الجمهورية الجزائرية الديمقراطية الشعبية

MINISTERE DE L'ENSEIGNEMENT SUPERIEUR
ET DE LA RECHERCHE SCIENTIFIQUE
UNIVERSITE DR. YAHIA FARES DE MEDEA

FACULTE DES SCIENCES ET DE LA TECHNOLOGIE
Département du Génie Électrique et de l'Informatique



Projet de Fin D'etude

Présenté pour l'obtention du Diplôme
de Master en Génie Électrique

Spécialité : **systèmes intelligents et communicants**

THEME

CONTRIBUTION A LA CONCEPTION D'UN CORRECTEUR AUTOMATIQUE DES REGLES DE RECITATION DU CORAN « TAJWID » (APPLICATION A L'IDGHAM)

Proposé et dirigé par,
M. ESSAID Billel

Préparés et présentés par,
Melle OULD RAMOUL Amina
Melle YAHIA HELALI Nouha

Dédicace

Je dédie ce modeste travail à:

Mes très chers parents: *vos prières et vos bénédictions m'ont été d'un grand secours pour mener à bien mes études. Aucune dédicace ne saurait être assez éloquente pour exprimer ce que vous méritez pour tous les sacrifices que vous n'avez pas cessé de me donner depuis ma naissance, durant mon enfance et même à l'âge adulte. Que Dieu, le tout puissant, vous préserve et vous accorde santé, longue vie et bonheur.*

A mon grand frère Hamza et sa future épouse.

A mes petits frères Abd El Raouf et Abd El Rahmen.

A ma grande sœur ***Baya*** *et son mari et leur fleur de Jasmin Yasmine.*

A ma petite adorable sœur ***Kaouther****.*

A tous les membres de ma famille, grands et petits.

A ma très chère partenaire Amina, vous étiez vraiment une formidable copine

A mes très chères amies Saliha, Meriem, Hadjer, Mounira Je vous dis : « Ouhiboukounna fi ***LLEH*** **».**

A mes très chères amies, Nawel, Amina, Kaouther, Kamilia, Soumia, Asma, Hanane.

Aux membres de ma promotion.

Nouha

Dédicace

Je dédie ce modeste travail à:

Mes très chers parents: vos prières et vos bénédictions m'ont été d'un grand secours pour mener à bien mes études. Aucune dédicace ne saurait être assez éloquente pour exprimer ce que vous méritez pour tous les sacrifices que vous n'avez pas cessé de me donner depuis ma naissance, durant mon enfance et même à l'âge adulte. Que Dieu, le tout puissant, vous préserve et vous accorde santé, longue vie et bonheur.

A mon fiancé Rachid et toute sa famille.

A mes frères Mohamed et Youcef.

A mes chères sœurs Yasmine et Maroua.

A tous les membres de ma famille, petits et grands

A ma très chère partenaire Nouha vous étiez vraiment une formidable copine.

Aux personnes qui m'ont toujours aidée et encouragée, qui étaient toujours à mes côtés, et qui m'ont accompagnaient durant mon chemin d'études supérieures, à mes aimables amis

Aux membres de ma promotion.

Amina

Remerciement

Tout d'abord,

*Louange à « **ALLAH** » qui nous a guidé sur le droit chemin tout au long du travail et nous a inspiré les bons pas et les justes reflexes. Sans sa miséricorde, ce travail n'aura pas abouti.*

*Nos remerciements s'adressent particulièrement au monsieur **ESSAID Billel**, promoteur de ce travail. Nous avons eu le privilège de travailler avec vous et d'apprécier vos conseils. Votre compétence et votre sens du devoir nous ont énormément marqués. Veuillez trouver ici l'expression de notre respectueuse considération.*

Nos sincères remerciements vont également aux membres de jury qui nous ont fait l'honneur de juger ce travail.

*Notre gratitude va aussi à l'ensemble des enseignants du département de génie électrique, pour leur encadrement universitaire en particulier Mr **SOLTANE Mohamed** qui nous a aidés dans notre travail.*

*Nous remercions aussi l'enseignante **YAHIA HELALI Baya** pour son considérable aide dès le début de notre travail.*

Nous remercions nos sœurs Naziha, Selma, Hadjer, Amina, Asma, Ahlem, Imen et Kenza qui nous ont aidés lors des enregistrements.

Que tous ceux qui nous ont aidés de près ou de loin, trouvent ici l'expression de notre gratitude.

A vous tous qui ont accepté d'endosser des irradiations pour l'accomplissement de ce travail, nous disons une fois de plus merci. Ce travail est aussi le votre.

Amina et Nouha **Juin 2014**

مُلخّص

يعتبر تعلم القرآن الكريم من أهم العلوم وأرفعها. فلابد لمتعلم القرآن من مراعاة أحكام الترتيل عند قراءة القرآن الكريم. ويأتي هذا البحث مدعماً لهذا العلم الرفيع. فمن خلاله قمنا بتطوير برنامج يمكن المستخدم من تعلم أحد أحكام النون الساكنة والتنوين وهو حكم الإدغام.

اعتمد البرنامج على لغة الماتلاب في تطبيقه ،حيث استخدمنا في المرحلة الأولى تقنية MFCC من اجل استخلاص السمات المميزة للصوت، أما في المرحلة الثانية فقد اعتمدنا على نماذج ماركوف الخفية HMM من اجل تدريب واختبار البرنامج بواسطة العينات الصوتية.

كلمات مفتاحية : *نماذج ماركوف الخفية، القرآن الكريم، التعرف الآلي على الكلام، استخلاص السيمات المميزة للصوت.*

Résumé

L'apprentissage du coran (tajwid) est considéré comme une science nécessaire et noble. Il *est* requis à celui qui apprend du Livre d'Allah, de veiller à respecter les règles de sa récitation. Ce travail de recherche soutient cette sublime science en développant un programme qui permet à l'utilisateur d'apprendre l'une des règles de « noun essakina et tanwin » et plus précisément la règle dite « el idgham ».

Ce programme s'est appuyé dans son application sur la langue de MATLAB. Où nous avons utilisé dans la première phase la technique MFCC pour extraire les caractéristiques des sons. Tandis que dans la deuxième phase, nous avons adopté les modèles cachés de Markov HMM afin d'entrainer et tester notre programme par le biais des échantillons vocaux.

Mots clés : *Les Chaines de Markov Cachées, Coran, Reconnaissance Automatique de la Parole, coefficients cepstraux MFCC.*

Abstract

Learning the Quran (Tajwid) is considered a necessary and noble science. It is required to he who learns the Book of Allah, to ensure compliance with the rules of the recitation. This research supports this sublime science by developing a program that allows the user to learn the rules of a "noun essakina and tanwin" and more specifically the rule known as "el Idgham."

This program is supported in its application to the MATLAB language. Where we used in the first phase of the MFCC technique to extract the characteristics of sounds. While in the second phase, we adopted the hidden Markov HMM to train and test our program through voice samples models.

Key words: *Hidden Makrov Models, Quran, Automatic Speech Recognition (ASR), Mel Frequency Cepstral Coefficients (MFCC).*

Table Des Matières

Liste des figures ..... **I**

Liste des tableaux ..... **II**

Liste des acronymes ..... **III**

Introduction générale **1**

Chapitre I : science de tadjweed et les methodes d'apprentissage

1. Introduction **3**

2. Le Mérite De La Lecture Du Coran **4**

2.1. La récompense de sa lecture de sa lecture 4

2.2. La récompense de son enseignement 4

2.3. Le Mérite de l'apprentissage de Coran et de sa mémorisation et du perfectionnement de sa lecture 4

2.4. La récompense de celui dont l'enfant étudie le Coran 5

2.5. L'intercession du Coran en faveur (comme en défaveur) de celui qui le récite le jour de la résurrection 5

2.6. La récompense de ceux qui se réunissent pour le lire et l'étudier 6

3. Définitions de la science du Tadjwîd **6**

4. LES BRANCHES DE LA SCIENCE DE TADJWÎD (**علم القراءات**) 7

4.1. La connaissance du tadjwîd (prononciation à la fois bonne et correcte) 7

4.1.1. علم التجويد النظري ('ilm at-tadjwîd an-nazary) 7

4.1.2. علم التجويد العملي (‘ilm at-tadjwîd al-'amaly) 7

4.2. La connaissance de plusieurs types de lecture 9

4.2.1. Hafs ibn Soulamane ibn al-moughayra ibn Daoud al-assadi al-koufi al-bazar 9

4.2.2. Warsh s'appelle 'Outhmane ibn Sa'id ibn 'Abdoullah ibn 'amrou ibn Soulayman ibn Ibrahim 10

4.3. Les différents modes de récitation du noble coran 10

4.3.1. El Tartil (الّترتيل) 10

4.3.2. El Hadr (الحَدْر) 10

4.3.3. El tadwir (التَدْوِير) 10

4.3.4. El Tahkik (التَّحْقِيق) 10

5. Règles du noûn as-sakîna « نْ » et du tanwîn (أحكام النّون السّاكنة و التنوين) **11**

5.1. L'assimilation 12
5.1.1. L'assimilation partielle avec nasillement (الإدغام بغنة) 12
➢ *La définition de la nasalisation (الغنة)* *12*
5.1.2. L'assimilation totale sans nasillement (الإدغام بدون غنة) 16
6. Le traitement automatique de la parole 17
6.1. Le signal de parole 17
6.2. Extraction des paramètres 18
6.2.1. Les paramètres MFCC (Mel Frequency cepstral Coefficients) 19
6.2.2 Les paramètres LPC (Linear Predictive Coefficients): 19
6.2.3. Les paramètres PLP (Perceptual Linear Prediction) 20
6.3. Méthodes de reconnaissances 20
6.3.1. Les modèles de Markov cachés HMM (Hidden Markov Models) 20
6.3.2. Machine à vecteurs de Supports SVM (supports vectors machine) 20
6.3.3. Déformation temporelle dynamique DTW (dynamic time warping) 21
6.3.4. Les réseaux de neurones 22
7. conclusion 22

Chapitre II : Analyse acoustique et apprentissage

1. Introduction 24
2. L'analyse acoustique par la méthode (MFCC) 24
2.1. La préaccentuation 25
2.2 Segmentation 27
2.2.1. Le chevauchement 27
2.3. Le fenêtrage 28
2.4. La transformer de Fourier discrète 30
2.5. Le banc du filtre MEL 31
2.5.1. La réalisation du banc du filtre MEL 32
2.6. La transformée en cosinus discrète (DCT) 35
3. Les chaines de Markov cachées 37
3.1. Définition de la chaine de Markov 37
3.2. Le principe de fonctionnement de HMM 39
3.2.1. Apprentissage/modélisation 40
3.2.1.1. L'initialisation 40

3.2.1.2. Les probabilités d'évaluation *43*
3.2.1.3. La ré-estimation *46*
3.2.1.4. Le modèle HMM *47*
3.2.2. L'identification 47
3.2.2.1. L'initialisation *47*
3.2.2.2. Les probabilités d'évaluations *48*
3.2.2.3. Résultat (score) *49*
3.2.3. La vérification 50
4. Conclusion 50

Chapitre III : La conception

1. Introduction ..... Erreur ! Signet non défini.
2. Description du système 51
2.1. Construction de la base de données 51
3. Organigramme 52
3.1. Collection des échantillons (enregistrement du son) 53
3.2. Extraction des paramètres acoustique par la méthode (MFCC) 54
3.3. Reconnaissance automatique de la parole par la méthode (HMM) 55
3.3.1. Apprentissage 56
3.3.1.1. Le nombre d'états (N) *57*
3.3.2. La reconnaissance 57
4. Proposition d'une interface graphique 65
5. Conclusion 68

Conclusion générale 69
Références 71
Annexe A 74
Annexe B 75
Annexe C 76

Liste Des Figures

Figure I.1. *Poème de l'Imâm Ibn Al-Jazary*. 11
Figure I.2. *Les deux catégories de l'assimilation.* 12
Figure I.3. *schéma descriptive de la nasalisation.* 13
Figure I.4. *L'assimilation avec la lettre «الياء».* 14
Figure I.5. *L'assimilation avec la lettre «الواو».* 14
Figure I.6. *L'assimilation avec la lettre «الميم».* 15
Figure I.7. *L'assimilation avec la lettre «النون».* 15
Figure I.8. *L'exception de la règle de l'assimilation avec nasalisation* 16
Figure I.9. *L'assimilation avec la lettre «اللام»* 16
Figure I.10. *L'assimilation avec la lettre «الراء».* 17
Figure I.11. *Schéma d'un système de reconnaissance basé sur la comparaison dynamique* 22
Figure II.1. *La chaine de traitement de l'analyse acoustique MFCC.* 25
Figure II.2. *Le code MATLAB pour la tache de préaccentuation de la méthode MFCC.* 26
Figure II.3. *la préaccentuation du signal enregistré « من يشاء ».* 26
Figure II.4. *Le principe de segmentation avec un chevauchement de 50%.* 27
Figure II.5. *Le code MATLAB pour la tache de segmentation de la méthode MFCC.* 28
Figure II.6. *Le principe de Fenêtrage avec un chevauchement de 50%.* 28
Figure II.7. *Le code MATLAB associé à la fenêtre de Hamming.* 29
Figure II.8. *La segmentation et le fenêtrage de la trame n°1 et n° 15 par la fenêtre de Hamming.* 29
Figure II.9. *La segmentation et le fenêtrage de toutes les trames par la fenêtre de Hamming* 30
Figure II.10. *Le code MATLAB associé à la transformer de Fourier discrète.* 30
Figure II.11. *: La transformer de Fourier discrète de la quinzième trame.* 31
Figure II.12. *: Le banc du filtre MEL.* 32
Figure II.13. *: Le code Matlab pour réaliser 13 fréquences séparées linéairement* 33
Figure II.14. *: Le code Matlab pour réaliser 27 fréquences séparées logarithmiquement.* 33
Figure II.15. *: La concaténation des deux fréquences linéaires et logarithmiques.* 33
Figure II.16. *: Les fréquences linéaires et logarithmiques du banc du filtre MEL.* 34
Figure II.17. *: Le calcul de l'espacement linéaire et logarithmique.* 34
Figure II.18. *: Le code Matlab pour tracer un triangle à surface unité.* 34
Figure II.19. *: Le code Matlab pour réaliser le banc du filtre MEL.* 35
Figure II.20. *: Le code Matlab le calcul de la fonction DCT.* 35
Figure II.21. *: Le code Matlab pour calculer l'énergie d'une trame.* 35
Figure II.22. *: Les coefficients spectraux du signale « من يشاء ».* 36
Figure II.23. *: Les énergies et les coefficients cepstraux de toutes les trames.* 36
Figure II.24. *Schéma descriptif du model de Markov caché.* 38
Figure II.25. *Schéma descriptif du système RAP « HMM ».* 40
Figure II.26. *: La chaine de prétraitement du processus d'apprentissage.* 40
Figure II.27. *: La matrice des coefficients cepstreaux du signal.* 41
Figure II.28. *: La modélisation du signal par une chaine de Markov caché.* 41
Figure II.29. *: Le code Matlab pour initialiser la matrice de transition.* 42
Figure II.30. *: Le code Matlab pour initialiser le vecteur de la distribution de la probabilité initiale.* 42
Figure II.31. *: Le code Matlab pour initialiser les matrices mu et sigma.* 43
Figure II.32. *: Le principe de fonctionnement de l'algorithme Forward.* 44
Figure II.33. *: Le principe de fonctionnement de l'algorithme backword.* 45
Figure II.34. *: Le code Matlab pour l'évaluation des probabilités.* 45
Figure II.35. *: Le principe de fonctionnement de l'algorithme Forward-Backward.* 46
Figure II.36. *: Le code Matlab pour la ré-estimation des matrices mu et sigma.* 46
Figure II.37. *: Le code Matlab pour générer les modèles λ = (A_, mu_, sigma_).* 47
Figure II.38. *: La chaine de prétraitement du processus de l'identification.* 47
Figure II.39. *: Le code Matlab pour initialiser la matrice A et le vecteur pi0.* 48
Figure II.40. *: la fonction de l'algorithme de Viterbi.* 48
Figure II.41. *: le résultat de l'algorithme de Viterbi pour 5 réciteurs.* 49

Figure III.1. *Schéma synoptique pour la construction de la base de données*. 51
Figure III.2. O*rganigramme.* 52
Figure III.3. *Le code MATLAB pour enregistrer le son.* 53
Figure III.4. *Le code MATLAB pour la lecture du son.* 53
Figure III.5. *L'analyse acoustique.* 54
Figure III.6. *Le code Matlab pour l'appelle de la fonction MFCC.* 55
Figure III.7. *La phase d'apprentissage en HMM.* 56
Figure III.8. *L'identification.* 58
Figure III.9. *L'identification du réciteur numéro 1.* 58
Figure III.10*. Répartition des scores de bonne récitation et mauvaise récitation d'un système parfait* 60
Figure III.11. *Influence du seuil de décision sur les erreurs d'un système de reconnaissance.* 61
Figure III.12. *la courbe DET.* 61
Figure III.13. *Correcteur automatique des règles de Tadjwîd.* 65
Figure III.14. C*hoisissez une bonne récitation pour l'écouter.* 66
Figure III.15. L'enregistrement du son. 66
Figure III.16. *L'analyse acoustique.* 67
Figure III.17. *La décision (bonne prononciation).* 67
Figure III.18. *La décision (mauvaise prononciation).* 68

Liste des tableaux

Tableau II.1. *Les paramètres du banc du filtre MEL.* 32
Tableau III.1. *la performance des différents systèmes RAP avec différents processus d'analyse acoustique*. 55
Tableau III.2. *La variation de nombre d'états pour une meilleure performance*.....57
Tableau III.3. *Le taux d'identification avec différentes réciteurs*. 59
Tableau III.4. *La performance du système.* 62
Tableau III.5. *Le taux de reconnaissance avec différentes sons (juste) des réciteurs de la base de données*.....63
Tableau III.6. *Le taux de reconnaissance avec différentes sons (faux) des réciteurs de la base de données.* 63
Tableau III.7. *Le taux de reconnaissance avec différentes sons (juste) des apprenants* 63
Tableau III.8. *Le taux de reconnaissance avec différentes sons (faux) des apprenants*.....64

Liste Des Acronymes

DCT: Discrete Cosine Transform

DET: Detection Error Trade-off curves

DFT: Discrete Fourier Transform

DTW: Dynamic Time warping

FFT: Fast Fourier Transform

FIR: Finite Impulse Response

FS: Sampling Frequency

GUI: Graphical User Interface

HMM: Hidden Markov Model

Hz: Hertz

LLR: Log Likelihood Ratio

LPC: Linear Predictive Coding

LPCC: Linear predictive coding cepstra

MFCC: Mel-frequency Cepstral Coefficients

NIST: National Institute of Standards and Technology

PLP: Perceptual Linear Prediction

RAP : Reconnaissance Automatique de la Parole

SVM: Support Vector Machine

Introduction générale

Introduction générale

La parole, comme un moyen de dialogue homme-machine efficace, a donné naissance à plusieurs travaux de recherche dans le domaine de la Reconnaissance Automatique de la Parole (désormais RAP). Un système de RAP est un système qui a la capacité de détecter à partir du signal vocal la parole et de l'analyser dans le but de le transcrire en une chaîne de mots ou phonèmes représentant ce que la personne a prononcé. [1]

La reconnaissance automatique de la parole pour la récitation des mots du noble Coran avec vérification des règles de Tadjwîd est un nouveau secteur de recherche. Car, la science de Tadjwîd est basée sur ***el Mouchafaha*** qui est décrite comme un apprentissage du face à face entre l'apprenant et le professeur qualifié, où le processus d'écoute, de la correction et de la répétition de la correcte récitation se fait dans un temps réel. [2]

La présence des professeurs qualifiés de nos jours est limitée dans le nombre (on trouve un à trois professeurs par mosquée) aussi bien que dans les heures d'éducations (maximum trois heures par semaine). Cela implique que les heures de pratique sont très courtes et ne satisfont pas les besoins de l'apprenant. Donc, ses technologies de l'information viennent renforcer cette insuffisance, et le système RAP peut être considéré comme une autre alternative d'apprendre la récitation. Ce système n'a pas l'ambition de remplacer les professeurs mais d'encourager les étudiants à s'entrainer individuellement car le système va juger si la récitation est juste ou fausse.

Dans le cadre de ce travail de recherche, nous présentons une forme d'une interaction homme-machine, celle liée à **l'identification** d'un signal de la parole et **la vérification** de sa bonne prononciation ; c'est dans ce cadre que s'inscrit notre projet de fin d'études.

La problématique à laquelle nous essayerons de répondre, tout au long de notre projet de fin d'étude, est la suivante :

- comment peut-on utiliser le domaine de traitement de la parole pour moderniser l'apprentissage de la récitation du noble coran?

Pour répondre à cette question d'ordre générale, il nous est indispensable de répondre à ses sous questions :

- quel est le processus utilisé pour extraire les caractéristiques de signale de la parole?
- Comment modéliser un signal par un processus d'apprentissage ?

- Comment en effet combiner ses deux processus pour avoir à la fin un système fiable?

Notre travail repose sur trois chapitres : Le premier présentera des généralités sur la science de Tadjwîd et la science de la reconnaissance de la parole, leurs principales caractérisations et règlementations. Le deuxième citera les deux processus utilisés pour réaliser notre application. Et pour terminer, nous allons consacrer le dernier chapitre pour la conception.

Nous conclurons notre modeste travail par un compte rendu de ce qui a été dit en rassemblant les résultats auxquelles nous avons arrivés à la fin de chaque chapitre et en répondant aux questions posées ci-dessus.

Chapitre 1:

Règles de récitation du Coran et les méthodes d'apprentissage

1. Introduction

Le Saint Qu'ran, ultime message d'Allâh سبحانه وتعالى a été révélé afin que les hommes y puisent leur guidance et leurs connaissances. [3]

Depuis sa révélation, et par la grâce d'Allâh سبحانه وتعالى cette source inépuisable n'a été sujette à aucune altération tant au niveau de son sens qu'au niveau de ses moindres lettres et signes de vocalisation. Allâh سبحانه وتعالى s'est porté garant de sa préservation de par les siècles et les générations successives [3] :

﴿ إِنَّا نَحْنُ نَزَّلْنَا الذِّكْرَ وَ إِنَّا لَهُ لَحَافِظُونَ ﴾ [1].

Cette préservation s'est produite sous deux formes: La préservation par voie écrite et la préservation par voie orale. Cette dernière qui nous importe concerne la façon dont le coran a été lu et rapporté dans ses moindres détails. Ainsi le coran nous a été rapporté par des chaînes de transmission fiables passant par le prophète ﷺ et remontant jusqu'à Allah. [3]

Ces chaînes de transmission nous garantissent la similitude de la lecture coranique du dernier maillon jusqu'au premier. [3]

Les savants par la suite ont mis en place un système de règles précises permettant une bonne lecture du coran, celle-ci ne pouvant se faire que grâce à l'accompagnement et l'encadrement par un professeur qualifié. [3]

Comme nous avons abordé dans l'introduction générale que la présence des professeurs qualifiés de nos jours est limitée et les heures d'éducations sont insuffisantes. La solution est d'utiliser les systèmes RAP.

Ce chapitre va d'abord définir la science de Tadjwed (Tartil) en citant quelques caractéristiques de ce dernier. Ensuite, présenter les systèmes les plus utilisés dans la reconnaissance de la parole et enfin expliquer brièvement leur fonctionnement.

Nous tenons à signaler que nous sommes convaincues du proverbe qui dit « traduire c'est trahir ». En effet, le présent chapitre contient des énoncés en arabe vue la difficulté de les traduire. Ces derniers se rapportent aux règles de récitation du coran qui, malgré nos tentatives de traduction, ont été gardés tels quels.

2. Le Mérite De La Lecture Du Coran

L'apprentissage du Coran, son enseignement et sa récitation ont beaucoup de mérites parmi lesquels :

2.1. *La récompense de sa lecture*

Le Prophète ﷺ a dit : « Celui qui récite une lettre du livre d'Allah obtient une bonne action, et toute bonne action est récompensé au décuple » **Rapporté par At-Tirmidhi.** [4]

عن أبي مسعود رضي الله عنه قال : قال رسول الله صلى الله عليه وسلم : ﴿ مَن قَرَأَ حَرْفًا مِن كِتَابِ اللهِ فَلَهُ بِهِ حَسَنَة، وَ الحَسَنَةُ بِعَشْرِ أَمْثَالِهَا لَا أَقُولُ ألم حَرْف، أَلِفٌ حَرْفٌ وَ لَامٌ حَرْفٌ وَ مِيمٌ حَرْفٌ ﴾ **رواه الترمذي والدارمي وغيرهما وصححه الألباني.**

2.2. *La récompense de son enseignement*

Le Prophète ﷺ a dit : « Le meilleur d'entre vous et celui qui apprend le coran et qui l'enseigne » **Rapporté par Al Boukhari.** [4]

عن عثمان بن عفان رضي الله عنه قال: قال الرسول صلى الله عليه وسلم : ﴿ خَيْرُكُم مَن تَعَلَّمَ القُرْآنَ وَ عَلَّمَهُ ﴾ **رواه البخاري وغيره.**

2.3. *Le Mérite de l'apprentissage de Coran et de sa mémorisation et du perfectionnement de sa lecture*

Le Prophète ﷺ a dit : « Celui qui récite le Coran en l'ayant appris par cœur est avec les Anges nobles et obéissants ; et celui qui le récite en cherchant à l'apprendre alors que cela lui est pénible aura une double récompense » **Rapporté par Al Boukhari et Mouslim.** [4]

عن عائشة رضي الله عنها قالت: قال رسول الله صلى الله عليه وسلم: ﴿ الذِي يَقْرَأُ القُرْآنَ وَ هُوَ مَاهِرٌ بِهِ مَعَ السَّفَرَةِ الكِرَامِ البَرَرَةِ، وَ الذِي يَقْرَأُ القُرْآنَ وَ يُتَعْتِعُ فِيهِ وَ هُوَ عَلَيْهِ شَاقٌ لَهُ أَجْرَانِ ﴾ **رواه البخاري ومسلم.**

Il a également dit : « Il sera dit à celui qui connait le Coran : lis et monte plus haut, récite lentement et clairement comme tu récitais dans le bas monde, car ta place se situe auprès du dernier verset que tu liras » **Rapporté par At-Tirmidhi**. [4]

عن عبد الله بن عمر رضي الله عنهما قال: قال رسول الله صلى الله عليه وسلم: ﴿يُقَالُ لِصَاحِبِ الْقُرْآنِ :اقْرَأْ وَ ارْتَقْ وَ رَتِّلْ كَمَا كُنْتَ تُرَتِّلُ فِي الدُّنْيَا ,فَإِنَّ مَنْزِلَتَكَ عِنْدَ آخِرِ آيَةٍ تَقْرَؤُهَا﴾ **رواه الترمذي و أبوداود و النسائي و أحمد و الحاكم وابن حبان**

2.4. *La récompense de celui dont l'enfant étudie le Coran*

Le Prophète ﷺ a dit : « Le Jour de la Résurrection, une couronne de lumière dont la clarté est comme la clarté du soleil sera portée sur la tête du père et de la mère de celui qui récite le Coran, l'étudie et le met en pratique, et ces derniers seront également vêtus de deux vêtements à côté desquels ce bas monde n'a aucune valeur. Alors ils diront : « Qu'est ce qui nous a valu de porter ceci ? » On leur dira : « Parce que votre enfant à appris le coran ». » **Rapporté par Al-Hakim.** [4]

عن أبي هريرة قال: قال رسول الله صلَّى الله عليه وسلم: ﴿يَجِيءُ الْقُرآنُ يَوْمَ القِيَامَةِ كَالرَّجُلِ الشَّاحِبِ يَقُولُ لِصَاحِبِهِ: هَلْ تَعْرِفُنِي ؟ أَنَا الذِي كُنْتُ أُسْهِرُ لَيْلَكَ وَ أُظْمِئُ هَوَاجِرَكَ ، وَ إِن كُلُّ تَاجِرٍ مِن وَرَاءِ تِجَارَتِهِ وَ أَنَا لَكَ اليَوْمَ مِن وَرَاءِ كُلِّ تَاجِرٍ فَيُعْطَى المُلْكُ بِيَمِينِهِ وَ الخُلْدُ بِشِمَالِهِ وَ يُوضَعُ عَلَى رَأْسِهِ تَاجَ الوَقَارِ ، وَ يُكْسَى وَالِدَاهُ حُلَّتَيْنِ لَا تَقُومُ لَهُمَا الدُّنْيَا وَ مَا فِيهَا ، فَيَقُولَانِ : يَا رَبْ أَنَّى لَنَا هَذَا ؟ فَيُقَالُ لَهُمَا : بِتَعْلِيمِ وَلَدِكُمَا القُرْآن﴾ **واه الطبراني في الأوسط.**

2.5. *L'intercession du Coran en faveur (comme en défaveur) de celui qui le récite le jour de la résurrection*

Le Prophète ﷺ a dit : « Lisez le Coran, car il viendra le Jour de la résurrection en intercesseur pour ceux qui le lisent » **Rappporté par Mouslim.** [4]

عن أبي أمامة رضي الله عنه قال: سمعت رسول الله صلى الله عليه وسلم يقول ﴿اقْرَؤُوا الْقُرْآنَ فَإِنَّهُ يَأْتِي يَوْمَ القِيَامَةِ شَفِيعًا لِأَصْحَابِهِ﴾ **رواه مسلم.**

Il a également dit : « Le Jeûne et le Coran intercéderont en faveur du serviteur (d'Allah) le Jour de la Résurrection » **Rapporté par Ahmad et Al Hakim**. Mais le coran le jour de la résurrection peut être en ta défaveur car tout va dépendre de la façon dont tu vas la mettre en pratique. [4]

2.6. *La récompense de ceux qui se réunissent pour le lire et l'étudier*

Le Prophète ﷺ a dit : « Tout groupe de personnes qui se rassemble dans une des mosquées d'Allah, pour lire le Coran et l'étudier entre eux : la quiétude descend sur eux ; ils sont enveloppés par la Miséricorde ; entourés par les Anges ; et Allah les évoque à ceux qui sont près de lui. » **Rapporté par Mouslim**. [4]

عن أبي هريرة رضي الله عنه قال: قال رسول الله صلى الله عليه وسلم: ....وَ مَا اجْتَمَعَ قَوْمٌ فِي بَيْتٍ مِن بُيُوتِ الله، يَتْلُونَ كِتَابَ الله، وَ يَتَدَارَسُونَهُ بَيْنَهُمْ، إِلَّا نَزَلَتْ عَلَيْهِمُ السَّكِينَةُ، وَ غَشِيَتْهُمُ الرَّحْمَةُ، وَ حَفَّتْهُمُ الملَائِكَةُ، وَ ذَكَرَهُمُ الله فِيمَنْ عِنْدَه رواه مسلم وغيره

Après avoir cité quelques mérites de la lecture du noble Coran, il est recommandé au lecteur du Coran de le lire soigneusement en donnant à chaque lettre sa mesure exacte en termes de diction conformément à la parole du Très-Haut : "Et récite le Coran, lentement et clairement." Le musulman ne doit pas appréhender la lecture du Coran ni craindre de se tromper en récitant. Au contraire, il doit prolonger sa compagnie avec le Coran pour atteindre un bon niveau de récitation par la répétition et la pratique. Le Très-Haut dit : "En effet, Nous avons rendu le Coran facile pour la méditation. Y a-t-il quelqu'un pour méditer ?" [4]

﴿ وَلَقَدْ يَسَّرْنَا الْقُرْآنَ لِلذِّكْرِ فَهَلْ مِنْ مُدَّكِرٍ ﴾ [I.2]

Il serait appréciable également que nous connaissons, ô croyant, quelques règles de tajwîd ou que nous entraînons à lire le Coran en présence d'un mémorisateur ou d'un maître-récitateur afin qu'il t'indique la manière de le restituer correctement. [4]

3. Définitions de la science du Tadjwîd

Le mot tadjwîd vient de la racine arabe « **jawwada - جَوَّدَ**» qui signifie « rendre meilleur », ou « améliorer ». **« التَّجْوِيدْ »** signifie donc : l'embellissement. [3]

Dans le langage technique, le terme « tadjwîd » a deux significations distinctes et peut faire référence à :

- ✓ Une bonne récitation prononcée correctement
- ✓ Un mode de récitation à vitesse moyenne

En ce qui concerne le Saint Qu'ran, le tadjwîd est donc une science dont l'objectif est de corriger sa récitation dans le but de la parfaire, en accordant à chaque lettre son point d'articulation correct (مخارج الحروف) et ses caractéristiques (صفات الحروف), sans exagération et

avec douceur, de la même manière qu'elle a été révélée et conformément à ce que les musulmans ont hérité du messager d'Allâh سبحانه و تعالى.

On peut atteindre de l'art du tadjwîd, en préservant la langue de toutes les erreurs dans la récitation du livre d'Allâh سبحانه و تعالى et tirer le bénéfice de Sa satisfaction.

Le premier qui a instauré cette science de manière pratique est notre prophète et envoyé Mouhamad ﷺ, qui l'a reçu de l'archange Jibrîl, qui lui-même l'a reçu de Allâh. سبحانه و تعالى

Le Saint Qu'ran a été en effet révélé avec une récitation qui lui est spécifique, et qui se lit différemment de tout autre texte.

Quant aux premiers qui ont établi les règles théoriques, ce sont les savants de la langue arabe tels al-Khalil Ibn Ahmed Al-Farahidi et son élève Sibawayh.

4. Les branches de la science de Tadjwîd (*علم القراءات*)

La science de la récitation inclut trois branches principales :

- ✓ La connaissance du tadjwîd.
- ✓ La connaissance de plusieurs types de lecture.
- ✓ Les différents modes d'apprentissage.

4.1. *La connaissance du tadjwîd (prononciation à la fois bonne et correcte)*

Elle se divise en deux parties :

4.1.1. *علم التجويد النظري ('ilm at-tadjwîd an-nazary)*

C'est-à-dire la connaissance des règles théoriques du tadjwîd, de ses normes, de ses conditions et de ses appellations, telle que les règles de l'allongement (أحكام المد), les règles du noûn as-sâkina wa at-tanwîn(أحكام النون الساكنة و التنوين) , et autres règles.

La connaissance de cette science incombe à la communauté islamique mais non à chaque musulman (fard kifâya – فرض كفاية). En d'autres termes, il suffit que quelques musulmans connaissent cette science pour que le reste de la communauté soit épargné du péché.

4.1.2. *علم التجويد العملي ('ilm at-tadjwîd al-'amaly)*

C'est la connaissance pratique de ces règles. C'est la manière correcte et juste de la récitation du Qu'ran, comme l'a récité le Prophète d'Allâh سبحانه و تعالى .

Cette connaissance est obligatoire à tout musulman ou musulmane qui désire lire ou apprendre quelques versets du Saint Qu'ran (fard 'ayn – فرض عين), dans la mesure de ses compétences.

Les experts du tadjwîd (psalmodie) et des lectures ainsi que les maîtres de la récitation coranique sont unanimes que le Saint Qu'ran doit être récité de façon spécifique, exactement comme il a été révélé par l'ange Jibrîl au prophète ﷺ, et comme l'ont reçu de lui la majorité de ses compagnons رضي الله عنهم, qui l'accentuait par une lecture particulière et singulière et l'ont transmis à leur tour jusqu'à ce qu'il nous soit parvenu.

Les preuves de son caractère obligatoire sont nombreuses :

Allah سبحانه و تعالى a dit :

﴿ وَ رَتِّلِ الْقُرْآنَ تَرْتِيلًا ﴾ [I.3]

﴿ وَقَالَ الَّذِينَ كَفَرُوا لَوْلَا نُزِّلَ عَلَيْهِ الْقُرْآنُ جُمْلَةً وَاحِدَةً ۚ كَذَٰلِكَ لِنُثَبِّتَ بِهِ فُؤَادَكَ ۖ وَرَتَّلْنَاهُ تَرْتِيلًا ﴾ [I.4]

﴿الَّذِينَ آتَيْنَاهُمُ الْكِتَابَ يَتْلُونَهُ حَقَّ تِلَاوَتِهِ أُولَٰئِكَ يُؤْمِنُونَ بِهِ ۗ وَمَن يَكْفُرْ بِهِ فَأُولَٰئِكَ هُمُ الْخَاسِرُونَ ﴾[I.5]

Notre prophète Mohamed ﷺ a dit :

روى البخاري ومسلم عن عائشة رضي الله عنها قالت: قال رسول الله صلى الله عليه وسلم: ﴿الذِي يَقْرَأُ القُرْآنَ وَ هُوَ مَاهِرٌ بِهِ مَعَ السَّفَرَةِ الكِرَامِ البَرَرَةِ، وَ الذِي يَقْرَأُ القُرْآنَ وَ يُتَعْتِعُ فِيهِ وَ هُوَ عَلَيْهِ شَاقٌ لَهُ أَجْرَانِ﴾ **رواه البخاري ومسلم.**

روى أبو داود والترمذي عن عبدالله بن عمرو بن العاص- رضي الله عنهما- عن -النبي -صلى الله عليه وسلم قال: ﴿يُقَالُ لصَاحِبِ القُرآن : إِقْرَأ وارتَقِ وَرَتِل كَمَا كُنتَ تُرَتِّلُ فِي الدُنيَا ، فإنَّ مَنزِلَتَكَ عِندَ آخِرِ آيَةٍ تَقْرَأُهَا﴾ **صحيح الجامع.**

روى زيدُ بن ثابت - رضي الله عنه - عن النبي - صلَّى الله عليه وسلَّم - أنه قال: ﴿إنَّ الله يُحبُّ أن يُقرأ القُرآنَ كَمَا أُنْزِلَ﴾ **في ضعيف الجامع.**

حدثنا نصر بن عاصم الأنطاكي حدثنا الوليد ومبشر يعني ابن إسمعيل الحلبي عن أبي عمرو قال يعني الوليد حدثنا أبو عمرو قال حدثني قتادة عن أبي سعيد الخدري وأنس بن مالك عن رسول الله صلى الله عليه وسلم قال: ﴿سَيَكُونُ فِي أُمَّتِي

اخْتِلَافٌ وَفُرْقَةٌ قَومٌ يُحسِنُونَ القِيلَ وَيُسِيئُونَ الفِعلَ يَقرَءُونَ القُرآنَ لَا يجاوز تَرَاقِيهِم يَمرُقُونَ من الدِّينِ مُرُوقَ السَّهمِ من الرَّمِيَّة لَا يَرْجِعُونِ حَتَى يَرْتَدَ عَلَى فَوْقِهِ هُمْ شَرُّ الخَلْقِ وَ الخَلِيقَةِ ... **صحيح أبى داود.**

Ceux donc, qui ne récitent pas le Saint Qu'ran conformément à la manière dont il a été révélé, ne le récitent pas comme il se doit.[3]

Cette manière de récitation consiste :

- à soigner son élocution lors de la lecture du Qu'ran,
- s'assurer de l'exactitude des points d'articulation des lettres,
- produire une récitation juste en donnant à chaque lettre le son précis qui lui correspond : (إعطاء كل حرف حقّه)
- en y mettant la manière pour obtenir le résultat le plus parfait possible.

4.2. *La connaissance de plusieurs types de lecture*

Au début, la révélation ne s'est faite que dans une seule « lettre », puis le messager d'Allah سبحانه و تعالى demanda à Jibrîl jusqu'à ce qu'il lui apprenne à réciter suivant 7 « lettres ». [3]

Ibn Abbas رضي الله عنهم rapporte que le Prophète ﷺ a dit : « Jibrîl m'a appris au début de réciter le Qu'ran suivant une seule «lettre» (harf). Et puis je l'ai sollicité de façon répétée, et il a porté les « lettres » à sept ». Rapporté par al-Boukhari, (3047) et Mouslim (819). [3]

Les deux modes de lecture les plus connus sont celles de Hafs et de Warsh :

4.2.1. *Hafs ibn Soulamane ibn al-moughayra ibn Daoud al-assadi al-koufi al-bazar*

Il est né en l'an 90 après al hidjra et est mort en 180 h. Sa chaîne de transmission : Il a pris sa lecture de l'Imam 'Assim qui était l'Imam de Koufa (mort en 127 ou 128 h). 'Assim a pris sa lecture de Abou 'abdi-Rahman 'Abdoullah ibn As-Soulami, de Abi Maryam Zir Al Assadi et de Abi 'Amrou Sa'id As-Shaybani. Les trois cités ont appris de 'Abdoullah ibn Mas'oud, As-Soulami et Zir رضي الله عنهم ont également appris de 'Outhman ibn Al 'Affan et aussi de 'Ali رضي الله عنهم. De plus As-Soulami a aussi appris de Oubayd ibn Ka'b ainsi que de Zayd ibn Thabit رضي الله عنهم. Enfin les cinq sahaba رضي الله عنهم ont appris du Prophète d'Allâh سبحانه و تعالى. [3]

4.2.2. *Warsh s'appelle 'Outhmane ibn Sa'id ibn 'Abdoullah ibn 'amrou ibn Soulayman ibn Ibrahim*

Il est né en 110 et mort en 197 h. Sa chaîne de transmission : Il a pris sa lecture de l'Imam Nafi' qui était l'Imam de Médine (mort en 170 h). Nafi' a appris de la part de 70 parmi les tabi'ines parmi lesquels Abou Ja'far, 'Abd-Rahman ibn Harmaz al-A'radj, Mouslim ibn Joundoub, Mouhamad ibn mouslim ibn Chihab Az-Zouhri, Salah ibn Khou-at, Chayba ibn Nasah et yazid ibn Roumane. Parmi ceux chez qui ont appris les tabi'ines mentionnés il y a 'Omar ibn Al khattab, 'Abdallah ibn 'Abbas, Abou Hourayra et 'Abdallah ibn 'ayach ibn Abi Rabi'a Al Makhzoumi رضي الله عنهم qui eux ont appris du Prophète d'Allâh سبحانه و تعالى. [3]

4.3. *Les différents modes de récitation du noble coran*

4.3.1. *El Tartil (الترتيل)*

وهو التوئدة في القراءة و الطمأنينة فيها و إعطاء كل حرف من الحروف حقّه من المخارج و الصفات، مصدقا لقوله تعالى « وَ رَتِّلِ الْقُرْآنَ تَرْتِيلاً» وقوله: «و رَتَّلْنَاه تَرْتِيلا» [5]

4.3.2. *El Hadr (الحَدْر)*

وهو الإسراع في القراءة نوعا ما مع مراعات الصفات و المخارج، و المدود و الغنات، وتبيين الحروف و اعتماد الإظهار.[5]

4.3.3. *El tadwir (التَدْوِير)*

وهو التوسط بين الترتيل و الحدر، طبعا مع مراعاة الصفات و المخارج، و كل الجوانب المتعلقة بحق الحرف.[5]

4.3.4. *El Tahkik (التَّحْقِيق)*

و يسمى أيضا التجويد، وهو القراءة بطمأنينة تامة، و تحقيق الحروف مع عدم الإسراف في إمداد المدود و الغنات كما يفعله بعض الجهلة كتجميل الصوت بطول الغنة و تعدي حق المد (نسال الله العافية) مع الإشارة إلى أن التجويد لم يجزه العلماء إلا في حق المتعلم.[5]

Finissons cette définition de la science de Tajwid par cette al-jazariya « االْجَزرِيَة » :

التجويد	
وَالأَخْذُ بِالتَّجْوِيدِ حَتْـمٌ لاَزِمُ	مَنْ لَمْ يُجَوِّدِ الْقُـرآنَ آثِـمُ
لأَنَّهُ بِهِ الإِلَـهُ أَنْـزَلاَ	وَهَكَـذَا مِنْـهُ إِلَيْنَا وَصَـلاَ
وَهُوَ أَيْضاً حِلْـيَةُ الـتِّلاَوَةِ	وَزِينَـةُ الأَدَاءِ وَالْقِـرَاءَةِ
وَهُوَ إِعطَاءُ الْحُرُوفِ حَقَّهَا	مِنْ صِـفَةٍ لَـهَا وَمُستَحَقَّهَـا
وَرَدُّ كُلِّ وَاحِـدٍ لأَصْلِـهِ	وَاللَّفْـظُ فِي نَظِيرِهِ كَمِثْـلِهِ
مُكَمِّلاً مِنْ غَيْرِ مَا تَكَلُّـفِ	بِاللُّطْفِ فِي النُّطْقِ بِلاَ تَعَسُّفِ
وَلَيْـسَ بَيْنَـهُ وَبَيْنَ تَرْكِهِ	إِلاَّ رِيَـاضَةُ امْـرِئٍ بِفَكِّـهِ

Figure I.1. *Poème de l'Imâm Ibn Al-Jazary.*

5. Règles du noûn as-sakîna « نْ » et du tanwîn *(أحكام النّون السّاكنة و التنوين)*

La lettre noûn sans voyelle نْ est appelée « noûn as-sakîna ».

Le tanwîn (التنوين) est le nom donné en arabe à la double vocalisation se trouvant à la fin des mots et dont le son est celui du noûn sakîna نْ. Son symbole c'est le doublement de voyelle (deux damma, deux fatha ou deux kasra). Soit, la terminaison d'un mot par :

- الفتحتانِ : an ً
- الكسرتان: in ٍ
- الضمتان: oun ٌ

Le noûn as-sakîna نْ est caché à l'écrit mais il se prononce à l'oral. [3]

Exemples

- مُبَارَكًا se prononce « Moubarak**an** »
- مُبَارَكٌ se prononce « Moubarak**oun**»
- مُبَارَكٍ se prononce « Moubarak**in** »

La lettre noûn sans voyelle نْ ou le tanwîn ً, ont deux points de sorties et suivent les même règles de tadjwîd :

La première, c'est quand le نْ est prononcé clairement, la pointe de la langue vient se poser sur la racine des incisives des dents supérieures. [3]

La deuxième, c'est quand elle est dissimulé ou assimilé avec une nasalisation, la langue ne touche pas la gencive, elle reste au milieu de la bouche. [3]

Il y a 4 règles concernant le noûn sakîna ou le tanwîn ً ٌ ٍ à appliquer selon le cas, lorsqu'ils sont suivis d'une des lettres de l'alphabet : [3]

- الإظهار (al-izhâr), la clarification
- الإدغام(al-idghâm), l'assimilation
- الإقلاب(al-iqlâb), la substitution
- الإخفاء (al-ikhfa'), la dissimulation

Dans notre corpus nous s'intéressons seulement à la deuxième règle de l'assimilation.

5.1. *L'assimilation*

Dans la langue arabe, al-idghâm signifie ***l'emboîtement, l'insertion***.

On applique cette règle lorsque le noûn as-sakîna نْ ou le tanwîn ً ٌ ٍ est suivi par l'une des six lettres de al-idghâm **(ي – و – ن– م – ل – ر).** Ses lettres sont regroupées dans le mot : **يَرْمَلُونَ.**

Techniquement, la règle d'al-idghâm consiste à faire emboîter la lettre du noûn as-sakîna نْ ou du tanwîn ً ٌ ٍ avec une de ces 6 lettres vocalisées - moutaharaka مُتَحَرِّكَة, de sorte à ne ***former plus qu'une seule lettre redoublée***, mouchadada (مُشَدَّدٌ).Al-idghâm se divise en 2 catégories : [3]

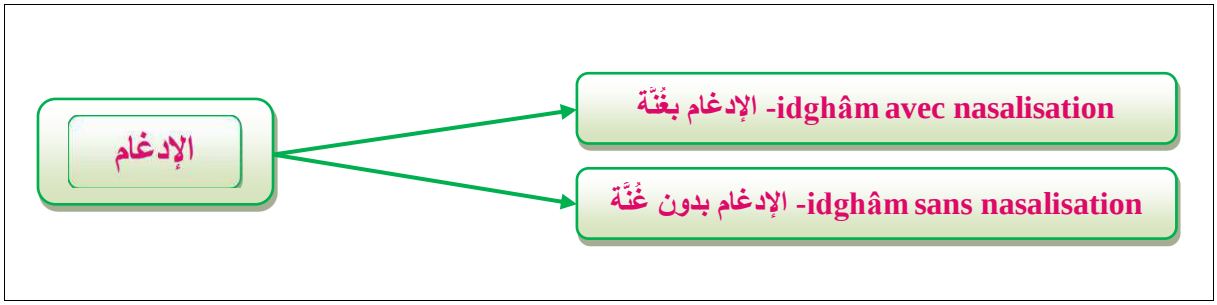


Figure I.2. *Les deux catégories de l'assimilation.*

5.1.1. *L'assimilation partielle avec nasillement (الإدغام بغنة)*

- La définition de la nasalisation (*الغنة*)

La ghounna est une caractéristique obligatoire pour le mîm م et le noûn ن, peu importe leur état (qu'elles soient vocalisées, portant une soukoûn ْ, lors du izhâr, du idghâm ou bien lors de l'ikhfâ). Son point de sortie est la cloison nasale (الخيشوم). Les deux figures représentent le point d'articulation du la nasalisation dans la lettre ن et la lettre م المُشَّددتين. [3]

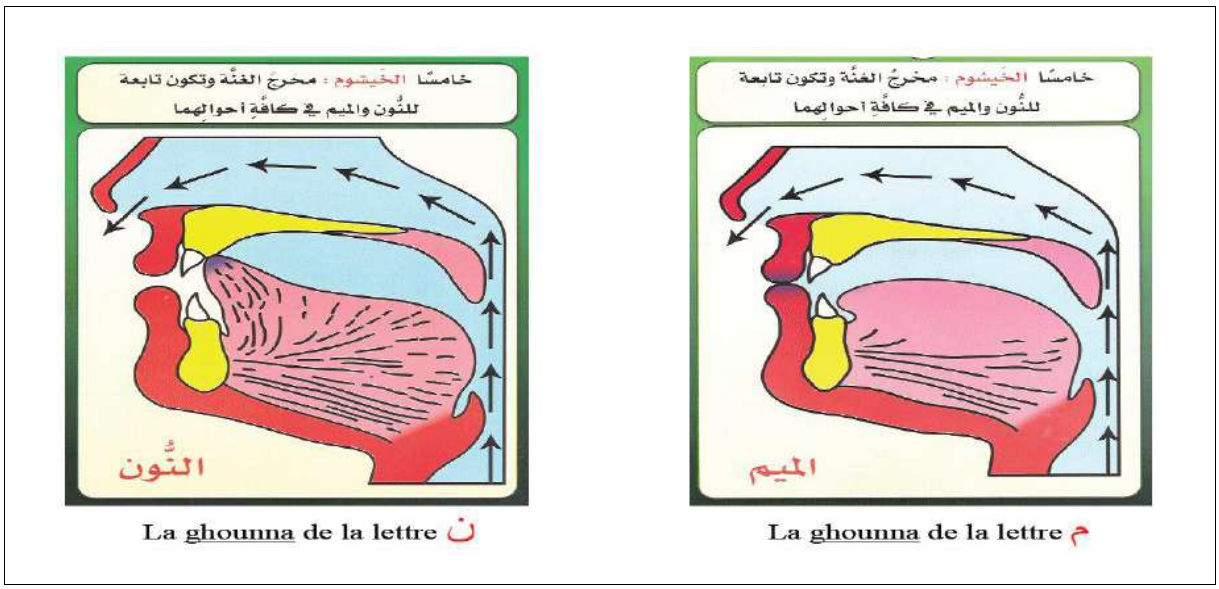


Figure I.3. *schéma descriptive de la nasalisation.*

On appelle cet idghâm : « nâqis » (الإدْغَامُ النَّاقِص) car il reste la caractéristique de la ghounna الْغُنَّة (le nasillement) du noûn as-sâkina نْ et du tanwîn lors de l'idghâm. [3]

Cet idghâm ne peut se faire qu'entre deux mots lorsque le noûn as-sakîna نْ ou le tanwîn ٌ ٍ ً est suivi d'une de ces 4 lettres : ي – و – ن – م. [3]

Ces lettres sont regroupées dans le mot: «يَنْمُو» pour faciliter l'apprentissage des règles de récitation. [3]

Techniquement, le noûn as-sakîna نْ ou le tanwîn ٌ ٍ ً disparaît donc mais on prononce la lettre qui suit avec nasillement, de deux temps ou (حَرَكَتَيْن). [3]

Exemples (الإدغام بغنة)

- L'assimilation avec la lettre «الياء»

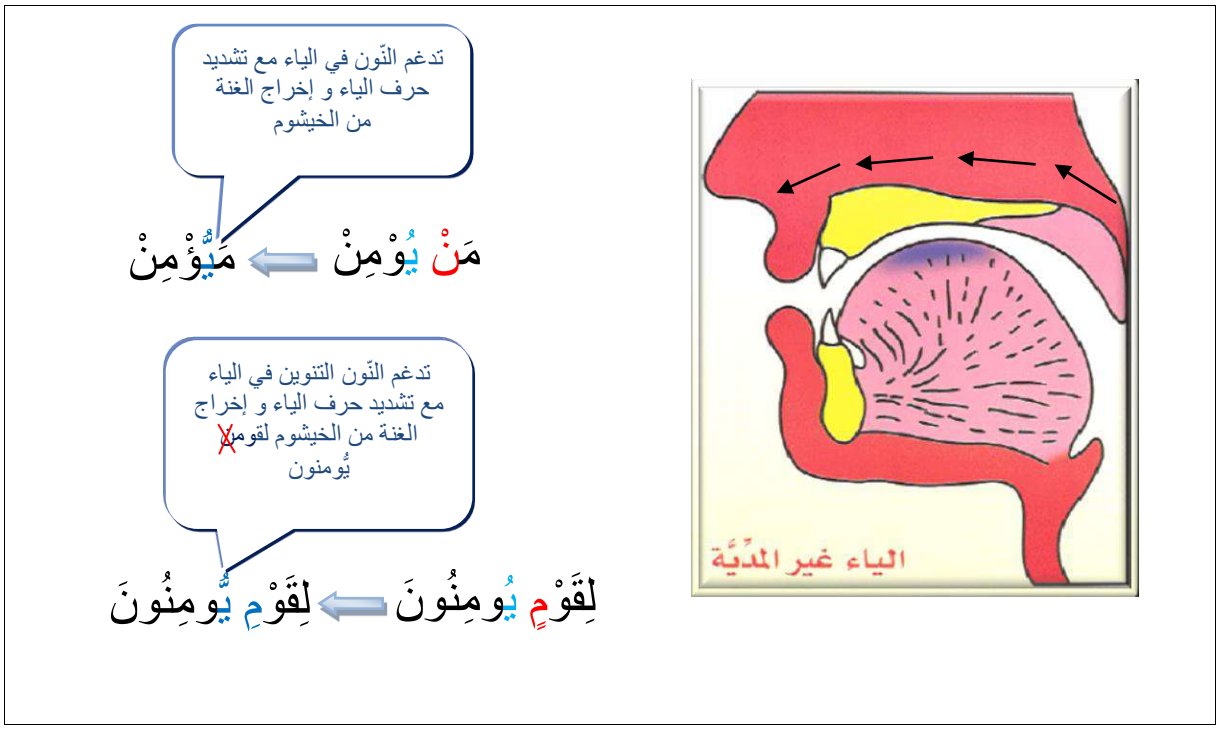


Figure I.4. *L'assimilation avec la lettre «الياء».*

- L'assimilation avec la lettre «الواو»

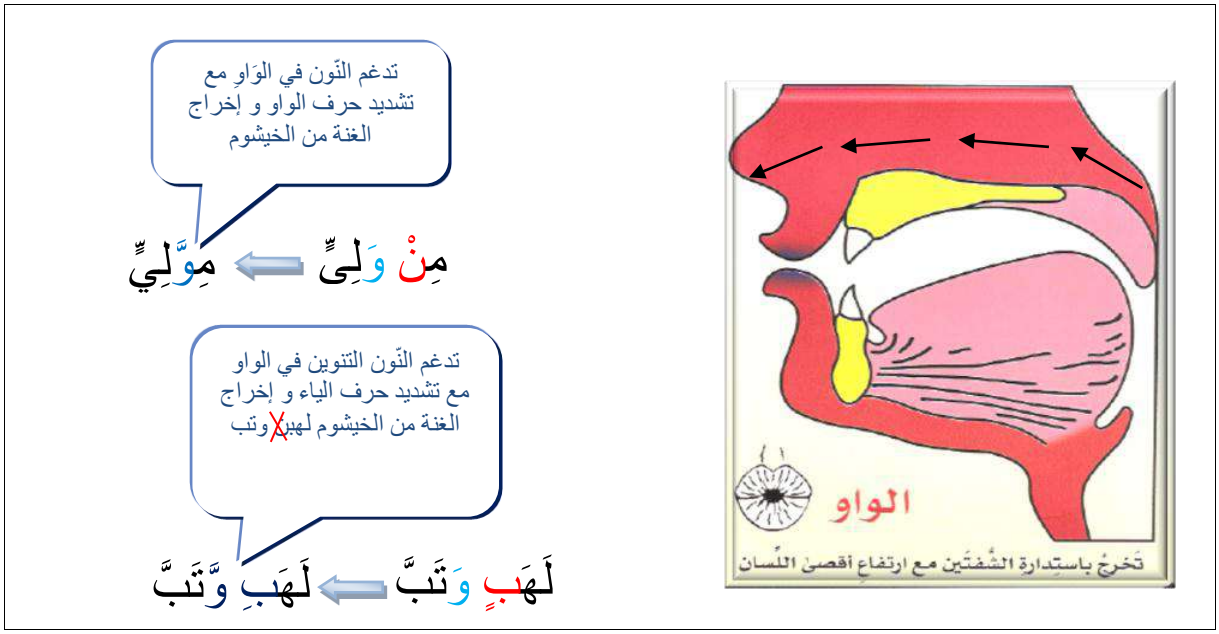


Figure I.5. *L'assimilation avec la lettre «الواو».*

- L'assimilation avec la lettre «الميم»

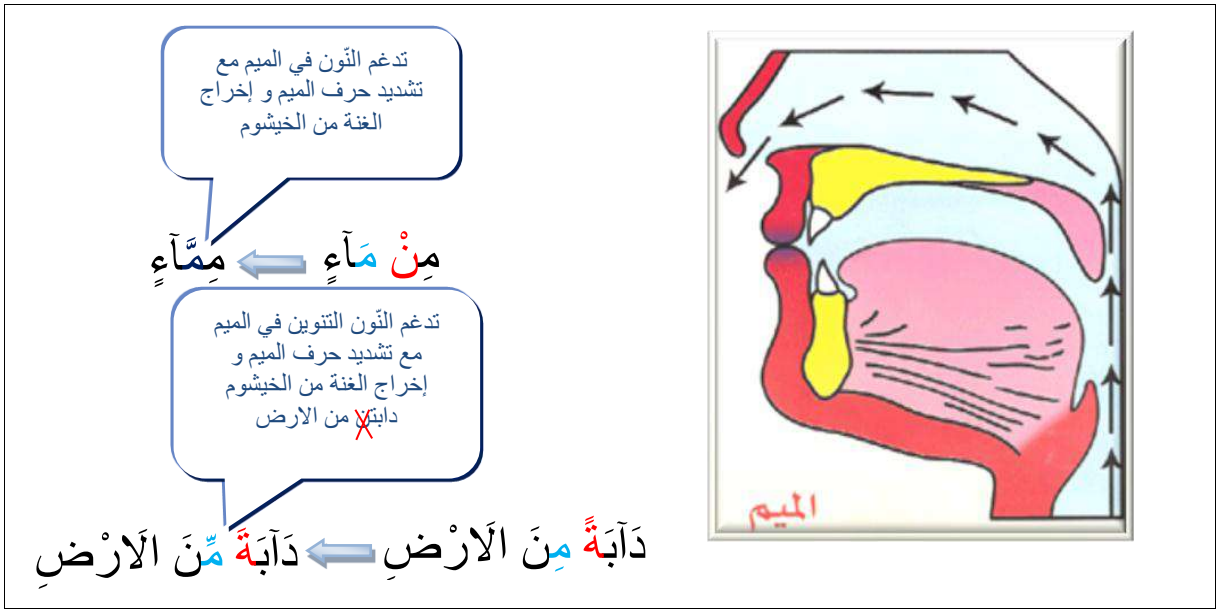


Figure I.6. *L'assimilation avec la lettre «الميم».*

- L'assimilation avec la lettre «النون»

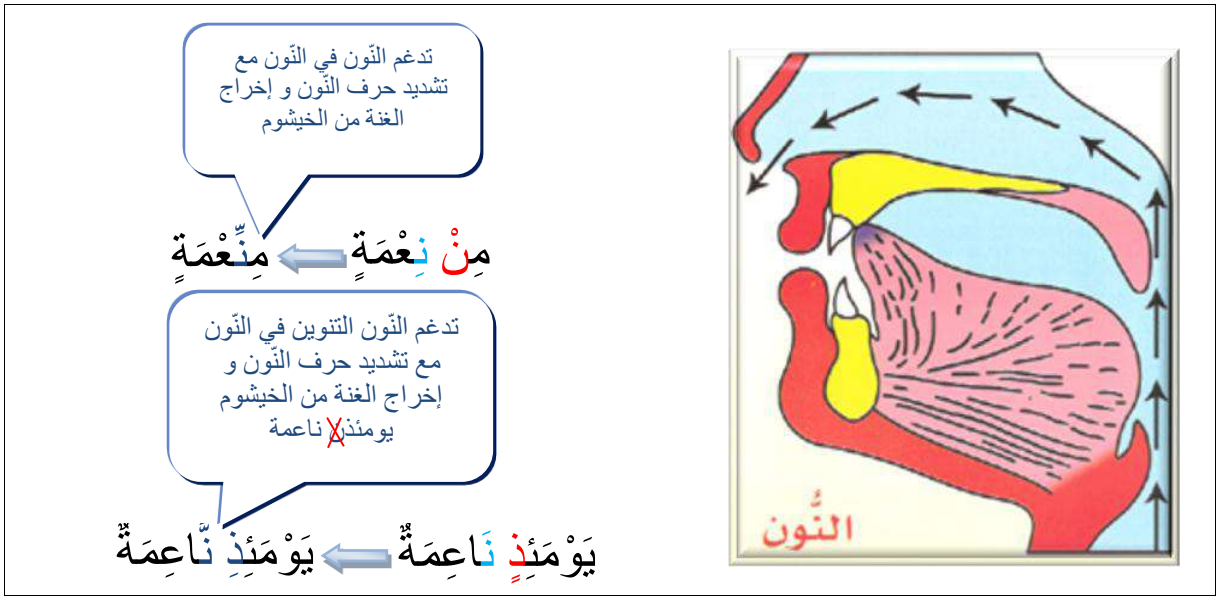


Figure I.7. *L'assimilation avec la lettre «النون».*

Exception

Les savants excluent quatre mots dans le Saint Qur'an où il n'applique pas la règle d'assimilation avec nasalisation car si on applique cette règle le sens du mot sera changer les quatre mots sont: (دُنْيَا، بُنْيَانٌ، صِنْوَانٌ، قِنْوَانٌ) ➔ (دُنْيَا، بُنْيَانٌ، صِوَّانٌ، قِوَّانٌ). [3]

الدُّنْيَا	بَلْ تُؤْثِرُونَ ٱلْحَيَوٰةَ ٱلدُّنْيَا Sourate Al-'Ala verset 16
قِنْوَانٌ	وَمِنَ ٱلنَّخْلِ مِن طَلْعِهَا قِنْوَانٌ دَانِيَةٌ وَجَنَّٰتٍ مِّنْ أَعْنَابٍ Sourate Al-An'am verset 99
صِنْوَانٌ	وَنَخِيلٌ صِنْوَانٌ وَغَيْرُ صِنْوَانٍ يُسْقَىٰ بِمَآءٍ Sourate Ar-Rad verset 4
بُنْيَانٌ	إِنَّ ٱللَّهَ يُحِبُّ ٱلَّذِينَ يُقَٰتِلُونَ فِى سَبِيلِهِۦ صَفًّا كَأَنَّهُم بُنْيَٰنٌ مَّرْصُوصٌ Sourate As-Saff verset 4

Figure I.8. *L'exception de la règle de l'assimilation avec nasalisation*

5.1.2. *L'assimilation totale sans nasillement (الإدغام بدون غنة)*

On appelle cet idghâm : « kâmil » (الإدغام الكامل) car le noûn as-sâkina نْ ainsi que ses caractéristiques disparaissent entièrement lors de l'assimilation et cela concerne ces 2 lettres : ر – ل . [3]

Techniquement, dans cette règle du l'idghâm, le noûn as-sakîna نْ ou le tanwîn ً ٍ ٌ s'effacent et on passe directement à la lettre qui suit sans faire intervenir la ghounna. [3]

Il n'y a pas de ghounna et la lettre du lâm ل ou du râ ر, sont renforcées avec chadda ّ (الشَّدَّةُ). Le noûn est directement emboîté dans la lettre qui suit. [3]

Exemples

- L'assimilation avec la lettre «اللام»:

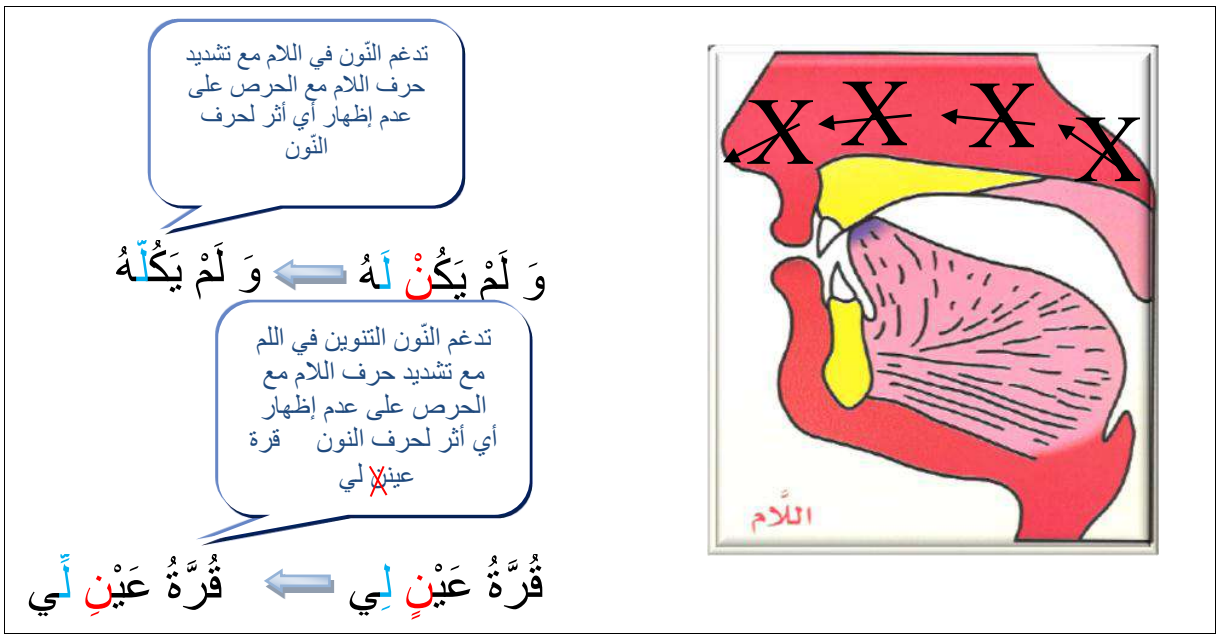


Figure I.9. *L'assimilation avec la lettre «اللام»*

- L'assimilation avec la lettre «الراء»:

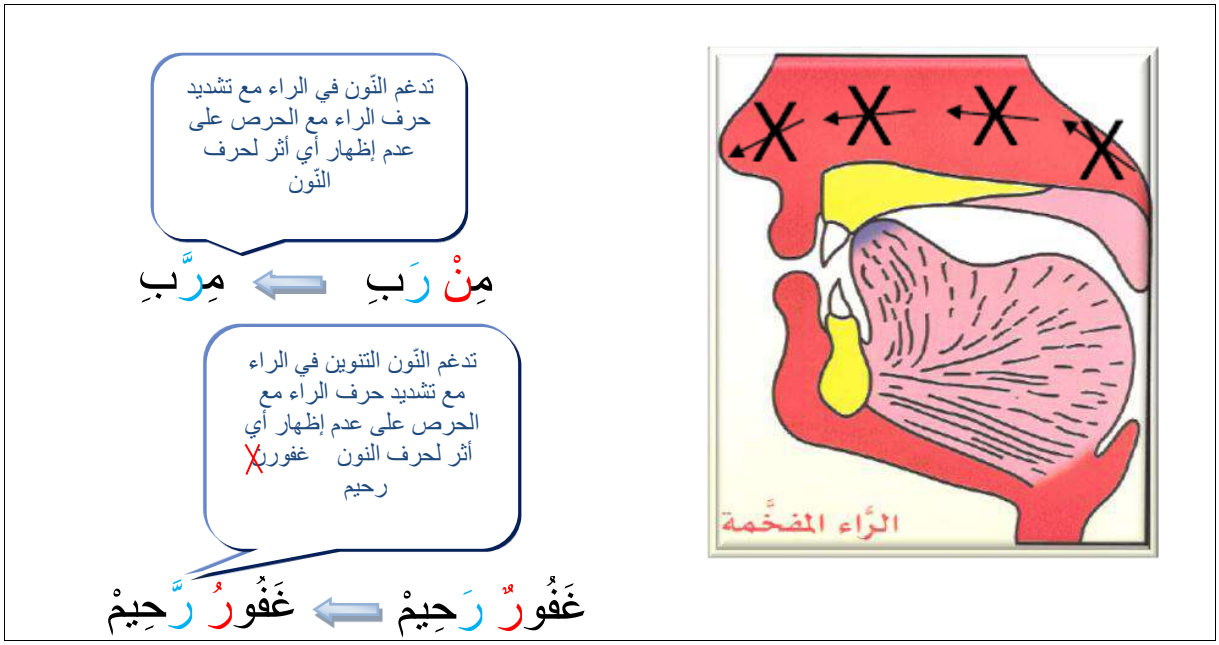


Figure I.10. *L'assimilation avec la lettre «الراء».*

6. Le traitement automatique de la parole

Le traitement automatique de la parole est un champ de recherche riche mais difficile. En effet, le signal de parole est complexe car sa structure résulte de l'interaction entre la production des sons et leur perception par l'oreille. Les chercheurs du domaine ont réussi malgré tout à réaliser quelques systèmes de reconnaissance ou de synthèse exploitables dans des conditions réelles. Les recherches dans ce sens sont toujours actives et nous assistons actuellement à des progrès liés à la reconnaissance de la parole continue multi-locuteurs opérant sur de grands vocabulaires et dans des milieux bruités. [6]

6.1. *Le signal de parole*

Le signal de parole est un signal acoustique produit par des vibrations des couches d'air. Les fluctuations de la pression de l'air produisent des variations de ce signal, en fonction du temps, qui peuvent être enregistrées de façon analogique ou digitale.

La parole est formée de phonèmes et de transitions entre ces phonèmes. Les phonèmes sont les éléments les plus brefs qui permettent de distinguer différents mots. Plusieurs formes du conduit vocal peuvent produire le même phonème. La forme que le conduit vocal prend pour la production d'un phonème, dans un contexte donné, est assez variable et surtout dépendante

de ce contexte. De ce fait, nous remarquons que les formes acoustiques associées à un phonème déterminé sont variables. Plusieurs types de phonèmes existent : les voyelles, les consonnes fricatives et les consonnes plosives, les nasales et les liquides.

Une grande difficulté de la reconnaissance de la parole réside dans la variabilité du signal vocal. Une même personne prononce rarement un mot exactement de la même façon. La vitesse d'élocution peut varier, la durée du signal est alors modifiée. Toute altération de l'appareil phonatoire peut modifier la qualité de l'émission (exemple : rhume, fatigue ...). De plus, la diction évolue dans le temps. La voix est modifiée au cours des étapes de la vie d'un être humain (enfance, adolescence, âge adulte ...). La variabilité interlocuteur est encore plus évidente. La hauteur de la voix, l'intonation, l'accent diffèrent selon le sexe, l'origine sociale, régionale ou nationale.

L'acoustique du milieu (milieu protégé ou environnement bruité), la qualité du microphone ou de la ligne téléphonique, les bruits de la bouche, les hésitations, les mots hors-vocabulaire sont autant d'interférences supplémentaires sur le signal de parole que le système de reconnaissance doit compenser.

L'aspect continu du signal de parole complique encore la tâche de reconnaissance. En effet, lorsqu'on écoute parler une personne, on perçoit une suite de mots, alors que l'analyse du signal vocal ne permet de déceler aucun séparateur. Le même problème de segmentation se retrouve à l'intérieur du mot lui-même. Celui-ci est perçu comme une suite de sons élémentaires.

Aussi, l'analyse du signal ne permet pas de découper en segments distincts le signal acoustique afin d'identifier les différents phonèmes qui le composent. [6]

6.2. *Extraction des paramètres*

Le signal de parole est complexe et redondant. Il possède une grande variabilité. Pour que le système de reconnaissance de la parole fonctionne efficacement, les informations caractéristiques et invariantes doivent être extraites du signal de parole. Cette procédure consiste à associer au signal de parole une ***série de vecteurs de paramètres généralement acoustiques***.

La phase d'extraction de caractéristiques doit être faite avec soin, car elle contribue directement aux performances du système global. Les codeurs les plus couramment utilisés sont le codage linéaire prédictif (Linear Predictive Coefficients « LPC »), les coefficients

cepstral (Mel Frequency Cepstre Coefficients « MFCC ») ou bien le codage linéaire prédictif perceptuel (Perceptual Linear Predictive PLP) .Le codage MFCC et le codage PLP ont la propriété d'intégrer des connaissances du modèle auditif humain

6.2.1. *Les paramètres MFCC (Mel Frequency cepstral Coefficients)*

Cette technique consiste à calculer les coefficients cepstraux sur une échelle en Mel qui se rapproche de la perception fréquentielle de l'oreille. Nous détaillons les différentes étapes de MFCC dans le chapitre II.

6.2.2 *Les paramètres LPC (Linear Predictive Coefficients):*

L'analyse par prédiction linéaire LPC est forte utilisée dans le codage de parole dans le but de réduire la redondance du signal vocal, ou pour extraire des paramètres pertinents pour la reconnaissance de parole. L'estimation des coefficients de la fonction de transfert du conduit vocal est faite en supposant connaître le signal d'excitation. Pour les sons non voisés, le signal d'excitation est un bruit blanc de moyenne nulle et de variance unité. Pour les sons voisés, cette excitation est une suite d'impulsions d'amplitude unité. La fonction de transfert du conduit vocal dans le domaine Z est donnée par :

$$\boldsymbol{H(z) = \frac{S(z)}{U(z)} = \frac{G}{1-A(z)}} \qquad \textbf{(I.1)}$$

$$\boldsymbol{A(z) = \sum_{k=1}^{p} a_k\, z^{-k}} \qquad \textbf{(I.2)}$$

Où A (z) est le prédicteur linéaire, (a_k) sont les coefficients de prédiction, S(z) est le signal de parole produit en sortie, U(z) est le signal d'excitation et G est un gain.

Le signal de parole S(n) à la sortie du modèle est donc représenté par la somme d'une combinaison linéaire des échantillons précédents et de la fonction d'excitation, tel que

$$\boldsymbol{S(n) = \sum_{k=1}^{p} a_k\, s(n-k) + Gu(n)} \qquad \textbf{(I.3)}$$

Le modèle de prédiction exploite le fait que les échantillons successifs du signal de parole sont corrélés ; d'où l'intérêt de ce modèle dans le codage de la parole dans le sens où il permet de représenter la parole juste par ses paramètres pertinents, sans redondance. Signalons également que les coefficients sont choisis de façon à minimiser l'erreur quadratique de prédiction sur chaque segment de la fenêtre d'analyse. [7]

6.2.3. *Les paramètres PLP (Perceptual Linear Prediction)*

La prédiction linéaire perceptuelle exploite les connaissances du système auditif humain pour paramétrer la parole en introduisant des mécanismes psycho acoustiques de l'oreille humaine.

L'analyse perceptive repose sur l'analyse par prédiction linéaire (LPC). Dans cette analyse, les coefficients sont calculés suite à la résolution d'équations obtenus par la transformée de Fourier inverse du module au carré de la transformée de Fourier du signal.

L'analyse perceptive introduit au niveau fréquentiel des bandes critiques. Cette intégration se fait avec un banc dont les fréquences centrales sont espacées linéairement selon l'échelle Bark. Cette intégration se justifie par le fait que l'oreille se comporte comme un banc de filtres. [8]

6.3. *Méthodes de reconnaissances*

6.3.1. *Les modèles de Markov cachés HMM (Hidden Markov Models)*

Les modèles de Markov cachés (Hidden Markov Models ou HMM) ont connu une grande importance depuis leur introduction en traitement de la parole. La plupart des systèmes de RAP utilisent les HMM pour modéliser les mots ou les unités élémentaires de la parole telle que les phonèmes, les syllabes. De nombreux travaux de recherches ont montré l'efficacité de ces modèles en reconnaissance de la parole continue ou isolée, indépendamment du locuteur pour de petits et grands vocabulaires, Nous détaillons les différentes étapes de HMM dans le chapitre II.

6.3.2. *Machine à vecteurs de Supports SVM (supports vectors machine)*

Les SVM sont des classificateurs binaires et statiques. C'est à dire qu'ils permettent de créer une surface de décision entre deux classes définies dans un même espace. Pour cela, ils construisent une frontière de décision par projection des caractéristiques provenant d'un espace d'origine dans un espace de caractéristiques de dimension supérieure (voir infini) dans le but de rendre les classes linéairement séparables.

L'hyperplan choisi est celui qui maximise la marge de séparabilité entre les deux ensembles de données. La sélection de l'hyperplan dans un espace de caractéristiques

nécessite d'évaluer un produit scalaire dans cet espace. Ce qui peut être très coûteux en temps et en complexité si l'espace est de très grande dimension. Heureusement, ce calcul n'est pas obligatoire grâce à une opération mathématique appelé kernel. Le kernel calcule le produit scalaire de deux points dans l'espace de dimension supérieur sans avoir à les projeter.

De plus, cette technique n'est pas appliquée seulement au SVM, mais aux groupes des machines à apprentissage appelés kernel machines. Toutes utilisent des kernels pour appliquer les données d'un espace d'entrée en un espace de caractéristiques de dimension supérieur. Ils sont capables de créer une séparation non linéaire d'un espace d'entrée. Il existe de nombreuses kernel machines, les plus connues Relevant Vector Machines, Bayes Point Machine, Gaussian processes, Least Squares SVM, méthodes de Boosting et Kernel Fisher Discriminant.

Les Kernel machines sont apparues comme des solutions viables et compétitives dans de nombreux domaines. Cependant, il existe de nombreux problèmes à leur utilisation pour la reconnaissance de la parole. [8]

6.3.3. *Déformation temporelle dynamique DTW (dynamic time warping)*

Cette méthode est la solution la plus simple à mettre en œuvre en reconnaissance de la parole. Elle consiste à effectuer une comparaison dynamique entre un vecteur de référence et un vecteur de test. On cherche à apparier l'ensemble des points du premier vecteur avec le second. On associe un poids à chaque appariement possible. L'utilisation de la programmation dynamique permet de trouver l'appariement minimisant ce coût. Les formes références sont obtenues par extraction des caractéristiques. La forme test peut avoir subit des transformations (rotations, translation, changement de taille). On utilise pour cela une méthode de mesure de similarité. La distance finale entre une forme de référence et une forme de test est calculée comme la somme des distances partielles entre vecteurs de paramètres le long du chemin optimal. Des poids liés à la suppression et à l'insertion d'informations, sont imposés pour calculer le meilleur chemin.

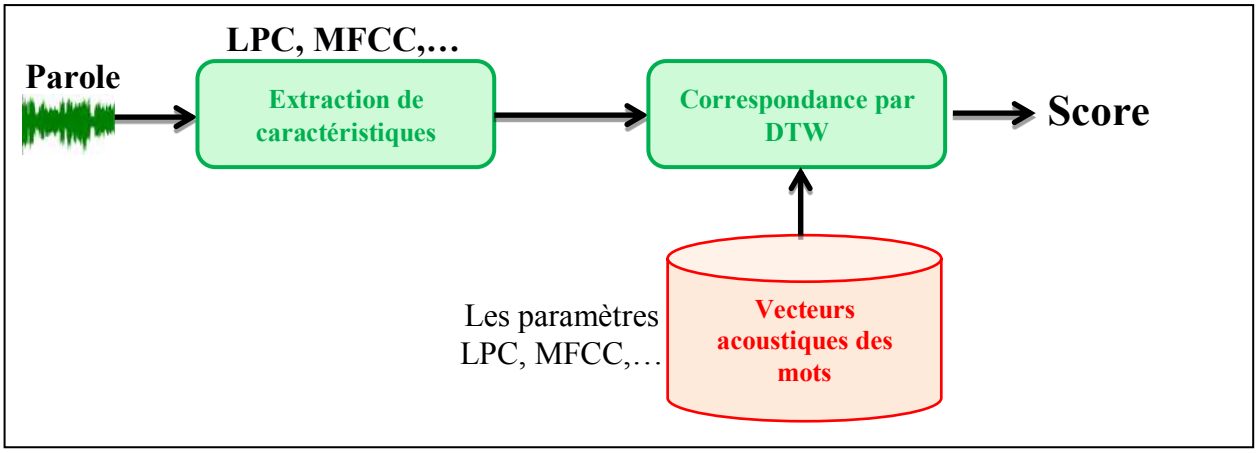


Figure I.11. *Schéma d'un système de reconnaissance basé sur la comparaison dynamique*

Cette méthode a l'avantage d'être simple et de nécessiter que peu de données d'apprentissage. Par contre, elle est limitée par la taille du vocabulaire à reconnaître et à la reconnaissance mono locuteur. [8]

6.3.4. *Les réseaux de neurones*

Les réseaux de neurones disposent de propriétés qui sont intéressantes du point de vue de la reconnaissance de la parole.

Ils peuvent apprendre des fonctions fortement non linéaires qui permettent de reconnaître une classe de forme et, simultanément, de rejeter les autres classes. Ils constituent des approximateurs de fonctions très puissants.

De plus, ils ne nécessitent aucune donnée statistique à l'initialisation contrairement aux modèles de Markov Cachés.

Enfin, ils présentent des structures parallèles régulières facilement implantables sous forme logicielle et matérielle. [8]

7. conclusion

A l'époque du Prophète ﷺ il n'y avait nul besoin de théoriser le tadjwîd car il s'agissait de leur propre langue. Avec l'arrivée des non-arabophones dans la communauté, puis l'évolution de la langue et des dialectes chez les arabophones, il a fallu très vite définir les règles qui régissent la langue et la récitation du Saint Qu'ran. [3] Ainsi l'arrivé de la technologie de l'information qui a le privilège d'être appliqué dans cette noble science pour donner un pouce pour apprendre la récitation comme Allah nous ordonne à lire.

Nous avons donc essayé dans ce chapitre de déterminer qu'est-ce qu'une science de Tadjwîd en particulier la règle de récitation « l'assimilation ». Nous avons également cité et définir les méthodes les plus utilisés dans une analyse acoustique et dans un système RAP, et expliqué comment ses technologies donnent un petit coup de main à l'apprenant pour devenir إن شاء الله un expert dans la récitation du noble Qu'ran.

Chapitre 2:

Analyse acoustique et apprentissage

1. Introduction

Le signal de la parole est l'un des signaux les plus complexes à caractériser ce qui rend difficile la tâche d'un système RAP. Cette complexité du signal de la parole provient de la combinaison de plusieurs facteurs, la redondance du signal acoustique, la grande variabilité inter et intra-locuteur, les effets de la coarticulation en parole continue et les conditions d'enregistrement. [1]

Pour surmonter ces difficultés, de nombreuses méthodes et modèles mathématiques ont été développés, parmi lesquels on peut citer : la comparaison dynamique, les réseaux de neurones artificiel (RNA), les machines à vecteurs supports (SVM), les modèles de Markov stochastiques et en particulier les modèles de Markov cachés (HMM). [1]

Ces méthodes et modèles travaillent à partir d'une information extraite du signal de la parole considérée comme pertinente. Cette extraction est effectuée par une analyse acoustique qui conduit à rassembler cette information sous le terme de ***vecteur de paramètres acoustiques*** dont la dimension et la nature sont déterminants pour accéder à de bonnes performances des systèmes de RAP. Les différents types de paramètres acoustiques couramment cités sont les coefficients : LPC, PLP, MFCC. [1]

Après avoir délimité les différentes difficultés rencontrées lors la conception du système RAP et proposer quelques solutions pour les surmonter, nous allons présenter dans ce chapitre la solution qui nous semble adéquate à notre problématique de recherche qui est le modèles de Markov cachés. Cette dernière travaille sur les extraits des signaux (le vecteur de paramètres acoustiques) effectuée par la méthode MFCC (*Mel Frequency Cepstral Coefficients*).

2. L'analyse acoustique par la méthode (MFCC)

De par sa complexité, l'information portée par le signal de parole ne peut, quelle que soit la tâche considérée, être utilisée dans sa totalité. C'est pourquoi l'utilisation d'un système automatique nécessite une sélection préalable des informations à exploiter pour la tâche qui lui est confiée. [9]

Cette phase correspond à la représentation acoustique du signal. L'analyse acoustique consiste à extraire du signal vocal un ensemble de ***paramètres pertinents*** dans le but de réduire la redondance du signal vocal pour une tâche de reconnaissance de la parole. Le nombre de ces paramètres doit rester raisonnable, afin d'éviter la nécessité d'un grand espace mémoire ce qui accroît le coût de calcul dans le module de décodage. Ces paramètres doivent être discriminants en rendant les sons de base facilement séparables. Ils doivent être robustes au bruit. [1]

Pour cette analyse, Nous utiliserons des coefficients cepstraux MFCC. Le calcul des paramètres acoustiques est réalisé par une chaîne de traitement selon les étapes illustré dans la figure suivante :

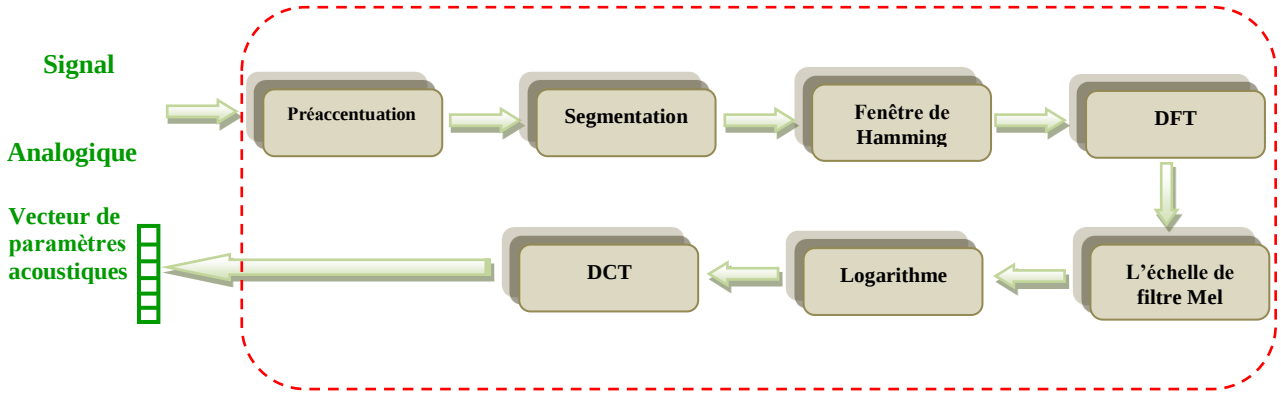


Figure 0.12. *La chaine de traitement de l'analyse acoustique MFCC.*

A l'entrée de la fonction de MFCC, on a le signal analogique qui peut être le son du locuteur (dans notre cas il s'agit de l'apprenant) ou un exemple d'apprentissage de la base de donnée, après avoir passé par les sept étapes illustrées dans la figure ci-dessus on obtient à la sortie un vecteur de paramètres acoustiques associé au son qui est représenté dans la (**figure II.22).**

2.1. *La préaccentuation*

La préaccentuation est considérée comme la première étape de MFCC. Elle fait la conversion du signal analogique vers un signal échantillonné. Les séquences d'échantillons x[n] sont obtenues à partir de signal continu x(t) en utilisant cette équation :

$$x[n] = x(nT) \qquad (0.4)$$

Tel que T est la période d'échantillonnage et « 1/T = fre_ech ». n représente le nombre d'échantillons, cette étape est faite dès qu'on enregistre le son ou lors de la lecture du son. Le signal échantillonné x[n] est, ensuite, pré-accentué afin ***de filtrer les hautes fréquences qui sont moins énergétiques que les basses fréquences***. Cette étape consiste à faire passer le signal échantillonné dans un filtre numérique à réponse impulsionnelle finie de premier ordre donné comme suit : [1]

$$H(z) = 1 - \alpha z^{-1} \quad \text{tq} \quad \alpha = 0.97 \qquad (0.5)$$

Ainsi, le signal pré-accentué est lié au signal échantillonné par la formule suivante :

$$Y[n] = x[n] - 0.97\, x[n-1] \qquad (0.6)$$

Le code en MATLAB est décrit comme suit :

```
%% ************************ La Préaccentuation ************************
%********filtrer le son enregisté
%********son_filtre = son_enregitre[n] - 0.97*son_enregitre[n-1]***
%pour enlever les hautes fréquences.
%pour but d' augmenter l'énergie du signal
son_filtre = filter([1 -.97], 1,son );
%wavplay(son_filtre,fre_ech);%----fonction pour écouter le son
```

Figure 0.2. *Le code MATLAB pour la tache de préaccentuation de la méthode MFCC.*

La figure ci-dessous montre l'effet de l'étape de préaccentuation sur un signal enregistré.

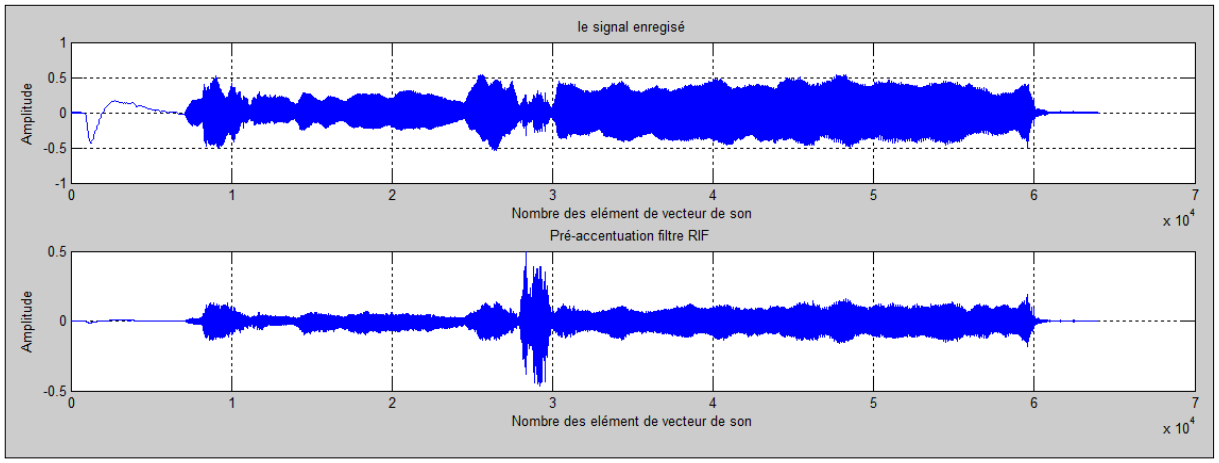


Figure 0.3. *la préaccentuation du signal enregistré « من يشاء ».*

2.2 *Segmentation*

Les méthodes du traitement de signal utilisées dans l'analyse du signal vocal (la transformer de Fourier) opèrent sur des signaux stationnaires, alors que le signal vocal est un signal non stationnaire. Afin de remédier à ce problème, l'analyse de ce signal est effectuée sur des trames successives de parole, de durée relativement courte sur lesquelles le signal peut en général être considéré comme quasi stationnaire. Dans cette étape de segmentation, le signal pré-accentué est ainsi découpé en trames de N échantillons de parole. En général N est fixé de telle manière à ce que chaque trame corresponde à environ ***20 à 30 ms*** de parole. Deux trames successives sont séparées de M échantillons correspondant à une période de l'ordre de la centi-seconde. [1]

2.2.1. *Le chevauchement*

Le chevauchement signifie que chaque nouvelle trame est faite pour contenir une partie de la trame précédente et une partie de la trame suivante. Il assure que le signal pré-accentué ne subit pas une perte de l'information lors de son découpage en trames. [10]

Le degré de chevauchement (généralement exprimé comme un pourcentage) décrit la quantité de la trame précédente qui se répète dans la trame suivante. (Exemple chevauchements de 25% et 50%). La figure qui suit explique le principe de segmentation avec un chevauchement de 50%. [10]

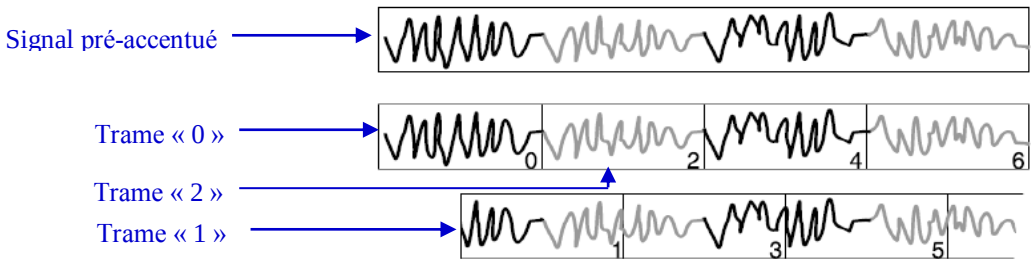


Figure II.4. *Le principe de segmentation avec un chevauchement de 50%.*

Pour ce qui concerne notre corpus d'étude la taille de la trame est égale à 20 ms. Quant au taux de chevauchement est égal à 25%. La figure suivante nous présente le code MATLAB permettant de fixer les paramètres nécessaire pour la tache de segmentation.

```
%% %%%%%%%%%%%%%%%%%%%%%%% La segmentation %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

% le nombre des elements du vecteur de trame est calculer par la formule
% suivante
taille_trame=floor(temps_trame/1000*freq_echantillonnage)+1;

% la valeur de chevauchement est egale a:
%(le taux de chevauchement * talle de trame).
chevauchement =floor(taux_chevauchement*taille_trame)+1;
```

Figure II.5. *Le code MATLAB pour la tache de segmentation de la méthode MFCC.*

Par la suite nous aurons un nombre de trames égal à Nbr trames (Nbr x la taille de la trame > la taille du signal) Ce qui signifie que notre signal s'accroit. Cela implique une discontinuité de ce signal. Ce dernier problème sera résolu par l'étape de fenêtrage.

2.3. *Le fenêtrage*

La segmentation du signal en trames produit des discontinuités aux frontières des trames. Dans le domaine spectral, ces discontinuités se manifestent par des lobes secondaires. Ces effets sont réduits en multipliant les échantillons de chaque trame par une fenêtre de pondération telle que la fenêtre de Hamming. [1]

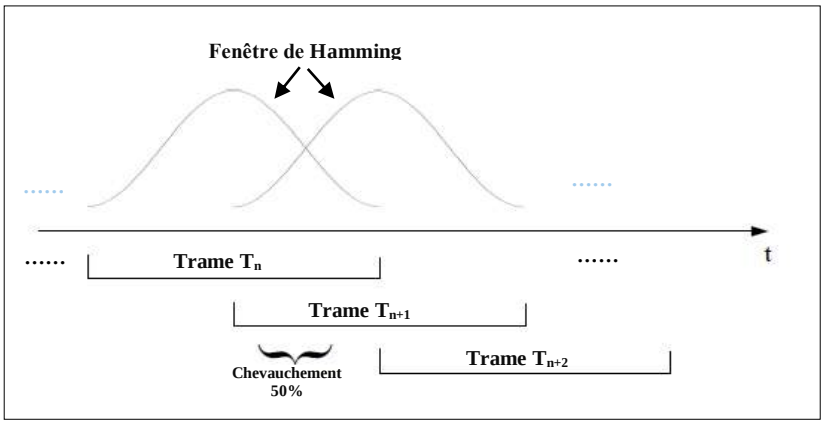


Figure II.6. *Le principe de Fenêtrage avec un chevauchement de 50%.*

« La fenêtre de Hamming est la plus convenable à la parole, car elle entraîne un minimum de distorsion spectrale du signal de parole, par rapport aux autres fenêtres (Atténuation du rapport du lobe principal au lobe secondaire est égale à –41 dB, c'est à dire que la concentration de l'énergie dans le lobe principal est égale à 99.96%) ». [13]

La fenêtre de Hamming est définie par le terme suivant :

$$W(n) \quad tq \quad 0 \leq n \leq (N-1) \qquad (0.7)$$

N représente le nombre des échantillons par trame. Le signal obtenu après le fenêtrage est y(n) tq :

$$Y(n)= x(n)*w(n) \quad 0\leq n\leq (N-1) \tag{0.8}$$

L'équation de Hamming est définie par l'équation suivante :

$$\begin{cases} 0.45 - 0.46\ \cos\left(\frac{2\pi n}{N-1}\right) & 0 \leq n \leq N-1 \\ 0 & ailleur \end{cases} \tag{0.9}$$

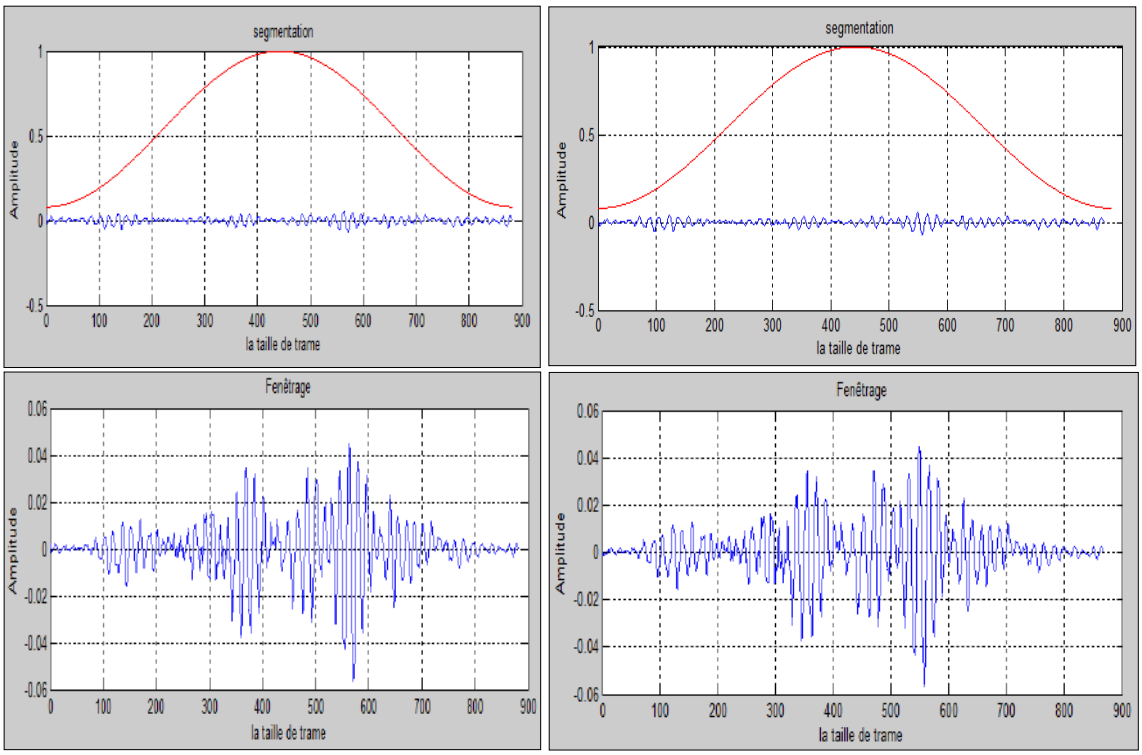
A la quelle nous pouvons associer le code MATLAB suivant :

```
%% ************************ La Fenetre de HAMMING ***************************
    a = .54;
    b = -.46;
    hamWindow = a + b*cos(2*pi*(0:taille_trame-1)/taille_trame);

%%************************************************************************%%
```

Figure 0.7. *Le code MATLAB associé à la fenêtre de Hamming.*

Pour une trame de taille 20 ms et un chevauchement de 0.25 (25%) nous arrivons aux résultats présentés par ces figures.



La trame numéro 1 *La trame numéro 15*

Figure II.8. *La segmentation et le fenêtrage de la trame n°1 et n° 15 par la fenêtre de Hamming.*

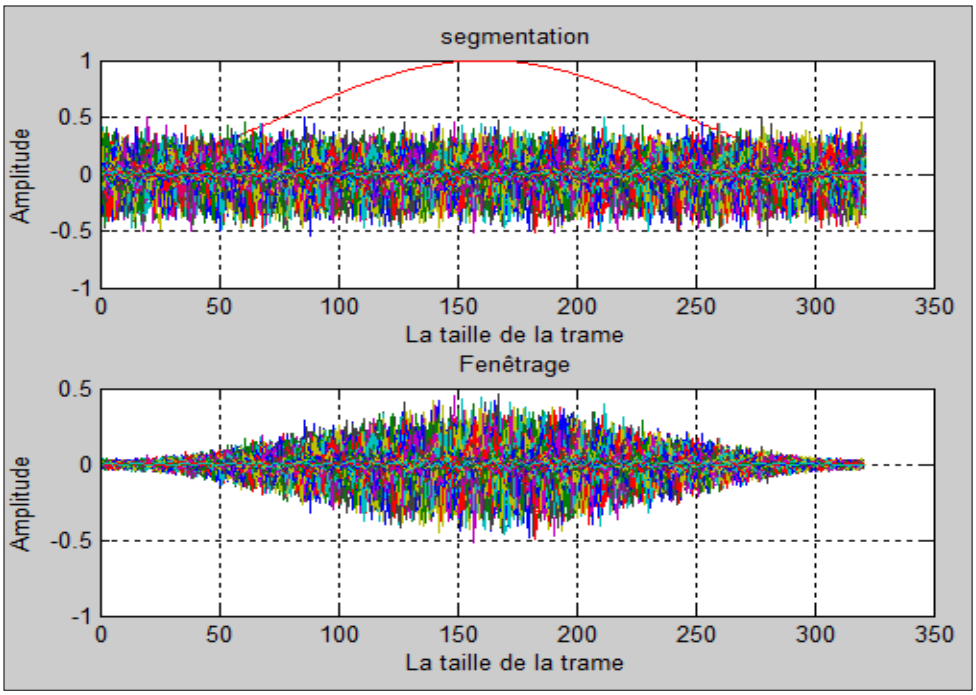


Figure II.9. *La segmentation et le fenêtrage de toutes les trames par la fenêtre de Hamming*

2.4. *La transformer de Fourier discrète*

La transformé de Fourier discrète est calculée à travers l'algorithme FFT (Fast Fourier Transform). L'objectif de cette transformé est de faire passer tous les trames de signal du domaine temporel vers le domaine fréquentiel, car le traitement dans le second est plus facile que dans le premier.

La commande Matlab est comme suit :

```
for start=1:nbr_trames
    fftData = zeros(1,2*taille_trame);
    %----------------- fenetre de haming -----------------
    fftData(1:taille_trame) = matrice_trames(:,start)'.*hamWindow;
    %---------------------- La DFT ----------------------
    fft_trame = abs(fft(fftData));
```

Figure II.10. *Le code MATLAB associé à la transformer de Fourier discrète.*

La fonction de la DFT est décrite par la formule suivante :

$$X_K = \sum_{k=0}^{N-1} x_k e^{-2\pi kn/N} \text{ tq n=0 ,1,2….N-1} \quad (0.10)$$

Le résultat est montré dans les figures suivantes :

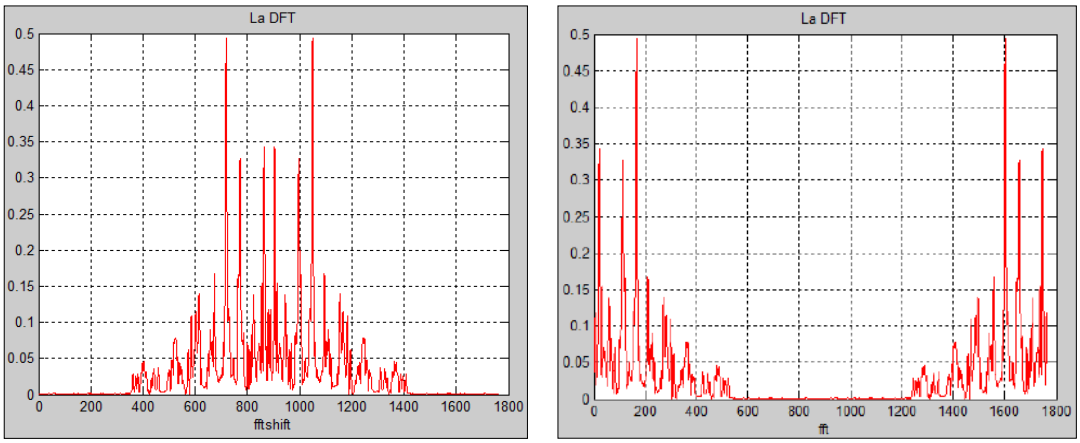


Figure II.11. *: La transformer de Fourier discrète de la quinzième trame.*

2.5. *Le banc du filtre MEL*

Cette technique consiste à calculer les coefficients cepstraux sur une échelle en Mel qui se rapproche de la perception fréquentielle de l'oreille.

Le banc de filtre Mel est considéré parmi les filtres auditifs[1], Les êtres humains perçoivent et font la différence beaucoup mieux des petits changements en hauts fréquences vers les basses fréquences que la machine. C'est la raison pour laquelle l'objectif de cette échelle est de découper le signal du domaine fréquentiel tout comme le fait parfaitement l'oreille humain.

La formule pour la conversion de fréquence à l'échelle Mel est:

$$M(f)= 2595 \log (1+ f/700) \quad (0.11)$$

Et la formule inverse pour la conversion de fréquence de l'échelle Mel vers les fréquences en Hertz est :

$$M^{-1}(m) =700(\exp(m/2595)-1) \quad (0.12)$$

Le banc du filtre Mel est un ensemble de 20-40 filtres triangulaires à une surface d'unité, que nous appliquons sur toutes les trames du signal. Il est caractérisé par ses deux fréquences basse et haute, ses dernières sont espacés linéairement jusqu'à 1 KHz, puis logarithmiquement

[1] Le système auditif réalise une analyse fréquentielle du signal à l'aide d'une série de filtres passe-bandes. Ces filtres, appelés filtres auditifs, se recouvrent continûment tout le long de la plage des fréquences audibles.

jusqu'aux fréquences maximales. Chaque banc du filtre (triangle) à son propre bas et haute fréquence.

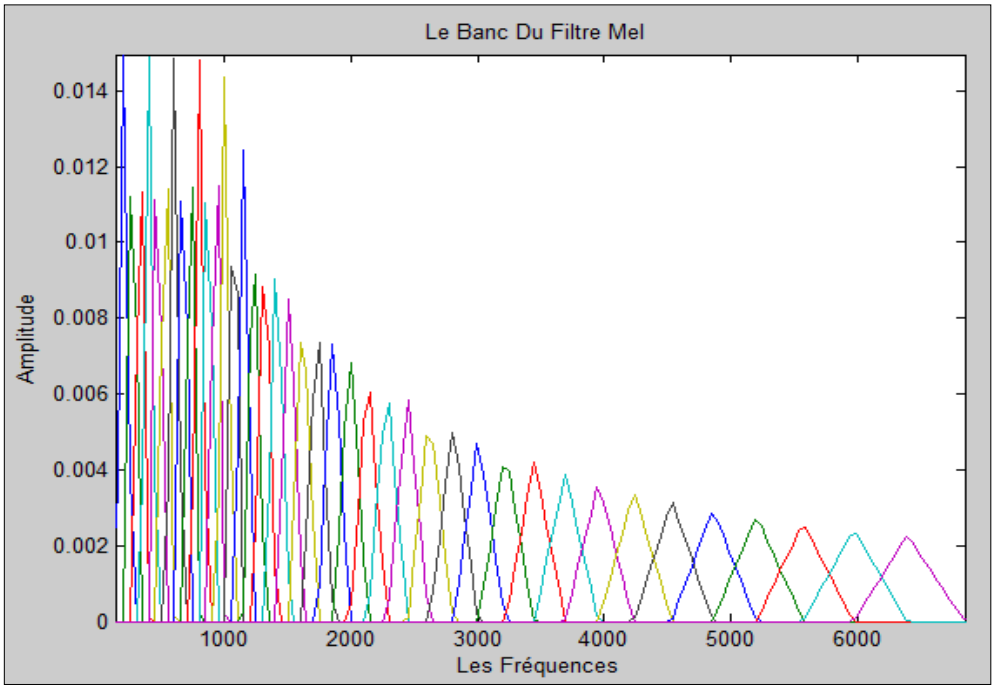


Figure 0.132.: *Le banc du filtre MEL.*

2.5.1. *La réalisation du banc du filtre MEL*

Pour ce qui concerne notre corpus d'étude nous avons choisi un banc de filtre qui est caractérisé par la liste des paramètres suivants [2]:

Tableau II.1. *Les paramètres du banc du filtre MEL.*

Paramètres	Valeurs
Filtre linéaire	13
Filtre logarithmique	27
La basse fréquence	133.3333
La fréquence intermédiaire	1000
La haute fréquence	6538

Pour obtenir les bancs de filtres présentés dans la **(figure II.12)** il faut d'abord calculer les 13 fréquences qui sont séparées linéairement par la formule suivante :

```
%les fréquences linéaires
freq_lin=linspace(freq_bas,inter_freq,nbf_lin+1)
```

Figure II.13. *: Le code Matlab pour réaliser 13 fréquences séparées linéairement*

La commande « linspace » permet de créer un vecteur de 14 valeurs entre la valeur freq_bas=133.3333Hz et la valeur inter_freq=999.8Hz.

Et aussi calculer les 27 fréquences logarithmiques qui sont séparé par une échelle logarithmique par la formule suivante :

```
%les fréquences logarithmiques
freq_log=logspace(log10(freq_lin(nbf_lin)),log10(freq_haut),nbf_log+3);
```

Figure II.14. *: Le code Matlab pour réaliser 27 fréquences séparées logarithmiquement.*

La concaténation des deux vecteurs nous donne un vecteur de 42 valeurs.

```
% la concatination des vecteurs de fréquences linéaires et logarithmiques
vec_freq(1:13)= freq_lin(1:13);
vec_freq(14:nbr_Totale_Filtres+2)=freq_log(2:30);
```

Figure 0.145.: *La concaténation des deux fréquences linéaires et logarithmiques.*

```
vec_freq =

   1.0e+03 *

  Columns 1 through 9

    0.1333    0.2000    0.2667    0.3333    0.4000    0.4667    0.5333    0.6000    0.6667

  Columns 10 through 18

    0.7333    0.8000    0.8667    0.9333    0.9998    1.0709    1.1471    1.2288    1.3162

  Columns 19 through 27

    1.4099    1.5102    1.6177    1.7329    1.8562    1.9883    2.1298    2.2814    2.4437

  Columns 28 through 36

    2.6177    2.8040    3.0035    3.2173    3.4463    3.6915    3.9542    4.2357    4.5371

  Columns 37 through 42

    4.8600    5.2059    5.5764    5.9733    6.3984    6.8538
```

- La première fréquence linéaire
- La dernière fréquence linéaire
- La première fréquence logarithmique
- La dernière fréquence logarithmique

Figure II.16. : *Les fréquences linéaires et logarithmiques du banc du filtre MEL.*

Donc on peut déduit que l'espacement linéaire et logarithmique sont égale respectivement à :

```
%% espacement linéaire et espacement logarithmique
esp_lin=freq_lin(2)-freq_lin(1)
esp_log=10^(log10(freq_log(2))-log10(freq_log(1)))
```

Figure II.17. : *Le calcul de l'espacement linéaire et logarithmique.*

D'après les deux commandes et la **(figure II.17)** nous remarquons que l'espacement linéaire est fixe dans les 13 premiers filtres au contraire au 27 derniers filtres ou l'espacement s'accroit avec le nombre de filtres.

Comme nous avons dit dans la définition du banc de filtre Mel, chaque triangle à une surface d'unité. Pour tracer ce triangle nous avons proposé cet exemple :

On veut tracer un triangle d'une surface unité avec une base [1 : 5] alors son sommet = 2/(5-1) pour avoir la surface = 1.

```
t=1:0.001:5;
temps_bas=t(1);
temps_haut=t(length(t));
temps_central=(temps_bas+temps_haut)/2;
hateur_triangle=2/(temps_haut-temps_bas);

%% le front montant du triangle
s1=(t>temps_bas & t<=temps_central).*hateur_triangle...
.*(t-temps_bas)/(temps_central-temps_bas);
subplot(311),plot(t,s1)

%% le front déscendant du triangle
s2=(t>temps_central & t<temps_haut).*hateur_triangle...
.*(temps_haut-t)/(temps_haut-temps_central);
subplot(312),plot(t,s2)

%% le tracé du triangle
subplot(313),plot(t,s1+s2)
```

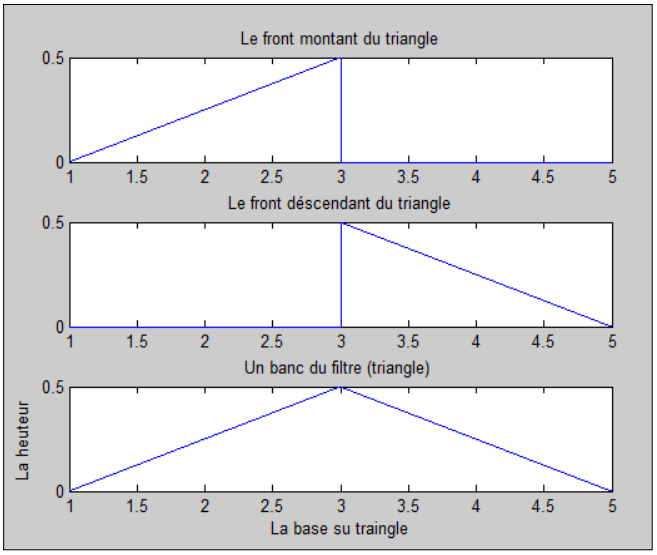


Figure 0.158. : *Le code Matlab pour tracer un triangle à surface unité.*

Maintenant, nous réalisons notre banc de filtres à partir du vecteur de fréquences. Le premier banc de filtres débutera au premier point, atteindre son sommet au second point, puis

retour à zéro à la 3ème point. Le deuxième banc de filtres commencera au 2ème point, atteindre son maximum à la 3ème point, puis retour à zéro à la 4ème point … etc. et nous traçons notre banc du filtre de la **(figure II.12)**.

```
for num_banc=1:40;
    B_D_Filtre_Mel(num_banc,:) = ...
  (axe_freqs > lower(num_banc) & axe_freqs <= center(num_banc)).* ...
   hateur_triangle(num_banc).*(axe_freqs-lower(num_banc))/(center...
   (num_banc)-lower(num_banc))+(axe_freqs > center(num_banc) &...
   axe_freqs < upper(num_banc)).*hateur_triangle(num_banc).*...
   (upper(num_banc)-axe_freqs)/(upper(num_banc)-center(num_banc));
end
plot(axe_freqs,B_D_Filtre_Mel')
```

Figure II.19. *: Le code Matlab pour réaliser le banc du filtre MEL.*

2.6. *La transformée en cosinus discrète (DCT)*

Sur une trame d'analyse du signal, les coefficients MFCC sont calculés à partir des énergies issues d'un banc de filtres triangulaires en échelle de fréquence Mel. Les coefficients cepstraux (en général sont choisis entre 10 et 15) peuvent être calculés directement en appliquant la transformée en cosinus discrète sur le logarithme des énergies sortant d'un banc de 40 filtres. [1]

La formule de DCT est décrite comme suit :

$$C_k = \sum_{i=1}^{M} log(E_i) \cdot cos\left[\frac{\pi k}{M}\left(i - \frac{1}{2}\right)\right] \qquad \text{k=0, …, d} \leq \text{M} \qquad \textbf{(0.13)}$$

Ei : est l'énergie de chaque trame.

```
%% ********************************** DCT ****************************************
% Les 13 coefficients cepstraux peuvent être calculés directement en
% appliquant la transformée en cosinus discrète sur le logarithme des
% énergies sortant d'un banc de 40 filtres

MFCC_DCT = 1/sqrt(nbr_Totale_Filtres/2)*cos((0:(nbr_D_coeff_cepstraux-1))' * ...
           (2*(0:(nbr_Totale_Filtres-1))+1) * pi/2/nbr_Totale_Filtres);
MFCC_DCT(1,:) = MFCC_DCT(1,:) * sqrt(2)/2;
```

Figure II.20. *: Le code Matlab le calcul de la fonction DCT.*

```
%---------------------- Le LOG ----------------------
fft_trame_log = log10(B_D_Filtre_Mel * fft_trame');

%---------------------- La DCT ----------------------
 Cepstrum(:,start+1) = MFCC_DCT* fft_trame_log
```

Figure II.21. *: Le code Matlab pour calculer l'énergie d'une trame.*

L'énergie est calculée par la multiplication du banc de filtre * la transformé de Fourier discret.

Et à la fin nous obtenons les coefficients cepstraux comme montre la figure suivante :

```
Cepstrum =

  Columns 1 through 9

  -19.4965  -20.1808  -20.3343  -19.5992  -19.7794  -19.1957  -19.1183  -20.0020  -19.7457
    0.3099    0.0955    0.5280    0.2650    0.9881    0.9542    0.6040    0.1541    0.3716
    0.6756    0.4002   -0.2565   -0.1405    0.5103    0.5048    0.5400    0.4465    0.7670
    3.3862    3.3645    3.3250    3.4872    3.4659    3.8550    2.9311    2.8725    3.4362
   -1.3947   -1.2528   -1.6622   -1.6736   -1.2567   -1.2485   -0.5566   -0.8083   -0.4456
    0.0437    0.0077   -0.0313   -0.1185   -0.0060    0.1072    0.1944    0.4353   -0.1129
   -0.0855   -0.4011    0.2983    0.0135    0.2049    0.2009    0.3083    0.2020    0.1022
   -0.0390   -0.3407   -0.6450   -0.5225   -0.5232   -0.3149    0.1655    0.1548   -0.0431
   -0.4467   -0.4466   -0.8941   -0.8913   -0.9674   -0.9182   -0.8664   -0.7743   -0.2803
    0.7211    0.3948   -0.1152    0.3480    0.4527    0.6272    0.0594    0.2157   -0.1921
   -0.5092    0.3148   -0.0398   -0.2923   -0.1531   -0.1474   -0.2176   -0.2640    0.1471
   -0.2132   -0.0480   -0.7273   -0.6062   -0.2315   -0.6105   -0.3779   -0.4476   -0.4608
    1.1060    0.4125    0.2685    0.3651    0.6937    0.8202    0.5804    0.6454    0.2094

  Columns 10 through 18

  -19.1991  -19.4330  -20.3566  -19.6229  -19.4678  -19.7964  -20.3703  -20.2916  -19.2573
    0.8890    0.7713    0.5776    0.6141    0.7575    0.5585    0.5128    0.9253    1.1679
    1.1326    0.2537    0.2546    0.7515    0.7507    0.5022    0.0361    0.6354    0.3817
    3.3179    3.3463    3.1624    3.3208    2.8241    2.5842    2.3042    2.8102    2.8893
   -1.5735   -0.8210   -1.4172   -0.8539   -0.0313   -0.4320   -0.3264   -0.1168   -0.2024
    0.1803    0.8676    0.3677    0.0141    0.0064    0.5879   -0.1571   -0.1042   -0.1924
    0.1743   -0.2619   -0.5646    0.0877   -0.0162    0.5759    0.7003    0.5313    0.5607
    0.1213   -0.7953   -0.3763   -0.4972   -0.4759   -0.5942   -0.2435   -0.4494   -0.7970
```

Figure 0.22. *: Les coefficients spectraux du signale « من يشاء ».*

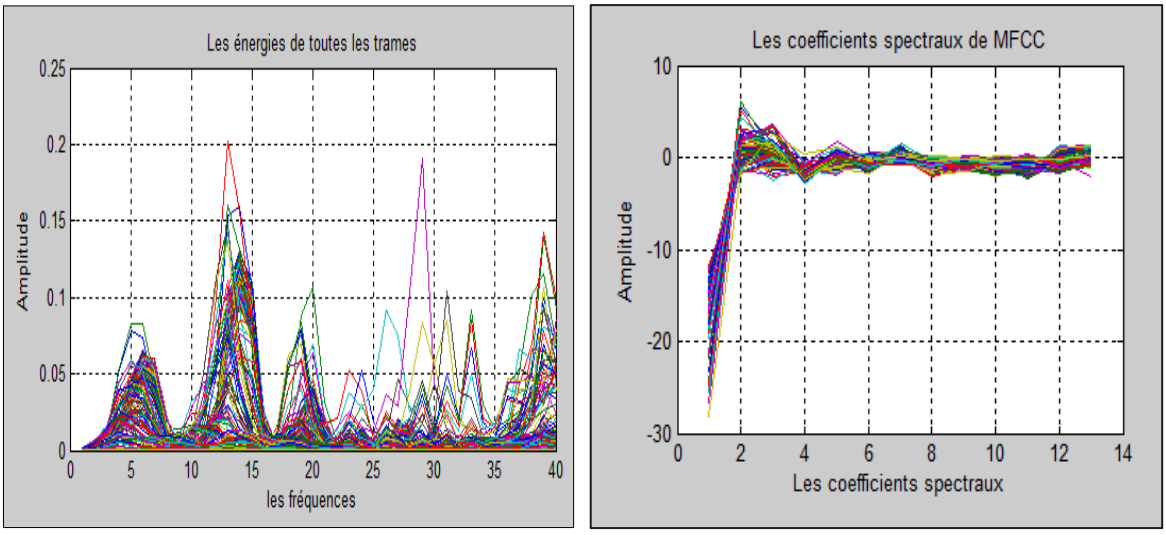


Figure II.23. *: Les énergies et les coefficients cepstraux de toutes les trames.*

Cette partie est la dernière étape de calcul des MFCC. Les coefficients spectraux de MEL sont regroupés dans un vecteur de paramètres acoustiques. Ses vecteurs sont alors considérés comme une entrée pour le prochain processus, qui se préoccupe d'entrainer ses vecteur pour la tache de la reconnaissance.

3. Les chaines de Markov cachées

Les processus réels produisent généralement des sorties en forme de signaux. Ces derniers sont, soit de nature discrète (caractères d'un alphabet fini ou des vecteurs quantifiés parmi un ensemble de valeurs, …), soit de nature continue (échantillon de parole, mesure de température, …). [11]

Un problème d'intérêt fondamental est de caractériser ces signaux en termes de modèle de signaux. [11]

A ce stade deux possibilités de modélisation se proposent à nous. Le modèle déterministe et le modèle stochastique, le modèle déterministe exploite généralement quelques propriétés du signal en vue de sa modélisation, comme par exemple la forme du signal (sinusoïde, somme d'exponentielles, amplitude, ...). La modélisation stochastique quant à elle, elle essaie de déterminer les caractéristiques statistiques du signal. On parle dans ce cas spécialement de fonctions de distribution de probabilité (gaussienne, poisson, …). [11]

Pour des applications de traitement de la parole, les deux modèles ont procuré d'excellents résultats. Dans notre corpus nous allons nous concentrer sur un seul type de modélisation stochastique, en l'occurrence les modèle de Markov cachés (HMM). [11]

Les modèles de Markov cachés, introduits fin des années 60, début des années 70, sont devenus la solution par excellence aux problèmes de la reconnaissance de la parole. En effet, ces modèles sont riches en structures mathématiques et par conséquent peuvent être utilisés dans un large domaine d'applications. En outre, ces modèles donnent de remarquables résultats en pratique quand ils sont correctement appliqués. [11]

3.1. *Définition de la chaine de Markov*

Un modèle HMM est défini comme un ensemble d'états, chacun d'entre eux associé à une distribution de probabilité. Les transitions entre les états sont régies par un ensemble de probabilités appelées *probabilités de transition* Dans un état particulier, un résultat ou *observation* peut être généré conformément à la distribution de probabilité associée. Par

opposition à un modèle de Markov classique où l'état est directement observable par un observateur externe, dans un modèle HMM, l'état n'est pas directement observable et seulement des variables influencées par l'état le sont. Les états sont donc cachés, d'où le nom de modèle de Markov caché. [11]

On retrouve bien le modèle de Markov : l'état début, les deux états état 1 et état 2 et les transitions avec leur probabilité associée. [16]

On y a ajouté deux observations : « A » et « B ». Chaque état peut émettre chacune des observations avec une certaine probabilité que nous appelons "probabilité d'émission". Cette probabilité peut éventuellement être nulle, ce qui signifie que l'état ne peut pas émettre l'observation concernée. [16]

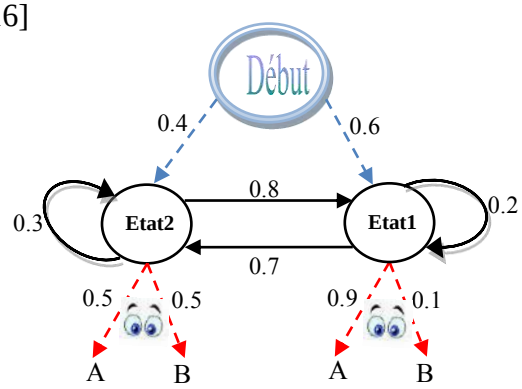


Figure II.24. *Schéma descriptif du model de Markov caché.*

Dans notre exemple, l'état 1 à 50% de chances (0.5) d'émettre un "A" et 50% de chances d'émettre un "B", tandis que l'état 2 à 90% de chances (0.9) d'émettre un "A" et 10% (0.1) d'émettre un "B".[16]

Comme pour les transitions partant d'un état, la somme des probabilités d'émission d'un état doit toujours être égale à 1. [16]

Un modèle HMM est défini par les éléments suivants :

- **N** : nombre d'état de la chaine de Markov.
- **pi0** : le vecteur de la distribution de la probabilité initiale.
- **A** : la matrice de transition.
- **X** : la matrice d'observation.
- **mu (μ)**: la matrice moyenne de tous les états pour les probabilités d'émission, ***mu (i, :)*** est le vecteur moyen correspond à le $i^{ème}$ état de HMM.
- **Sigma (∑)** : la matrice de covariance de tous les états pour la probabilité d'émission.

3.2. *Le principe de fonctionnement de HMM*

Le HMM est décrit par trois processus on les cite comme suit :

a. Apprentissage / modélisation

On associé dans cette étape à chaque échantillons dans la base de données un modèle HMM λ= (A, pi0, mu, sigma). [2]

b. Identification :

Cette étape consiste à comparer la matrice d'observation (qui est obtenus par une analyse acoustique MFCC) avec tous les modèles HMM de la base de données afin de trouver le meilleur score ***(le maximum de vraisemblance)*** ce dernier identifié le signal à reconnaitre. [2]

c. Vérification

Cette étape consiste a jugé si le mot a été bien prononcé ou non, ce jugement est basé sur ***un seuil*** si le score est supérieur à cette variable le signal est reconnu si non le signal ne sera pas reconnu. [2]

Le schéma suivant explique ces trois processus :

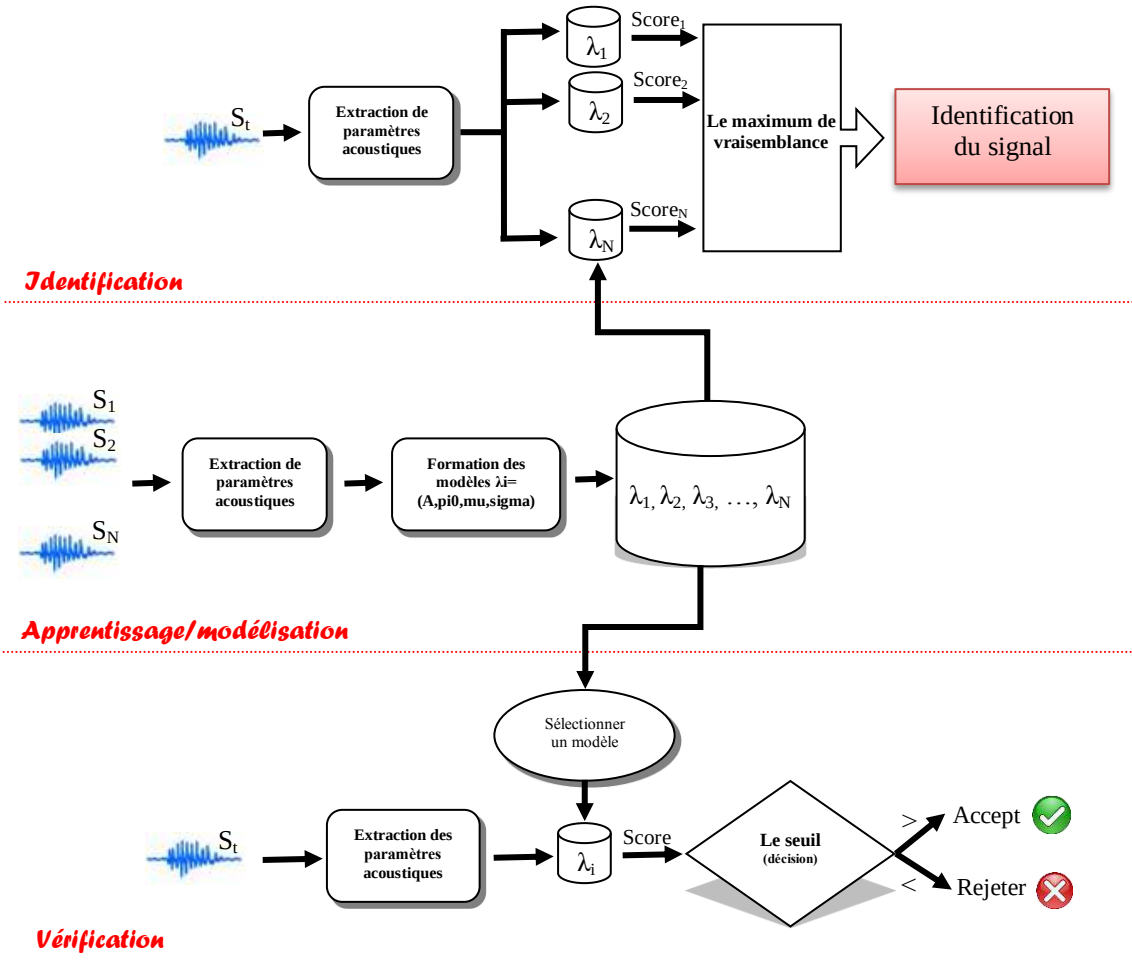


Figure II.25. *Schéma descriptif du système RAP « HMM ».*

3.2.1. Apprentissage/modélisation

La figure suivante montre la chaine de prétraitement de cette étape :

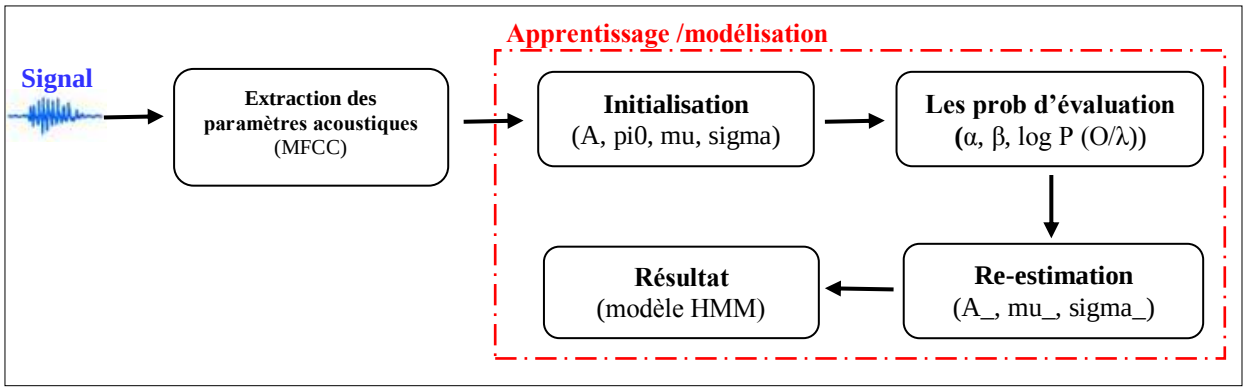


Figure 0.26. : *La chaine de prétraitement du processus d'apprentissage.*

3.2.1.1. L'initialisation

Comme dans tout algorithme de maximisation, le problème du choix de la valeur initiale est crucial. Dans cette étape consiste à l'initialisation de tous les variables (la matrice de transition A, le vecteur de la distribution de la probabilité initiale pi0, la matrice moyenne de tous les états pour les probabilité d'émission mu, la matrice de covariance de tous les états pour les probabilité d'émission sigma). [14]

- **Le nombre d'état (N) de la chaine de Markov :**

Le choix du nombre d'états de la chaine de Markov n'a pas un critère précisé. Il est calculé selon la nature du problème, dans certain problème l'application nous exige de modifier le nombre d'état jusqu'à arriver au meilleur résultat. Dans le cas des mots isolé un critère est considéré pour le calcul de nombre d'état selon la description phonétique du son, il est généralement entre 4 à 12 états. [15]

- ***La modélisation d'un signale de parole par un modèle de Markov caché:***

La modélisation d'un signal par un modèle de Markov se fait par l'intervention d'une analyse acoustique qui est déjà expliquée dans la première partie de ce chapitre. Cette analyse nous génère une matrice de trames ou chaque trame contient 13 coefficients associés à elle.

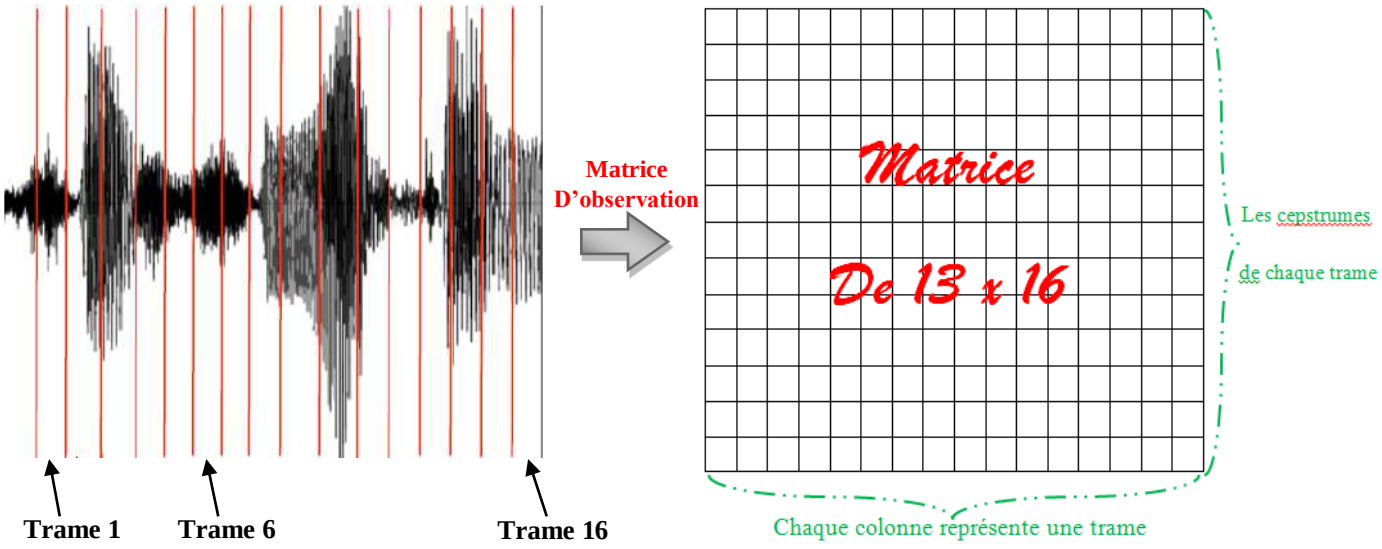
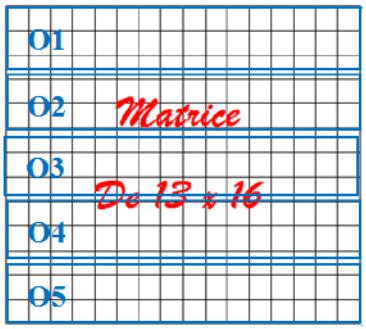


Figure II.27. : *La matrice des coefficients cepstreaux du signal.*

La matrice des coefficients spectraux est ensuite considérer comme une observation {O} dans le processus de Markov. Où elle se devise en séquences d'observation comme montre le schéma suivant, si on veut modéliser le signal précède par 5 états, on devise notre matrice à 5 régions ou chaque région forme une séquence d'observation O= {O1, O2, O3, O4, O5}.



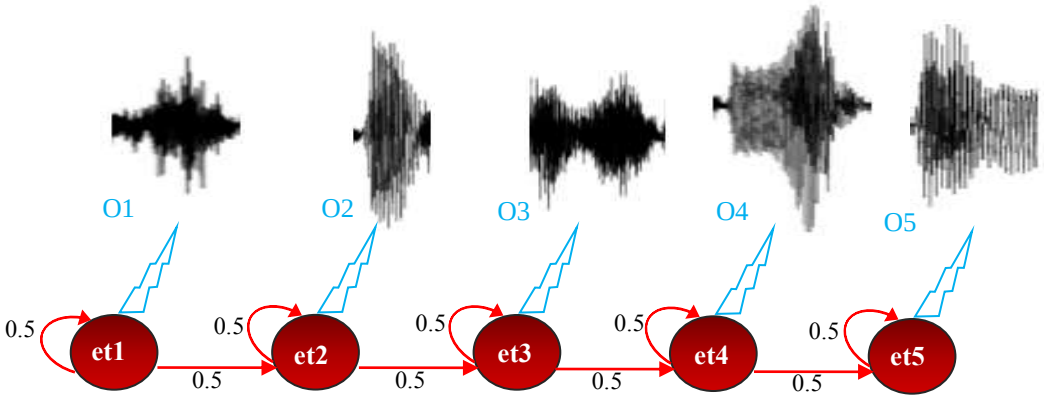


Figure 0.28. : *La modélisation du signal par une chaine de Markov caché.*

➢ ***La matrice de transition A :***

Les transitions entre les états sont régies par un ensemble de probabilités appelées probabilités de transition. Ces derniers sont regroupés dans une matrice de transition A (NxN). Ou aij représente la probabilité de passer de l'état i vers l'état j.

On utilise le modèle gauche-droite car il est le model utilisé pour la représentation de la parole. Pratiquement, on utilise deux transitions pour décrire un état. La matrice est définie comme suit :

$$\begin{cases} aij \neq 0 \; pour \; j = i \; et \; j = i + 1. \\ aij = 0 \; ailleur. \end{cases} \quad \textbf{(0.14)}$$

Dans ce modèle les probabilités de transitions forcent les observations de commencer dans le premier état ensuite d'aller vers le deuxième état jusqu'à l'arriver au dernier état. [12]

Nous avons initialisé nos matrices aléatoirement par l'utilisation de la fonction randn du Matlab.

```
%phase 1: INITIALISATION

% La matrice de transition (il faut que la somme de chaque ligne=1 )
% et il faut qu'elle soit triangulaire supérieure
A=abs(randn(1,2));
A=A./(sum(A)*ones(size(A)));
A=diag(ones(1,N).*A(1))+diag(ones(1,N-1).*A(2),1);
A(N*N)=1
```

Figure II.29. ***:*** *Le code Matlab pour initialiser la matrice de transition.*

➢ ***Le vecteur de la distribution de la probabilité initiale pi0 :***

Ce vecteur est initialisé en affectant la valeur 1 au premier état et 0 pour les autres (nous avons considéré qu'au début de chaque observation nous sommes dans l'état 1).

```
% Le vecteur de la distribution de la probabilité initiale
pi0 = [1 zeros(1,N-1)]
```

Figure 0.30. ***:*** *Le code Matlab pour initialiser le vecteur de la distribution de la probabilité initiale.*

➢ ***La matrice moyenne (mu) et covariance (sigma) :***

Les observations sont modélisées par des distributions gaussiennes (une matrice des moyennes et une matrice des covariances). Ces derniers sont initialisés par la fonction hmm_mint.

```
% Appelle de la fonction d'initialisation des paramatres mu et sigma
% Nous avons utilisé le toolbox publié sur ce site
% http://perso.telecom-paristech.fr/~cappe/Code/H2m/

[mu,sigma]=hmm_mint(X,st,N,1)
```

Figure II.31. ***:*** *Le code Matlab pour initialiser les matrices mu et sigma.*

Où X est la matrice d'observation, st est le vecteur des séquences et N le nombre d'état de la chaine markovienne.

3.2.1.2. Les probabilités d'évaluation

Ce problème se pose notamment quand nous avons, par exemple, plusieurs HMMs décrivant différents systèmes, et une séquence d'observations. Nous voulons ainsi connaître quel est le HMM ayant la plus forte probabilité d'avoir généré cette séquence. En d'autres termes, pour un modèle λ = (pi0, A, B) et une séquence d'observations O= {o1, o2,…., oT}, nous avons à calculer la probabilité p (O|λ). Un calcul de cette probabilité implique un nombre d'opérations de l'ordre de NT. Heureusement, une autre méthode, ayant une complexité inférieure, existe. [12]

Cette méthode utilise une variable intermédiaire appelée variable "avant" ou forward, d'où le nom de l'algorithme Forward (ou avant). [12]

Algorithme Forward : Cet algorithme est utilisé pour calculer la probabilité d'une observation de longueur T : [12]

$$\mathbf{O= \{o1, o2,…., oT\}} \qquad \mathbf{(0.15)}$$

Avec chaque (o) est un élément de l'ensemble observable. La variable intermédiaire $\alpha_t(i)$ est définie comme la probabilité de la séquence d'observation partielle Ot = {o1, o2, …, ot}, t ≤ T, qui se termine à l'état i. Les probabilités intermédiaires (ou partielles) sont calculées de manière récursive en calculant premièrement ces probabilités pour tous les états à t = 1. [12]

$$\boldsymbol{\alpha_1(j) = pi0(j).b_j(1),\ \ pour\ 1 \le j \le N} \qquad \mathbf{(0.16)}$$

Ensuite, pour chaque instant, t = 2, ..., T, les probabilités partielles sont calculées pour chaque état par la relation récursive suivante :

$$\boldsymbol{\alpha_{t+1}(j) = \sum_{i=1}^{N}(\alpha_t (i)\alpha_{ij})b_j (t),\ \ pour\ 1 \le j \le N,\ \ 1 \le t \le T-1} \qquad \textbf{(0.17)}$$

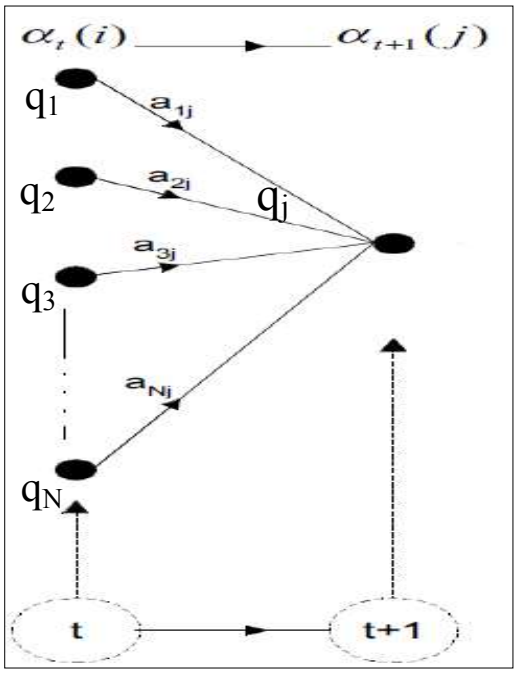


Figure II.32. : *Le principe de fonctionnement de l'algorithme Forward.*

Avec cette relation, nous pouvons alors calculer la probabilité intermédiaire à l'instant T pour chaque état j, α_T (j). Et finalement, la somme de toutes les probabilités partielles à l'instant T fournit la probabilité requise :

$$\boldsymbol{p(O|\lambda) = \sum_{i=1}^{N} \alpha_T (i)} \qquad \textbf{(0.18)}$$

Pour récapituler, chaque probabilité partielle (à l'instant t > 2) est calculée à partir de tous les états précédents. De façon similaire, nous pouvons définir une variable « arrière » ou backward βt (i) comme la probabilité de la séquence d'observation partielle {ot+1, ot+2, ..., oT}, étant donné que l'état courant est i. Pour calculer les βt(i), il existe aussi, comme pour les αt(i), une relation récursive : [12]

$$\boldsymbol{\beta_t(i) = \sum_{j=1}^{N}(\beta_{t+1} (j)\alpha_{ij})b_j (t+1),\ \ pour\ 1 \le j \le N,\ \ 1 \le t \le T-1} \qquad \textbf{(0.19)}$$

Avec :

$$\boldsymbol{\beta_T(i) = 1,\ \ pour\ 1 \le i \le N.} \qquad \textbf{(0.20)}$$

Si nous cherchons un lien entre les deux variables intermédiaires βt(i) et αt(i), nous pouvons remarquer que :

$$\boldsymbol{\alpha_t(i)\beta_t(i) = p(O, o_t = i|\lambda),\ \ pour\ 1 \le j \le N,\ \ 1 \le t \le T} \qquad \textbf{(0.21)}$$

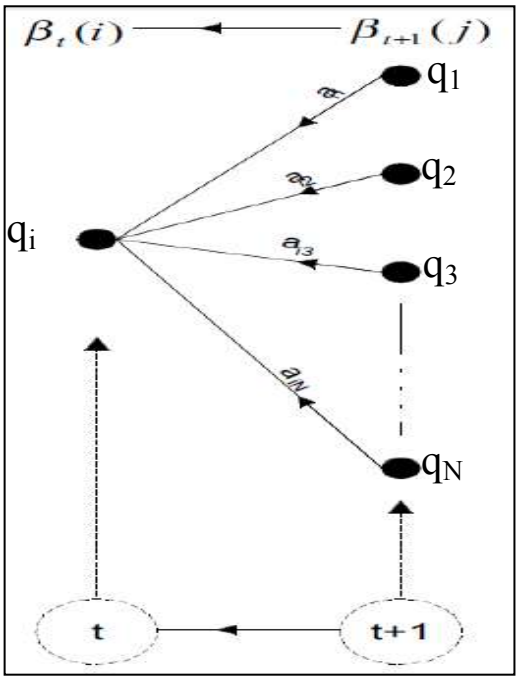


Figure II.33. *: Le principe de fonctionnement de l'algorithme backword.*

Ainsi, la somme de ce produit donne une autre façon pour calculer la probabilité p (O|λ), tout en utilisant les probabilités forward et backward : [13]

$$\boldsymbol{p\ (O|\lambda) = \sum_{i=1}^{N} p(O, o_t = i|\lambda) = \sum_{j=1}^{N} \alpha_t(i)\beta_t(i)\quad pour\quad 1 \le t \le T} \qquad \textbf{(0.22)}$$

Nous avons utilisé la fonction hmm_fb de la bibliothèque h2m pour calculer les variable α et β.

```
%phase 2:
% Evaluation des probabilités
%appelle de la fonction qui calcule les probabilités
[alpha, beta, logl, dens] = hmm_fb (X, A, pi0, mu, sigma);
```

Figure 0.34. *: Le code Matlab pour l'évaluation des probabilités.*

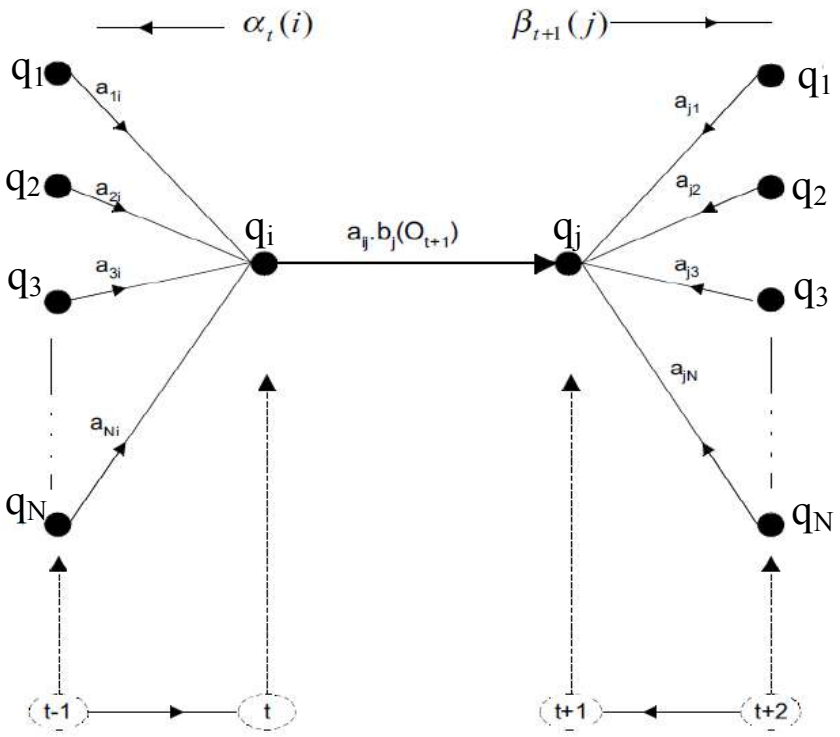


Figure II.35. ***:*** *Le principe de fonctionnement de l'algorithme Forward-Backward.*

3.2.1.3. La ré-estimation

L'algorithme utilisé pour la ré-estimation du modèle, λ = (A, pi0, mu, sigma), est l'algorithme itératif de ***Baum-Welch***. Cet algorithme maximise la fonction de vraisemblance pour le modèle λ = (A, pi0, mu, sigma). Ici, pour chaque itération de l'algorithme de Baum-Welch la ré-estimation des paramètres du HMM est faite à une valeur maximale de ***p (o/λ)***. L'algorithme de Baum-Welch est basé sur une combinaison de l'algorithme forward et bachword, qui ont été expliqués avant. [15]

Cette phase a pour but de re_estimer la matrice de transition A_, la matrice de moyenne mu_ et la matrice de covariance sigma_. Avec un nombre fini d'itération n_it on re_estime ces valeurs en utilisant ces deux fonctions de la bibliothèque « ***h2m*** » *hmm_mest* pour ré_estimer la matrice de transition et *mix_par* pour re_estimer les matrices mu et sigma.

```
%phase 3:
%ré-estimation
for i = 1:n_it
   [A_, logl(i), gamma] = hmm_mest(X, st, A_, mu_, sigma_, QUIET);
   [mu_, sigma_] = mix_par(X, gamma, DIAG_COV, QUIET);
%   fprintf(1, 'Iteration %d:\t%.3f\n', i, logl(i));
end
```

Figure II.36. ***:*** *Le code Matlab pour la ré-estimation des matrices mu et sigma.*

3.2.1.4. Le modèle HMM

Le model Hmm se constitue de trois variables ***λ= (A_, mu_, sigma_)*** sera enregistré dans une base de donnée sous format de fichier ***.MAT*** par la fonction ***save***.

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%                                                                                                     %
%                                         La Base de Donnée                                           %
%                                                                                                     %
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
save('reciteur1.mat','A_','mu_','sigma_')     % "yassine al djazairi"
save('reciteur2.mat','A_','mu_','sigma_')     % "khalil el hossari"
```

Figure II.37. : *Le code Matlab pour générer les modèles λ = (A_, mu_, sigma_).*

3.2.2. *L'identification*

Le problème d'identification se pose quand, étant donné une série d'observations, nous avons à trouver la séquence la plus probable des états cachés d'un modèle HMM. Ce problème est d'autant plus intéressant que dans plusieurs cas, les états cachés du HMM représentent quelque chose de non observable directement. Pour déterminer la séquence des états cachés la plus probable, étant donné une séquence d'observations, O = {o_1, o_2, o_3, …o_4}, et un HMM λ = (A_, mu, sigma) l'algorithme de ***Viterbi*** est le plus utilisé. Dans cette méthode, la séquence complète des états avec le maximum de vraisemblance est trouvée. [15]

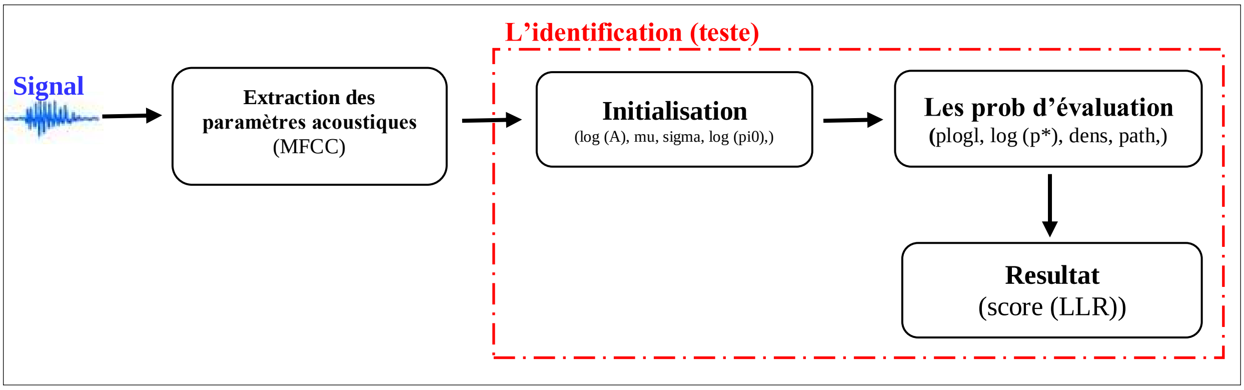


Figure II.38. : *La chaine de prétraitement du processus de l'identification.*

3.2.2.1. L'initialisation

Cette étape consiste à initialiser les variables log (A), log (pi0), mu et sigma pour la phase d'identification.

- **Log (A) et log (pi0):**

L'initialisation se fait en chargeant la matrice A_ et le vecteur pi0 à partir de fichier « .MAT » de modèle d'apprentissage λ, Après calculé leurs logarithme. Et comme la matrice

de transition est de type gauche droit, il est impossible de calculer le logarithme de zéro. Pour résoudre ce problème on additionne à chaque élément de la matrice A_ avec un petit nombre comme par exemple « ***realmin*** » en Matab. La figure suivante montre le code matlab pour initialiser ses variables.

```
% HMM probabilities in log
A = log(A + realmin);
pi0 = log(pi0 + realmin);
```

Figure II.39. *: Le code Matlab pour initialiser la matrice A et le vecteur pi0.*

- **mu (μ) :** Charger la variable **mu_** du modèle d'apprentissage **λ** à partir du fichier .MAT.
- **Sigma (sigma) :** Charger la variable **sigma_** du modèle d'apprentissage **λ** à partir du fichier .MAT.

3.2.2.2. Les probabilités d'évaluations

- **Log (p*) :** le calcul de la probabilité de la séquence d'état le plus probable.
- **Plogl :** L'état qui donne le plus grand log (P*) au temps T.
- **Path :** la séquence optimale d'états.
- **Dens :** la densité de probabilité.

La figure dessous montre la fonction pour calculer ces probabilités :

```
function [logl, path, plogl, logdens] = hmm_vit (X, A, pi0, mu, Sigma, QUIET)
```

Figure II.40. *: la fonction de l'algorithme de Viterbi.*

Le problème de l'identification est de trouver le chemin qui correspond à la séquence d'état la plus probable au sens de la probabilité de vraisemblance de la séquence d'observations X. l'algorithme Viterbi résoudre ce problème par excellence. [13]

Algorithme de Viterbi : L'algorithme peut se résumer formellement de la façon suivante : Pour chacun des états, calcul par récurrence de la variable intermédiaire :

$$\delta t\,(i) = \max p\,(q_1, q_2, \ldots, q_{t-1}, q_t = i, o_1, o_2, \ldots, o_{t-1}|\lambda) \qquad (0.23)$$

Le maximum étant calculé sur toutes les séquences d'états possibles, q_1, q_2, ..., q_{t-1}. Ce calcul se fait de manière récursive en deux étapes :

a. Initialisation :

$$\delta_1(j) = pi0(j).bj(1) \text{ , pour } 1 \leq j \leq N \qquad (0.24)$$

b. Relation récursive :

$$\delta t+1(j) = bj\,(t+1)\,\{max1_i_N_t\,(i).aij\}, \text{ pour } 1 \leq j \leq N,\ 1 \leq t \leq T-1 \qquad (0.25)$$

Calcul de δ_T (j), $1 \leq j \leq N$, en utilisant cette dernière récursion et en retenant toujours un pointeur sur l'état « élu » dans une opération de maximisation.

c. Détermination de l'état final du système (t = T) le plus probable :

$$i_t = arg\ max_{1 \leq j \leq N}\ .\ (\delta_T\,(i)) \qquad (0.26)$$

d. Suivi du chemin le plus probable en revenant en arrière, soit : Si on note :

$$\Phi(i) = arg\ max_{1 \leq j \leq N}\ .\ (\delta_{t-1}\,(i)) \qquad (0.27)$$

La séquence d'état la plus probable peut être trouvée par :

$$i_t = \Phi_{t+1}\,(i_{t+1}) \qquad (0.28)$$

Et en fin, la séquence « i_1, i_2, ..., i_T » est la séquence la plus probable des états cachés pour la séquence d'observation considérée.

3.2.2.3. Résultat (score)

Le résultat de l'algorithme de Viterbi est interprété dans la variable score qui est égale à plogl dans le programme, cette dernière représente l'état qui donne la plus grande probabilité de la séquence d'état la plus probable au temps T.

```
score =

   1.0e+04 *

   -0.5683   -1.1609    0.0029   -0.7713   -1.0062
```

Figure II.41: *le résultat de l'algorithme de Viterbi pour 5 réciteurs.*

La figure ci-dessous montre la comparaison du réciteur numéro 3 par tous les 5 réciteurs de la base de données, où score (3) est le log (p (O/λ)) du réciteur numéro 3 qui a la vraisemblance maximale car on a un son qui existe dans la base de donnée.

3.2.3. *La vérification*

Comme le décrit la partie précédente, l'algorithme de Viterbi 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. [9]

Un module de décision doit, à partir de ce score, fournir une décision qui constituera la réponse finale du système de reconnaissance. [9]

Le module de décision reçoit, en entrée et pour chaque test, un score. Celui-ci résulte de la comparaison entre les caractéristiques de l'utilisateur testé et la référence apprise lors de la phase d'apprentissage. Un score élevé signifiera que la probabilité pour que l'utilisateur testé corresponde à l'identité qu'il annonce est élevée (figure II.43) et un score faible signifiera que cette probabilité est faible. La décision binaire qui constitue la sortie du module résulte de la comparaison de ce score avec un « ***seuil*** **»** défini à l'avance. Si le score est supérieur au seuil, l'utilisateur est accepté et s'il est inférieur au seuil, l'utilisateur est rejeté. [9]

En pratique, le seuil de décision doit être déterminé avant utilisation du système. Il est choisi de manière « ***empirique*** **»**. [9]

4. Conclusion

Dans ce chapitre, nous avons commencé tout d'abord par la représentation acoustique du signal de parole, Nous avons étudié particulièrement les coefficients MFCC (Mel Frequency Cepstrum Coefficients) utilisés dans la majorité des systèmes de reconnaissance actuelle.

Cette technique consiste à calculer les coefficients cepstraux sur une échelle Mel qui se rapproche de la perception fréquentielle de l'oreille humaine. Après l'application d'une transformée de Fourier à court terme, l'énergie est calculée dans des bandes critiques modélisées par des filtres triangulaires. L'échelle des fréquences est exprimée en Mel. Le cepstre est ensuite calculé, Après nous avons décrit l'approche statistique pour la reconnaissance fondée sur l'utilisation des modèles de Markov cachés (HMM).

Une fois qu'un système est décrit comme un HMM, trois problèmes doivent être résolus ; le problème d'apprentissage du modèle peut être résolu par l'algorithme Baum-Welch (ou Forward-Backward). Le problème de l'évaluation des probabilités est résolu par l'algorithme Forward. Enfin, Le problème de l'identification peut être résolu en utilisant l'algorithme de Viterbi.

Chapitre 3:

La Conception

1. Introduction

Ce chapitre met l'accent sur la conception d'une application ayant pour objectif de réaliser un système de reconnaissance automatique de la parole pour la récitation des mots du noble Coran, avec vérification des règles de Tadjwîd. Il couvrira tous les aspects de la partie pratique de notre travail de recherche en englobant les divers schémas. Ces derniers, à travers des algorithmes et des méthodes appliqués, illustrent dans la mesure du possible tout ce qui a été présenté dans la partie théorique.

2. Description du système

Notre système est sous la forme d'une application de reconnaissance automatique de la parole, il est appliqué dans la science de Tadjwîd. Ce système a la capacité de juger si la récitation de l'apprenant est juste ou fausse. C'est la raison pour laquelle il était indispensable de préparer une collection des sons pour construire une base de données.

Comme nous avons dit dans le chapitre I, nous nous intéressons qu'à la règle de l'assimilation, en particulier, le Noun Essakina. Le mot choisi pour la construction de la base de données est : مِن وَلِي. Le même principe pourrait être appliqué pour les autres mots.

2.1. *Construction de la base de données*

Le mot مِن وَلِي a été enregistré par 5 réciteuses. Chacun a répété le mot 10 fois. Voici ci-dessous un schéma illustratif :

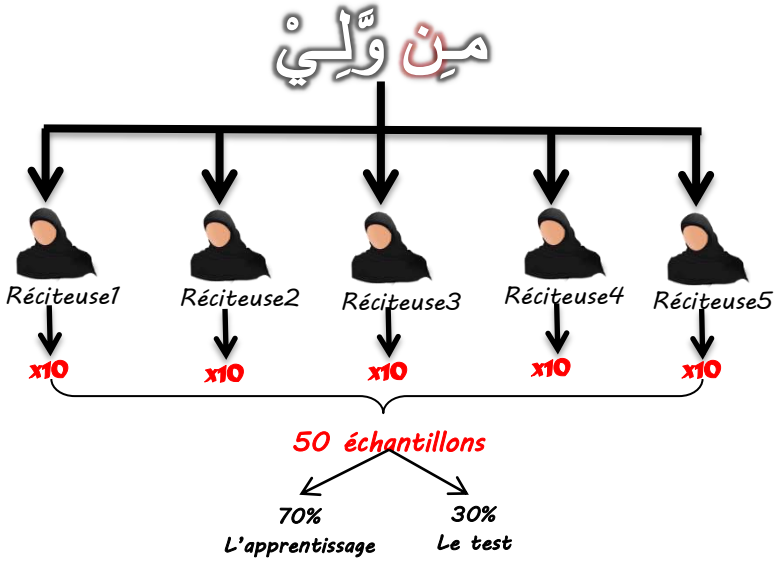


Figure 0.1. *Schéma synoptique pour la construction de la base de données.*

Notre base de données est, donc, constituée de 40 échantillons d'enregistrements. Nous consacrons 70% d'entre eux pour faire l'apprentissage des chaines de Markov afin de construire les modèles λ (A, mu, sigma). Le reste qui représente 30% est consacré au test afin de calculer le seuil. Ce dernier va juger si la récitation de l'apprenant est vraie ou fausse.

3. Organigramme

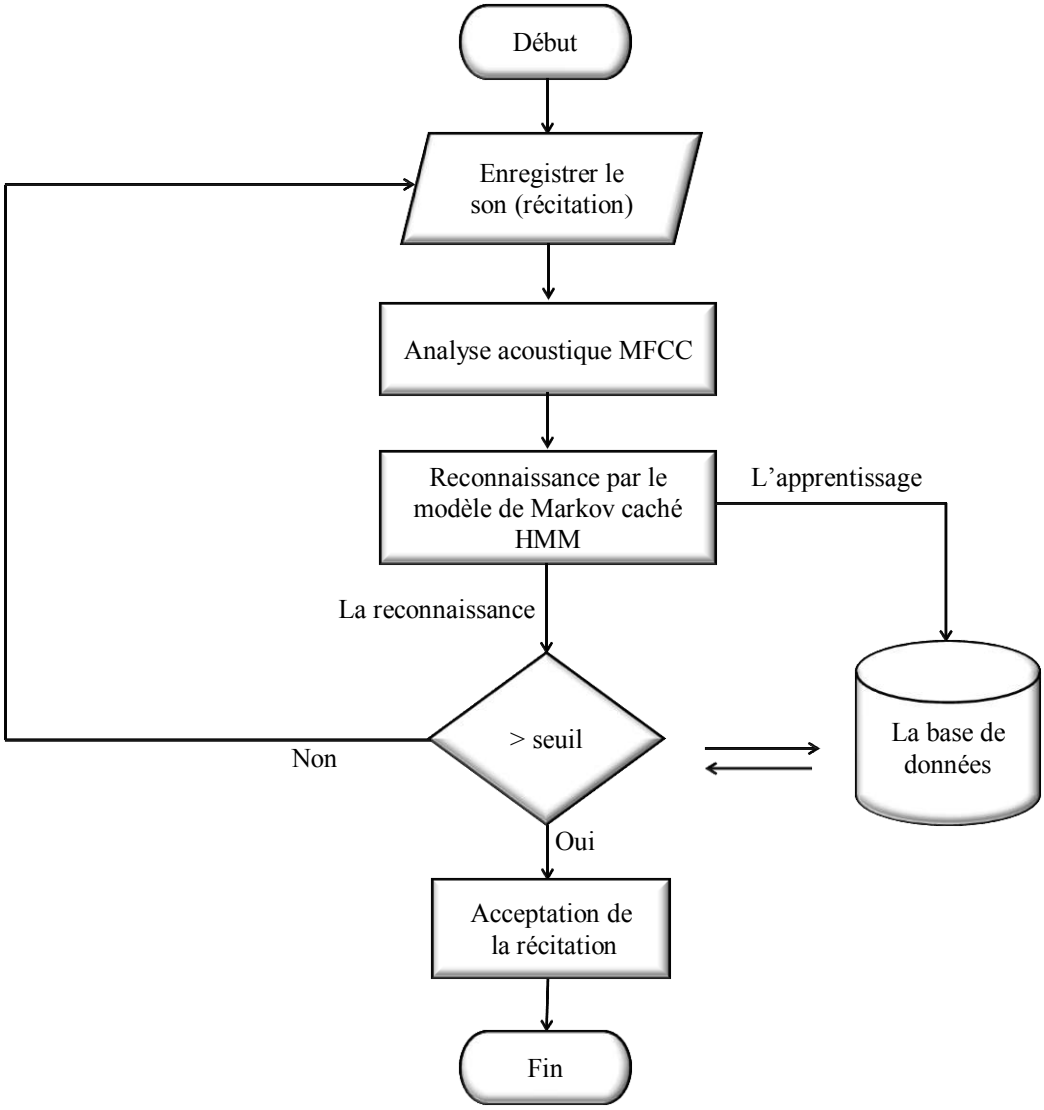


Figure III.2. *Organigramme.*

3.1. Collection des échantillons (enregistrement du son)

Dans cette partie nous proposons le code MATLAB pour enregistrer le son à partir de microphone et aussi pour lire un son. La figure ci-dessous montre le code associé pour l'enregistrement.

```
%------------------Enregister le son-----------------------------
%--le son est automatiquement enregistré sous fomat.wav

fre_ech=16000; %la fréquence d'echantillonnage
temps=4;  %le temps pour enregistrer le son en secondes

%-----fonction d'enregistrement
son_enregitre= wavrecord(temps*fre_ech,fre_ech,'double')

wavplay(son_enregitre,fre_ech);%----fonction pour écouter le son
```

Figure III.3. *Le code MATLAB pour enregistrer le son.*

D'abord, la fonction d'enregistrement exige au programmeur la définition de certains paramètres avant d'enregistrer le son à savoir : le nombre d'échantillons, le nombre de bit par échantillon et la fréquence d'échantillonnage. Le format de la commande est comme suit:

```
« son_enregitré = wavrecord (le nombre d'échantillon, fréquence
d'échantillonnage, nombre de bit par échantillon) »
```

Ensuite, nous choisissons aussi bien le temps d'enregistrement « temps = 4 secondes», que la fréquence d'échantillonnage « fre_ech = 16000 HZ ». Nous savons que le nombre d'échantillon est égal à : la durée d'enregistrement multipliée par la fréquence d'échantillonnage. Nous obtenons à la fin la taille de vecteur « son_enregitré = 64000 x1 ».

Pour lire un son à partir d'un disque dur, nous utilisons la commande «waveread» qui associe un vecteur de valeurs au son lu. La figure qui suit montre le code Matlab associé pour la lecture.

```
%-----fonction de lecture d'un son
[son_enregistre,FS,NBITS]=wavread('Souret El Zomar-al ghamidi.wav');
wavplay(son_enregistre,FS);%----fonction pour écouter le son
```

Figure III.4. *Le code MATLAB pour la lecture du son.*

Où FS est la fréquence d'échantillonnage et NBITS est le nombre de bit par échantillon.

3.2. Extraction des paramètres acoustique par la méthode (MFCC)

Le programme choisi dans notre application est le MFCC. Ce dernier est le plus utilisé pour l'extraction des paramètres acoustiques. Nous citons, en premier lieu, la référence [1] dans laquelle ils ont fait une étude comparative entre les principaux processus d'analyse acoustiques et ils ont arrivés aux résultats suivants :

> « *Une étude comparative classique a été effectuée entre plusieurs représentations du signal vocal: cepstre en sortie d'un banc de filtres en échelle Mel (MFCC) ou en échelle linéaire (LFCC), coefficients de prédiction linéaire (LPC) ou de réflexion (RC), cepstre calculé à partir des coefficients auto-régressifs (LPCC). Ces représentations ont été appliquées dans un système de reconnaissance fondé sur un alignement DTW entre un mot de test et des mots de référence. Dans cette étude, les MFCC donnent les meilleurs résultats, ce qui montre plus généralement l'intérêt d'un pré-traitement par banc de filtres, d'une échelle fréquentielle non linéaire, et de la représentation cepstrale [6]. »*

En second lieu, la référence [2] affirme :

> **"***MFCC is perhaps the best popular features extraction method used recently (Bateman, D. et al ; 1992; Ehab, M. et al., 2007), and MFCC also one of the most popular feature extraction techniques used in speech recognition, whereby it is based on the frequency domain of Mel scale for human ear scale (Chetouani, M. et al., 2002)***"**

C'est la raison pour laquelle, nous avons choisi cette analyse acoustique qui est basé sur le calcul des cepstrume dans l'échelle MEL linéaire et logarithmique. Cette analyse consiste à calculer les coefficients spectraux du signal d'apprentissage et du signal du test. La figure suivante résume son fonctionnement.

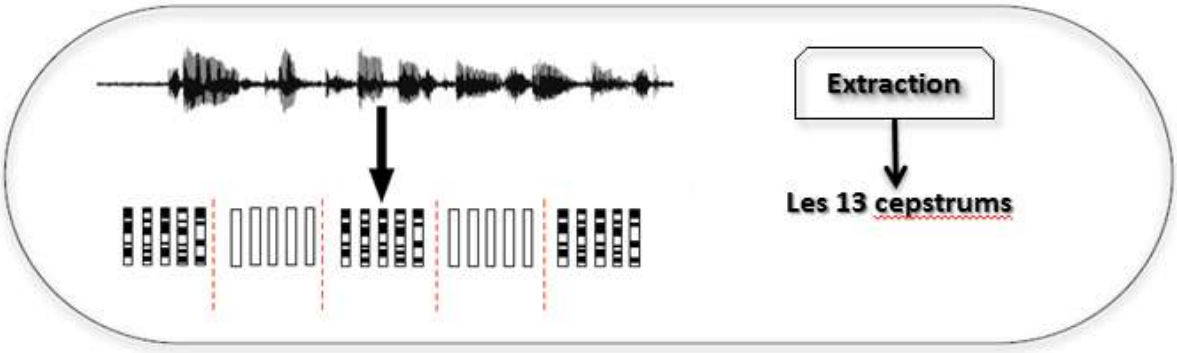


Figure III.5. *L'analyse acoustique.*

La **(figure III.6)** montre le code Matlab pour l'appel de la fonction MFCC qui a été expliquée dans le chapitre précèdent. Nous avons choisi le temps de trame égale à 20ms, le taux de chevauchement égale à 25% et le nombre de coefficients spectraux égale à 13 coefficients.

```
temps_trame=20;
taux_chevauchement=0.25;
nbr_D_coeff_cepstraux=13;
%appel de la fonction MFCC
Cepstrum=MFCC(son_enregistre,FS,temps_trame,taux_chevauchement,nbr_D_coeff_cepstraux);
```

Figure III.6. *Le code Matlab pour l'appelle de la fonction MFCC.*

3.3. Reconnaissance automatique de la parole par la méthode (HMM)

Les modèles de Markov cachés travaillent sur une information extraite du signal qui est faite par une analyse acoustique MFCC. Le tableau suivant montre la performance des différents systèmes RAP avec différents processus d'analyse acoustique.

Tableau 0.1. *la performance des différents systèmes RAP avec différents processus d'analyse acoustique***.**

References	Pre-Processing	Feature extraction method	Classification / recognition techniques	Performance
[Tabbal, H. et al.' 06]	Pre-Emphasis Filter	MFCC	Hidden Markov Model (HMM)	85%-92%
[Youcef, A. &, Emam , O.'04]	-	MFCC	Hidden Markov Model (HMM)	90.2%
[Ahmad, A.M. et al.'04]	Digital Filtering	MFCC LPCC	Recurrent Neural Network (RNN)	MFCC 95.9%-98.6% LPCC 94.5%-99.3%
[Bashir, M.S. et al.'03]	Preemphasis Filtering (bandpass filter)	Spectrographic Analylis	Spectrographic Analylis based on different frequency band of intensity	93.33%

[kirchoff, K.et al.'04]	Kneser –Ney Smoothing	Not stated	Hidden Markov Model (HMM)	Not stated
[Hasan, M.R.et al.'04]	-	MFCC	Vector Quantization (VQ)	57%-100%
[Podder. S.K.'97]	-	LPCC	VQ and HMM	62%-96%
[Bhotto, M.Z.A & amine, M.R.'04]	-	MFCC	Vector Quantization (VQ)	70%-85%

D'après le tableau [2] les systèmes basés sur la combinaison de processus MFCC avec les modèles de Markov cachés ont atteint une performance supérieure à 90%. C'est pour cette raison que nous avons opté pour ces deux processus pour notre conception.

3.3.1. Apprentissage

Cette étape a pour but de maximiser la fonction de vraisemblance pour les modèles ***p (séquence observation/ un modèle λ)*** qui seront par la suite stockés dans la base de données. Cette étape se déroule en 3 phases : l'initialisation des variables {A, pi0, mu, sigma}, ensuite le calcul des probabilités d'évaluations et à la fin la re-éstimation de ses variables.

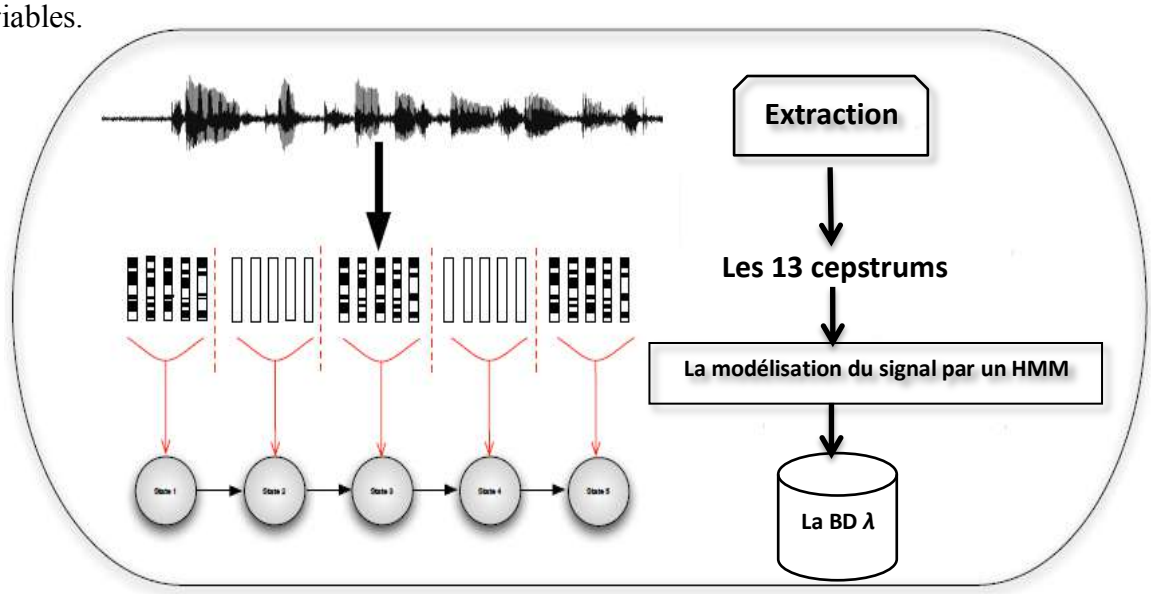


Figure III.7. *La phase d'apprentissage en HMM.*

3.3.1.1. Le nombre d'états (N)

Nous avons choisi le nombre d'état égal à 14 états en se basant sur la référence [17].

Ils ont fait la reconnaissance automatique de la parole pour la récitation des mots du noble Coran avec vérification des règles de Tadjwîd, plus particulièrement la règle de la substitution (الإقلاب). Ils ont choisi 15 états pour modéliser leurs signaux.

Tableau 0.2. *La variation de nombre d'états pour une meilleure performance.*

(1,18,1)	(1,15,1)	(2,6,1)	(1,6,1)	
%100	%94	%66	%83	أنبأك
%100	%72	%0	%39	إذ إنبعث
%97	%97	%0	%94	بذنبهم
%97	%100	%89	%100	من بعد
%94	%97	%0	%91	أن بورك
%69	%100	%100	%55	يؤمن بالله
%94	%100	%0	%86	زوج بهيج
%100	%100	%11	%100	مشاءٍ بنميم
%100	%100	%22	%100	سميع بصير
%94.5	%95	%32	%83	**دقة النموذج**

3.3.2. La reconnaissance

La reconnaissance se devise en 2 parties : l'identification et la vérification

- l'identification

Cette étape consiste à comparer la matrice d'observation avec tous les modèles HMM de la base de données afin de trouver le meilleur score ***(le maximum de vraisemblance),*** pour identifier le signal à reconnaitre qui correspond à la séquence d'état la plus probable.

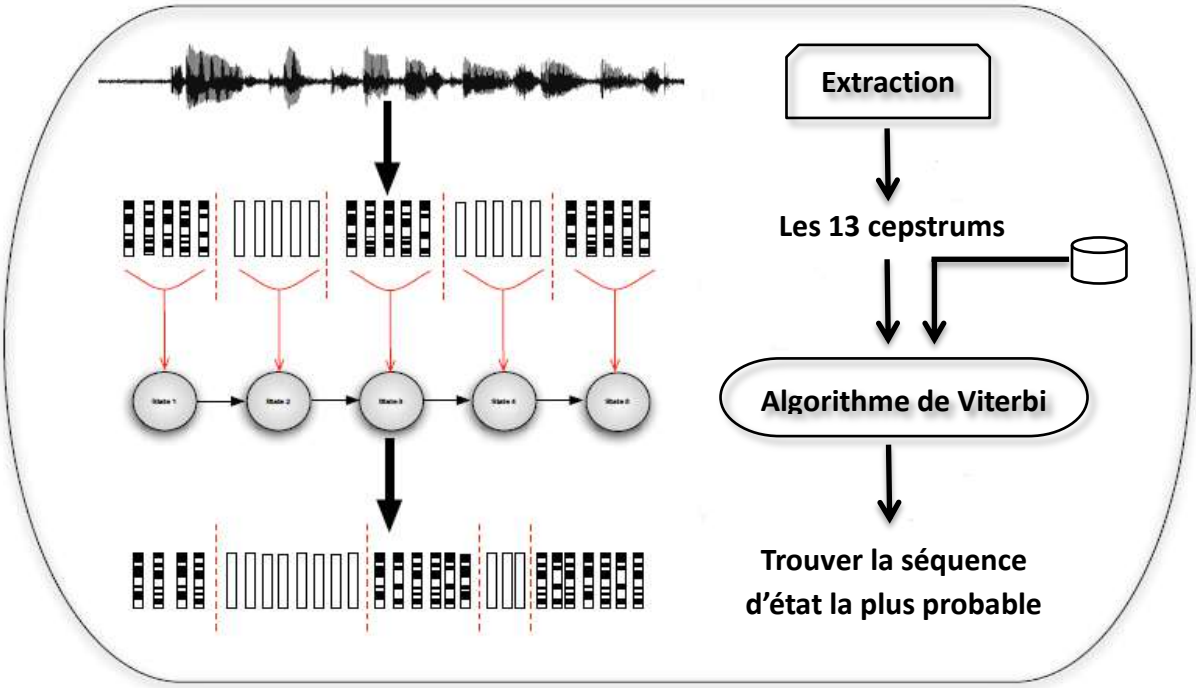


Figure III.8. *L'identification.*

L'étape de l'identification permet d'identifier chaque signal de test (30%), la figure ci-dessous identifié le son de test de réciteur1 avec les sept modèles de la base de données. Et le meilleur score est obtenu avec l'enregistrement numéro 7.

```
score =

   1.0e+04 *

  Columns 1 through 7          Réciteur1

   -0.2288   -0.2233   -0.2280   -0.1640   -0.1776   -0.2037   -0.1517

  Columns 8 through 14         Réciteur2

   -0.5633   -0.5629   -0.5601   -0.5519   -0.5038   -0.5403   -0.4850

  Columns 15 through 21        Réciteur3

   -0.6157   -0.6093   -0.5915   -0.6128   -0.5859   -0.6068   -0.7172

  Columns 22 through 28        Réciteur4

   -1.2976   -0.8113   -1.2500   -1.0609   -0.9379   -1.0843   -1.2461

  Columns 29 through 35        Réciteur5

   -0.5107   -0.5754   -0.5541   -0.7100   -0.4409   -0.6205   -0.6548
```

Figure III.9. *L'identification du réciteur numéro 1.*

Le tableau suivant exprime le taux de l'identification pour tous les autres tests :

Tableau III.3. *Le taux d'identification avec différentes réciteurs***.**

Réciteurs	**Taux d'identification**
Réciteur 1	**100%**
Réciteur 2	**100%**
Réciteur 3	**100%**
Réciteur 4	**100%**
Réciteur 5	**100%**

- La vérification :

Cette étape consiste a jugé si le mot a été bien prononcé ou non, ce jugement est basé sur ***un seuil,*** si le score est supérieur à cette variable la récitation sera accepter si non elle sera rejeter. Nous avons utilisé deux méthodes pour calculer le seuil.

En vérification, deux types de tests, deux types d'erreurs [9]

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

Les tests de bonne récitation lors desquels l'échantillon présenté au système correspond à une bonne récitation. [9]

Test de mauvaise récitation lors desquels l'échantillon présenté au système provient d'une mauvaise récitation. [9]

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 : [9]

Faux rejet (FR) erreur commise lorsque le système rejette, à tort, une bonne récitation (i.e. erreur commise lors d'un test de bonne récitation). [9]

Fausse acceptation (FA) erreur commise lorsqu'une mauvaise récitation est malencontreusement acceptée en tant qu'utilisateur légitime (i.e. erreur commise lors d'un test mauvaise récitation). [9]

Le module de décision reçoit, en entrée et pour chaque test, un score. Celui-ci résulte de la comparaison entre les caractéristiques de son de l'utilisateur testé et la référence apprise lors de la phase d'enrôlement. Un score élevé signifiera que la probabilité pour que l'utilisateur testé corresponde à l'identité qu'il annonce est élevée et un score faible signifiera que cette probabilité est faible. La décision binaire qui constitue la sortie du module résulte de la

comparaison de ce score avec un seuil défini à l'avance. Si le score est supérieur au seuil, l'utilisateur est accepté et s'il est inférieur au seuil, l'utilisateur est rejeté. [9]

Le choix d'un seuil a une incidence directe sur les performances du système. Pour un système idéal, les scores obtenus par ***une bonne récitation*** seront tous plus élevés que les scores obtenus par ***une mauvaise récitation***. Dans ce cas, le seuil à fixer se situe entre le score de mauvaise récitation le plus élevé et le score d'une bonne récitation le plus faible, assurant ainsi une authentification parfaite. [9]

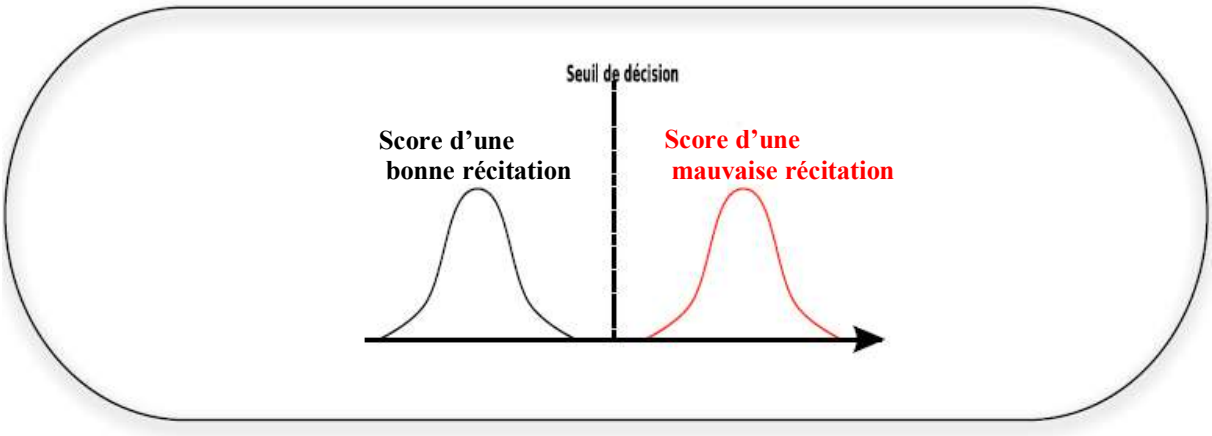


Figure III.10. *Répartition des scores de bonne récitation et mauvaise récitation d'un système parfait*

À chaque valeur de seuil est associé un couple (p(FA), p(FR)) et l'ensemble des couples obtenus peut être représenté sous la forme d'une courbe courbe DET (Detection Error Tradeoff). La courbe DET est basée sur une distribution normale qui se substitue à l'échelle linéaire. [9]

Selon l'application visée, le coût d'un faux rejet ou d'une fausse acceptation ne sont pas les mêmes. L'estimation des performances d'un système de vérification pour une application spécifique se doit de tenir compte des coûts différents de ces deux types. [9]

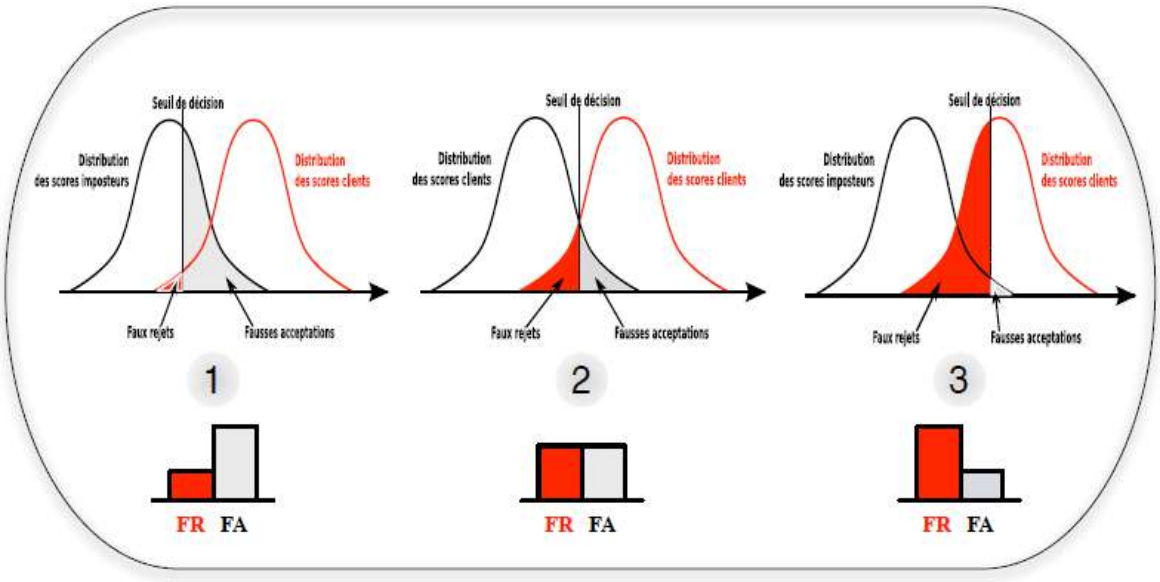


Figure III.11. *Influence du seuil de décision sur les erreurs d'un système de reconnaissance.*

1 - Seuil de décision choisi dans le but de réduire le nombre de faux rejets
2 - Seuil de décision choisi pour obtenir autant de faux rejets que de fausses acceptations (EER)
3 - Seuil de décision choisi dans le but de réduire le nombre de fausses acceptations

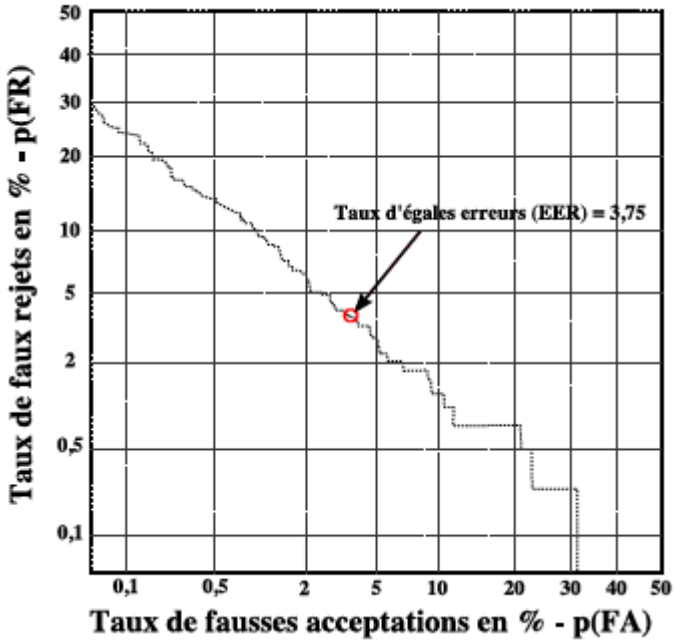


Figure III.12. *la courbe DET.*

A travers le graphe représenté dans la figure III.12 le seuil optimale correspond à la valeur (EER) le taux d'égales erreurs qui est représenté par un cercle rouge. Dans ce point la probabilité pour avoir une fausse acceptation est égale à la probabilité de faux rejet.

Méthode 1 :

A partir d'un ensemble de test nous avons opté a un seuil égale à : « ***-3.7754e+03*** » et l'algorithme DET affirme cette valeur. Ou le principe de DET est de trouver le seuil optimal à travers deux fichiers. Le premier fichier « true_score » qui contient que les scores de la bonne prononciation, et le deuxième fichier « false_score » qui contient que les score de la fausse prononciation. A partir de ces fichiers l'algorithme va varier la valeur de scores jusqu'à arriver à la valeur optimale.

Nous avons consacré 20 enregistrements pour la bonne prononciation et 30 enregistrements pour la fausse prononciation et la performance de notre système est montrée dans le **(tableau III.8)** où nous sommes arrivés à une performance de 85% pour la reconnaissance des mots avec une bonne prononciation. Et 40% pour la reconnaissance des mots avec une fausse prononciation.

Tableau III.4. *La performance du système***.**

	Nombre totale	**Nbr d'acceptation**	**Nombre de rejet**	**Performance**
Bonne prononciation	20	17	3	**85%**
Fausse prononciation	30	12	18	**40%**

Méthode 2 :

L'idée est de fixer un seuil pour chaque réciteurs de la base de données, donc on obtient 5 seuils et à chaque fois si on veut tester un nouveau son le programme va faire une identification par la maximisation de la vraisemblance (**max LLR**) pour trouver le réciteur le plus proche du son de test, après on va comparer le avec le seuil adéquat.

- **Résultats de test**

Les résultats de test avec des ***nouveaux Sons juste*** mais pour les mêmes réciteurs de la base de données sont montrés dans le tableau suivant:

Tableau III.5. *Le taux de reconnaissance avec différentes sons (juste) des réciteurs de la base de données.*

Réciteurs	Nbr répétition	Accepter (oui)	Rejeter (non)	pourcentage
1	6	5	1	**83,33%**
2	3	2	1	**66.66%**
3	2	2	0	**100%**
4	4	4	0	**100%**
5	1	0	1	**0%**
Somme	16	13	3	**81.25%**

Les résultats de test avec des ***nouveaux Sons faux*** mais pour les mêmes réciteurs de la base de donnes sont montrés dans le tableau suivant:

Tableau III.6. *Le taux de reconnaissance avec différentes sons (faux) des réciteurs de la base de données.*

Réciteurs	Nbr repitition	Accepter (oui)	Rejeter (non)	pourcentage
1	3	2	1	**33,33%**
2	3	0	3	**100%**
3	4	3	1	**25%**
4	4	3	1	**25%**
Somme	14	8	6	**42.85%**

Nous remarquons que les résultats de test avec les même réciteurs de la base de donnes est 81% pour les sons juste, et de 42% pour les sons faux.

Les résultats de test avec des ***nouveaux Sons juste*** des apprenants sont montrés dans le tableau suivant:

Tableau III.7. *Le taux de reconnaissance avec différentes sons (juste) des apprenants*

Apprenants	Nbr répétition	Accepter (oui)	Rejeter (non)	pourcentage
1	4	2	2	**50%**
2	4	3	1	**75%**
3	8	2	6	**25%**
4	3	1	2	**33.33%**
5	1	0	1	**0%**
Somme	20	8	12	**40%**

Les résultats de test avec des ***nouveaux Sons faux*** des apprenants sont montés dans le tableau suivant:

Tableau III.8. *Le taux de reconnaissance avec différentes sons (faux) des apprenants*

	Nbr repitition	**Accepter (oui)**	**Rejeter (non)**	**pourcentage**
Faux_minwaliy	33	13	20	**60.60%**

Nous remarquons que les résultats de test est 40% pour les sons juste, et de 60% pour les sons faux.

4. Proposition d'une interface graphique

L'interface graphique a été réalisée par la fonction « guide » définie dans Matlab. Les figures qui suivent illustrent étape par étape le fonctionnement de notre interface.



Figure III.13. *Correcteur automatique des règles de Tadjwîd.*



Figure III.14. *Choisissez une bonne récitation pour l'écouter.*



Figure III.15. L'enregistrement du son.

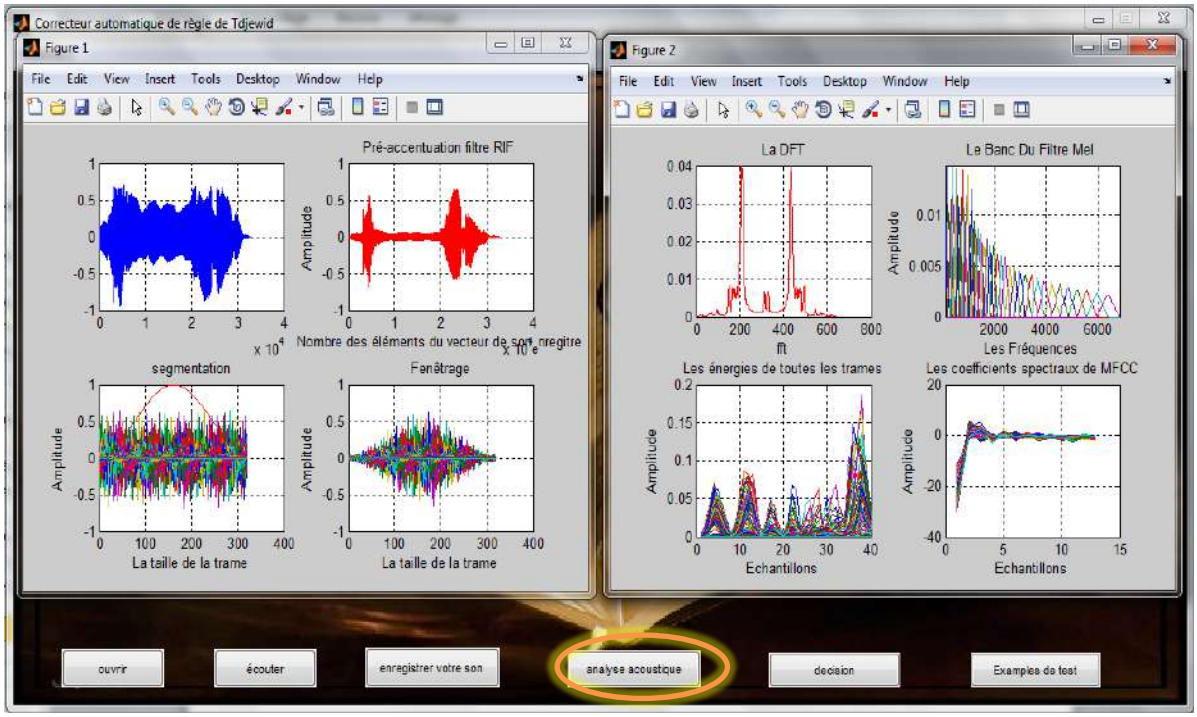


Figure III.16. *L'analyse acoustique.*



Figure III.17. *La décision (bonne prononciation).*



Figure III.18. *La décision (mauvaise prononciation).*

5. Conclusion

Dans ce chapitre nous avons décrit notre système en expliquant comment on a créé notre base donnée, ensuite nous avons cité l'ensemble des processus appliqués pour la conception de notre application. Nous avons commencé par un organigramme qui décrit notre système puis, nous avons expliqué les étapes en commençant par l'enregistrement du son jusqu'à l'arrivé à la décision.

La décision a été calculée par deux méthodes, où nous avons arrivé à un taux de reconnaissance égale à 40% pour reconnaitre une bonne récitation et 60% pour la mauvaise récitation dans la première méthode. Et un taux de reconnaissance égale à 85% pour reconnaitre une bonne récitation et 40% pour la mauvaise récitation dans la deuxième méthode.

À la fin de notre chapitre, nous avons présenté notre interface graphique et son fonctionnement. Cette dernière a été réalisée par la fonction « ***Guide*** » du Matlab

Conclusion générale

Conclusion générale

Le travail que nous avons effectué consiste à concevoir un système RAP qui permet la correction automatique des règles de Tdjwid. Ce système peut être considéré comme une autre alternative d'apprendre la récitation du noble coran, son importance est de servir la religion islamique et encourager les musulmans à faire progresser leurs récitations. Notre proposition aide les apprenants à apprendre les règles de Tdjwid d'une façon plus rapide. Elle peut être, aussi, un outil d'auto-apprentissage pour les adultes ayant des contraintes de temps leur empêchant d'apprendre le noble coran.

Pour réaliser notre système RAP, notre travail s'étale sur 3 chapitres : le premier est une sorte d'initialisation de notre travail en définissant la science de Tdjwid et les différents modes de récitation du noble coran ainsi que les règles du Tdjwid. Puis nous citons les différentes méthodes utilisées dans le domaine de reconnaissance automatique de la parole.

Dans le deuxième chapitre, nous avons commencé tout d'abord par la représentation acoustique du signal de parole. Nous avons étudié particulièrement les coefficients MFCC (Mel Frequency Cepstrum Coefficients) utilisés dans la majorité des systèmes de reconnaissance actuelle. De plus, nous avons décrit l'approche statistique pour la reconnaissance fondée sur l'utilisation des modèles de Markov cachés (HMM).

Nous avons consacré le dernier chapitre à la conception et la mise en œuvre de notre projet. Nous avons vus les différents mécanisme et outils utilisés pour la réalisation de notre système RAP.

Quant aux questions posées dans l'introduction nous somme arrivées aux réponses suivantes :

- le processus utilisé pour extraire les caractéristiques de signale de la parole est la technique MFCC (Mel Frequency Cepstrum Coefficients) ;
- le signal de parole est modélisé par des modèles de Markov cachés (HMM) ;
- Le HMM fait l'apprentissage de la matrice de cepstrom qui est générer par l'analyse acoustique MFCC et la combinaison entre ces deux processus (MFCC_HMM) a prouvé leur efficacité pour réaliser un système fiable ;

Nous espérons que ce modeste travail pourrait être utile pour d'autres étudiants et qu'il mérite d'être enrichi et développé. Pour cela nous proposons les perspectives suivantes :

- Poursuivre le développement du correcteur automatique des règles de Tdjwid pour que le système RAP soit capable de connaitre toutes les règles de Tdjwid, et non seulement la règle de al_idgham.
- Améliorer la performance du système. Malgré le fait que nos résultats soient encourageants, il y a beaucoup de choses qui devraient être mis au point jusqu'à ce que nous atteignons un niveau de performance adapté au caractère sacré du Saint Coran.
- Agrandir la base de données des mots pour avoir une bonne performance.
- Développer l'application pour reconnaitre un verset coranique et non seulement un mot.

Références

[I.1] سورة الحجر (مكية) ، سورة رقم 15، عدد آياتها 99، آية رقم 15

[I.2] سورة القمر (مكية) ، سورة رقم 54، عدد آياتها 55، آية رقم17

[I.3] سورة المزّمّل (مكية) ، سورة رقم 73، عدد آياتها 20، آية رقم 04

[I.4] سورة الفرقان (مكية) ، سورة رقم 25، عدد آياتها 77، آية رقم 32

[I.5] سورة البقرة (مدنية) ، سورة رقم 02 عدد آياتها 286 آية رقم 121

[1] Abdenour Hacine-Gharbi, " Sélection de paramètres acoustiques pertinents pour la reconnaissance de la parole," PDF, faculté de technologie, Sétif, Algérie, 09/12/2012. [En ligne]. Disponible: http://hal.archives-ouvertes.fr/docs/00/84/36/52/PDF/abdenour.hacinegharbi_2151_vm.pdf . [Consulté le 22 Décembre 2013].

[2] Noor Jamaliah Binti Ibrahim, "automated tajweed checking rules engine for quranic verse recitation", university of malaya, kuala lumpur, Malaisie, 2010.

[3] Inconnu. inconnu, *résumer des règle de tadjwîd*, 1, inconnue: inconnu, inconnue. [Livre électronique]. Disponible: http://www.sahab.fr/livre/regles_tajwid.pdf. 3

[4] Inconnu. inconnu, *Le mérite de la lecture du Coran*, 1, inconnue: inconnu, inconnue. [Livre électronique]. Disponible: http://www.action-islamique.com/science/merite_coran.pdf.

[5] طلاب مسجد النور, "مذكرة أحكام التجويد", مسجد النور, المدية, الجزائر

[6] Salma Jamoussi, "Méthodes statistiques pour la compréhension automatique de la parole", Doctorat de l'université Henri Poincaré Nancy 1, Nancy, France ,2004 [En ligne]. Disponible: http://www.freewebs.com/salmajamoussi/These.pdf [consulter le 23 mais 2014].

[7] Asmaa Amehraye,"Débruitage perceptuel de la parole", l'Ecole Nationale Supérieure des Télécommunications de Bretagne, France, mai 2009. [En ligne]. Disponible:http://perso.telecombretagne.eu/pastor/data/Theses/Memoire_These_Amehraye.pdf [consulter le 23 mais 2014].

[8] Teddy DIDÉ, " Réalisation d'un framework pour la reconnaissance de la parole ," ECOLE POLYTECHNIQUE DE L'UNIVERSITÉ DE TOURS, Tours, France, 2006-2007. [En ligne]. Disponible: https://stagemaster.googlecode.com/files/Rapport_stage.doc[Consulté le 23 mais 2014].

[9] Anthony LARCHER, "Modèles acoustiques à structure temporelle renforcée pour la vérification du locuteur embarquée", université d'avignon et des pays de vaucluse, 24 septembre 2009. [En ligne]. Disponible: http://biobimo.eurecom.fr/publications/these_antho.pdf , [Consulté le 24 Mai 2014].

[10] MCLOUGHLIN Ian, *Applied speech and audio processing with Matlab examples*, édition, Singapore: Éditeur, Année. [Livre électronique]. Disponible: http://www.sahab.fr/livre/regles_tajwid.pdf 9

[11] Vincent Luba, AinaneYounes, " Modèles de Markov cachés Reconnaissance de la parole," PDF, faculte polytechnique de mons, Ville, Province, Pays, 2004-2005. [En ligne]. Disponible: http://tcts.fpms.ac.be/cours/1005-07-08/speech/projects/2004/ainane_luba.pdf . [Consulté le 20 Mars 2014].

[12] Noureddine Aboutabit, "Reconnaissance de la Langue Française Parlée Complétée (LPC) : Décodage phonétique des gestes main-lèvres," PDF, institut national polytechnique de Grenoble, le 11 décembre 2007. [En ligne]. Disponible: http://tel.archives-ouvertes.fr/docs/00/27/01/62/PDF/these_aboutabit.pdf . [Consulté le 06 Avril 2014].

[13] MAAZOUZ Mouhamed, " Implémentation Temps Réel d'un Vocodeur à Faible Débit," PDF, université Dr. yahia fares, Médéa, Algérie.

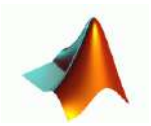
[14] Maurice Charbit, *Reconnaissance de mots isolés (utilisation des modèles HMM)*, édition, Singapore: Éditeur, 25 Octobre 2002. [Livre électronique]. Disponible: http://www.sahab.fr/livre/regles_tajwid.pdf

[15] Waleed Abdulla et Nikola Kasabov, *The Concepts of Hidden Markov Model in Speech Recognition*, 99/09, DUNEDIN NEW ZEALAND: ISSN 1177-455X, May 1999. [Livre électronique].Disponible : http://otago.ourarchive.ac.nz/bitstream/handle/10523/997/dp1999-09-online.pdf?sequence=3 .

[16] Initiale de l'auteur. Nom de l'auteur (ou Nom de la collectivité auteur), "Hidden Markov Model - Description du modèle," *Hidden Markov Model*, [En ligne]. Disponible: URL. [http://www-igm.univ-mlv.fr/~dr/XPOSE2012/HiddenMarkovModel/description.html].

[17] رماح معتصم ابراهيم, رواء عوض محمد , سوسن عبد اللطيف أحمد, هالة فتح الرحمن عباس, *استخدام نماذج ماركوف الخفية في التعرف على حكم الإقلاب* " université sciences et technologies soudan, Khartoum, Soudan, 2008 .[En ligne]. Disponible: http://www.arabcomputersociety.org/archives/indexACS&page=article&op=view&path%5B%5D=124.htm

Logiciels



Matlab, *version 2012*, [Logiciel], Montréal, Qc.: la Cie. 2012.



Gold Wave, *version 5,7*, [Logiciel].

Annexe A

La boite à outils H2M

H2M est un ensemble de fonctions MATLAB / OCTAVE qui mettent en œuvre l'algorithme EM dans le cas des modèles de mélange ou modèles de Markov cachés avec distribution conditionnelle de l'État gaussienne multivariée. Plus précisément, trois cas particuliers ont été considérés

1. Mélanges gaussiens.
2. Modèles gaussiens ergodique (ou totalement) connectés de Markov cachées.
3. Des modèles de Markov cachés gaussiennes gauche -droite.

On a utilisé ce tool box pour modéliser notre signal par des modèles de Markov cachés. La bibliothèque H2M contient des fonctions MATLAB parmi elles :

- **hmm_mint :** Initialise les paramètres de distribution à l'aide de multiples observations (modèle gauche-droite)
- **hmm_fb :** calcul les probabilités forward et backward
- **hmm_mest :** ré-estimations les paramètres de transition pour plusieurs séquences d'observation.
- **Mix_par :** ré-estimations des Paramètres mélange.
- **Hmm_vit :** Calcul la séquence la plus probable d'états (algorithme de Viterbi).

Cette bibliothèque est publiée sur ce site :
http://perso.telecom-paristech.fr/~cappe/Code/H2m/

Annexe B

La boite à outils DETware

NIST (*National Institute of Standards and Technology)* fournit DETware, c'est un ensemble de fonctions MATLAB qui permet de tracer DET-courbes (Detection Error Trade-off curves), nous avons utilisé ce toolbox pour déterminer notre seuil de décision.

La bibliothèque ***DETware*** contient les fonctions suivant :

- **Compute_DET :** calcul des probabilité fausse acceptation/faux rejet
- **Plot_DET**: tracer la courbe DET
- **Min_DCF**: Trouver fonction de coût minimal de détection
- **Set_DET_limits**: initialise le min / max et tracer des limites

Cette bibliothèque est disponible dans ce site :http://www.itl.nist.gov/iad/mig/tools/

Annexe C

La boite à outils Auditory

Ce toolbox contient beaucoup de fonctions utilisé dans le domaine de reconnaissance du parole parmi elles on trouve la fonction **MFCC**, cette dernière a été utilisé dans notre application avec des modifications afin de calculer les coefficients cepstraux MFCC (Mel Frequency Cepstrum Coefficients).

Cette bibliothèque est disponible dans ce site :
https://engineering.purdue.edu/~malcolm/interval/1998-010/