méthode matrice-crayon apparaît plus robuste, plus facile à utiliser et permet d'obtenir les mêmes performances en termes de résolution spectrale. La méthode du crayon matriciel a été appliquée avec succès à l'analyse des sons percussifs et à l'identification des courbes d'admittance et nous avons exposés les différentes expériences de discrimination menées sur notre base de données, et nous avons détaillés et discuter les différents résultats de obtenus. Deux séries d'expériences ont été étudiées en utilisant les paramètres de la matrice de Pencil. Le but de ces expériences est d'analyser un signal de parole par la matrice de Pencil pour extraire les différentes caractéristiques acoustique, les résultats obtenus ont montré l'efficacité de cette approche.

## *Conclusion et perspective*

Les avantages de l'algorithme MP par rapport aux autres techniques d'estimation comprennent la super résolution du spectre, la faible variance des paramètres se rapprochant et le fait qu'il s'agit d'un processus non stochastique qui utilise une approche de données directes pour obtenir le modèle. Paramètres. L'inconvénient de l'algorithme MP est qu'il s'agit d'un processus de calcul chronophage.

Bonnes performances sur des fenêtres d'analyse courtes : La puissance de la méthode matrice- crayon est plus remarquable lorsque la longueur de l'enregistrement de données est petite. Dans ce cas, les méthodes basées sur la transformée de Fourier et la sélection de pics ne fonctionnent pas bien en raison de la petite taille de la fenêtre d'analyse, et des schémas plus complexes sont nécessaires, comme décrit ci-dessus. Paramètres sinusoïdaux intégrés : un autre avantage de la méthode matrice-crayon par rapport aux méthodes basées sur la FFT est sa capacité inhérente à se concentrer sur les composants importants du signal et à extraire directement les paramètres d'intérêt (fréquence, facteur d'amortissement, amplitude et phase).

Enfin, l'algorithme MP doit faire l'objet d'avantage de réflexion afin d'en améliorer les performances. A titre d'exemple on cite : D'augmenter le nombre d'exemples (combinaisons inter et intra-locuteur) pour l'apprentissage .l'hybridation de l'algorithme ; modification des paramètres de l'algorithme ; il est aussi intéressant de modifier l'algorithme pour qu'il puisse traiter des problèmes multi-objectifs. D'utiliser d'autres classifieurs basés sur l'intelligence artificielle tel que ; les réseaux de neurones, la logique floue, etc..., pour une éventuelle comparaison et / ou fusion.

## Références Bibliographiques

---

- [1] G. Fant, 1960. Acoustic Theory of Speech Production. Hague : Mouton's Co.
- [2] R. Boite. Traitement de la parole. Collection Electricité. Presses Polytechniques et Universitaires Romandes, 2000.
- [3] <https://theses.univ-oran1.dz/document/TH2426.pdf>
- [4] Mr. Abdelghani HARRAG, Extraction des données d'une base : Application à l'extraction des traits du locuteur, thèse de doctorat à Département Electronique Université Ferhat ABBAS – Sétif, 2011.
- [5] Alexandre. PRETI. Surveillance de réseaux professionnels de communication par la reconnaissance du locuteur. , thèse de doctorat de l'école Doctorale 166 I2S, Décembre 2008.
- [6] Jain et al, an introduction to biometric recognition: Mouton's Co, IEEE Security Privacy Mag., vol. 14, no. 2, pp. 4 - 20, 2004.
- [7] Sondes Abdelmouleh, (2011), « traitement du signal ». Repéré à <http://www.technologuepro.com/cours-genie-electrique/cours-14-traitement-signal/>
- [8] Christian JUTTEN, (2009), « théorie du signal », Université Joseph Fourier Polytech'Grenoble
- [9] J.P. DELMAS, Éléments de théorie du signal : les signaux déterministes, Ellipses, Paris, 1991.
- [10] J.N. Rayner, Spectral Analysis, The Ohio State University, Columbus, USA, 2001, Pages 14861-14864
- [11] Alain DUCASSE, Corinne MAILHES, Francis CASTANI E, " Estimation de fréquences : panorama des méthodes paramétriques" in Traitement du Signal 1998 – Volume 15 - n°2.
- [12] <http://www.i3s.unice.fr/~rendas/esinsa/TEST>.
- [13] Jean Laroche, "The use of the matrix pencil method for the spectrum analysis of musical signals "Spectrum Analysis of musical signals, Vol. 94, No. 4, October 1993.

- [14] Khakim Ikramov, " Matrix pencils: Theory, applications, and numerical methods " in Journal of Soviet Mathematics · April 1993.
- [15] Michael L. Van Blaricum AND Raj Mittra, " A Technique for Extracting the Poles and Residues of a System Directly from Its Transient Response " e transactions on antennas an propagation, vol. ap-23, no. 6, November 1975.

## الملخص:

يتكون تمييز المتحدثين من تمييز كلام متحدث أو أكثر في الملفات الصوتية، من أجل اكتشاف لحظة تغيرات المتحدث وكذا عدد المتحدثين. يمكن أن يحتوي هذا المجال على تطبيق فعال في علم الجريمة أو أو الوسائط المتعددة. يمكن تنفيذ هذه المهمة باستخدام طريقة مصفوفة Pencil للمعالجة لتحليل وتوليف إشارة الكلام لكل متحدث. طريقة مصفوفة Pencil هي إحدى طرق التحليل الطيفي عالية الدقة التي تم تطويرها بواسطة Hua و Sarkar في عام 1989. تعتمد طريقة مصفوفة Pencil على افتراض أن الإشارة تتكون من مجموع أشباه الجيوب، ويعطي تقديرات لتردداتها وعوامل التخميد (أو عرض النطاق الترددي). في هذه المذكرة، قمنا بدراسة مجموعتين من التجارب باستخدام محددات MFSC. الهدف من هذه التجارب هو تحليل إشارة الكلام بواسطة مصفوفة Pencil لاستخراج الخصائص الصوتية المختلفة.

## Résumé

La discrimination de locuteur consiste à distinguer la parole de deux locuteurs ou plus dans un fichier audio, afin de détecter les instants de changements de locuteurs et le nombre de locuteurs. Ce domaine peut avoir une application Criminologie, sécuritaires, ou multimédia. Cette tâche peut être effectuée en utilisant la méthode de la matrice de Pencil pour faire l'analyse et la synthèse du signal de parole d'un locuteur. Cette méthode est l'une de plusieurs méthodes d'analyse spectrale à haute résolution qui ont été développées par Hua et Sarkar en 1989. Elle repose sur l'hypothèse que le signal est constitué d'une somme de décroissance exponentielle, sinusoïdes, et donne des estimations de leurs fréquences et facteurs d'amortissement. Dans ce mémoire, deux séries d'expériences seront réalisées en utilisant les paramètres de la matrice de Pencil. Le but de ces expériences est d'analyser un signal de parole par la matrice de Pencil pour extraire les différentes caractéristiques acoustique.

## Abstract

Speaker discrimination is the process of distinguishing the speech of two or more speakers in audio files, in order to detect the times of speaker changes as well as the number of speakers. This field can have a Criminology, security, or multimedia application. This task can be performed using the Pencil Matrix method to process each speaker's speech signal for analysis and synthesis. The Pencil Matrix Method is one of the many high-resolution spectral analysis methods that were developed by Hua and Sarkar in 1989. The Pencil Matrix Method is based on the assumption that the signal consists of a sum of exponential decay, sinusoids, and gives estimates of their frequencies and damping factors (or bandwidth). In this paper, two sets of experiments will be studied using the parameters of the Pencil matrix. The aim of these experiments is to analyze a speech signal through the Pencil matrix to extract the different acoustic characteristics.