

République Algérienne Démocratique et Populaire
Ministère de L'Enseignement Supérieur et de la Recherche Scientifique

UNIVERSITE MOULOUD MAMMERI DE TIZI-OUZOU



FACULTE DE GENIE ELECTRIQUE ET D' INFORMATIQUE
DEPARTEMENT D'ELECTRONIQUE

**Mémoire de Fin d'Etude
de MASTER ACADEMIQUE/PROFESSIONNEL
Spécialité : RESEAUX ET TELECOMMUNICATION**

Présenté par

MEZIOUD Djamila

Thème

**Evaluation de la qualité de la voix
téléphonique en utilisant les tests MOS
et model E**

Mémoire soutenu publiquement le 11 JUIN 2015 devant le jury composé de :

M F. OUALOUCH

M A A, UMMTO, Président

M M. MAZRI

M C A, UMMTO, Rapporteur

M D.ALLOUACHE

M A B, UMMTO, Examineur

M M.SEHAD

M A A, UMMTO, Examineur



Remerciements

Au terme de ce travail, je tiens à exprimer ma profonde gratitude et mes sincères remerciements à mon encadrant MR LAZERI pour sa patience, sa disponibilité et surtout ses judicieux conseils.

Mes profonds remerciements vont à Melle TALEM pour toute la documentation qu'elle m'a fournie.

Mes plus sincères remerciements vont à tous les membres du jury qui m'ont fait l'honneur de juger mon travail.

Mes remerciements vont enfin à toute personne qui a contribué de près ou de loin à l'élaboration de ce travail.

Dédicace

Je dédie ce modeste travail à :

Mes très chers parents qui m'ont toujours soutenu.

Ma chère Grand-Mère.

*Mes chère sœurs **NADIA, ZAHRA** et **FAZIA**.*

*Mes chers frères **IBRAHIM** et **NOUR-EDDINE**.*

*Mon adorable nièce **HIND**.*

*Mon adorable neveu **YANI**.*

Tous mes ami(e)s ainsi qu'à tous ce qui me sont chers.

Et à toute personne m'ayant fait part de son savoir.

MEZIOUD Djamila

Glossaire

TIC : Technologies de l'Information et de la Communication

UIT : Union International des Télécommunications

MOS : Mean Opinion Score

DMOS : Degradation Mean Opinion Score

CMOS : Comparaison Mean Opinion Score

ACR: Absolute Category Rating

DCR: Degradation Category Rating

CCR: Comparison Category Rating

RTC : Réseau Téléphonique Commuté

RNIS : Réseau Numérique à Intégration de Service

PCM : Pulse Code Modulation

DCME : Digital Circuit Multiplication Equipement

ADPCM :Adaptative Differntial Pulse Code Modulation

VOIP : Voice Over Internet Protocol

ZAA : Zone à Autonomie d'Acheminement

ZTP : Zone de Transit Secondaire

ZTS : Zone de Transit Secondaire

DTMF : Dual Time Multi Frequency

SNR: Signal to Noise Ratio

NRZ: No Return to Zero

ETTD : Equipement Terminal de Traitement de Données

ETCD : Equipement Terminal de Circuit de Données

DTE : Data Terminal Equipment

DCE : Data Communication Equipment

TPS: Television Par Satellite

UTP: Unshielded Twisted Pair

STP: Shielded Twisted Pair

DRT: Diagnostic Rhyme Test

STI: Speech Transmission Index

SII : Speech Itelligibility Index

PESQ: Perceptual Speech Quality Measure

PAQM : Perceptual Audio Quality Measure

BSD: Bark Spectral Distorsion

MBSD: Modified Bark Spectral Distorsion

QOS: Quality Of Service

ANOVA: Analysis Of Variance

Ie: Impairment equipment

Liste des figures

Figure I.1: principe du RTC

Figure I.2: architecture du RTC

Figure I.3. Phase d'établissement d'un appel

Figure I.4: la tension sinusoïdale de la tonalité

Figure I.5: chronogramme de numérotation décimale

Figure I.6: numérotation par DTMF

Figure I.7: le signal sinusoïdal de la sonnerie

Figure II.1. schématisation d'une chaîne de transmission

Figure II.2 : modulation d'amplitude

Figure II.3 : modulation de fréquence

Figure III.4: modulation de phase

Figure IV.5: signal analogique

Figure II.6 : signal échantillonné

Figure II.7 : signal analogique

Figure II.8 : signal quantifié

Figure II.9 : Pas de quantification

Figure II.10 : L'allure du bruit de quantification

Figure II.11 : Quantification défaut

Figure II.12 : erreur de quantification

Figure II.13 : Quantification centrée

Figure II.14 : Bruit de quantification

Figure II.15 : quantification logarithmique

Figure II.16 : Evolution du SNR

Figure II.17 : Signal original

Figure II.18 : Codage NRZ

Figure II.19 : Signal original

Figure II.20 : Codage bipolaire simple

Figure II.21. Transmission en mode simplex

Figure II.22. Transmission en mode half duplex

Figure II.23. Transmission en mode full duplex

Figure III.1. Conversion MOS/R

Figure IV. 5. Interface graphique du test utilisant l'échelle MOS

Figure IV.6. Interface graphique du test utilisant l'échelle R du modèle E

Figure IV.7. Interface graphique du test utilisant l'échelle de comparaison par paires

Figure IV.8. Courbe d'interpolation des I_e obtenus pour test MOS1

Figure IV.6. Moyenne pour chaque condition avec écart type

Figure IV.6. Résultats MOS2 pour chaque condition

Liste des tableaux

Tab II.1: *Caractéristiques physiques des différents câbles.*

Tab III.1 : Échelle MOS

Tab III.2 : Echelle CMOS

Tab III.3. Echelle DMOS

Tab III.4 : Les mesures objectives de qualité des signaux les plus utilisées

Tab III.5. Echelle MOS

Tab IV.1. Caractéristiques des conditions

Tab. IV.2. Répartition des conditions dans chaque test

Tab IV.3. Notes attribuées aux locuteurs du test MOS

Tab IV.4. Moyenne pour chaque condition avec écart type

Tab IV.5. Calcul des valeurs normalisées sur l'échelle R des conditions du test MOS1

Tab IV.6. Notes attribuées aux locuteurs du test MOS2

Tab IV.7. Résultats du test MOS2

Tab IV.8. Comparaison des résultats entre MOS1 et MOS2

Tab IV.9. Résultats du test avec l'échelle R pour les locuteurs allemands

Tab IV.10. Résultats du test avec l'échelle R pour les locuteurs français

Tab IV.11. Résultats de la comparaison entre conditions

Sommaire

Introduction

Chapitre I : généralités sur la téléphonie

Préambule	01
I.1 La téléphonie.....	01
I.1.1 Téléphonie analogique	01
I.1.2 Téléphonie digital	02
I.2 Quelques généralités.	02
I.2.1 CoDec.	02
I.2.2 Bande passante	02
I.2.3 Wideband	03
I.3 Evolution de la téléphonie par rapport à la bande passante	03
I.4 La parole et importance de la bande passante	03
I.4.1 Production de la parole	03
I.4.2 Perception de la parole	04
I.5 Le réseau téléphonique commuté RTC	04
I.5.1 Principe du réseau RTC	04
I.5.2 Gestion du réseau téléphonique commuté.....	05
I.5.3 Hiérarchie du réseau téléphonique commuté	06
I.5.4 Phase d'établissement d'une communication téléphonique.....	08
I.5.5 Signalisation.....	14
I.5.6 L'acheminement	14
Discussion.	15

Chapitre II : Généralités sur la transmission de données

Préambule	16
II.1 Chaîne de transmission	16
II.2 Description des fonctions utilisées	17

II.2.1 Traduction.....	17
II.2.2 Traitement.....	17
II.2.3 La modulation.....	17
II.2.4 Amplification.....	18
II.2.5 Canal	18
II.2.6 Démodulation	18
II. 2.7 Filtre	18
II.3 Les techniques de transmission de la voix.....	18
II.3.1 Transmission analogique	18
II.3.1.1 Transmission par modulation d’amplitude de la porteuse	19
II.3.1.2 Transmission par modulation de fréquence	19
II.3.1.3 Transmission par modulation de phase.....	21
II.3.1.4 Avantage et inconvénients de la transmission par modulation	22
II.3.2 Transmission numérique.....	22
II.3.2.1 L’échantillonnage	23
II.3.2.2 La quantification.....	24
II.3.2.3 Le codage.....	28
II.3.3 Transmission en bande de base.....	30
II.3.3.1 Principe	30
II.3.3.2 Les inconvénients de la transmission en bande de base.....	30
II.4 Modes de transmission	30
II.4.1 Transmission série et transmission parallèle.....	30
II.4.1.1 Liaison parallèle.....	30
II.4.1.2 Liaison en série.	31
II.4.2 Transmission synchrone et asynchrone	31
II.4.2.1 Liaison synchrone.....	31
II.4.2.2 Liaison asynchrone	31
II.5 Sens de transmission.....	32
II.5.1 Mode simplex	32

II.5.2 Mode half-duplex.....	32
II.5.3 Mode full duplex.....	32
II.6 Supports physique de transmission.....	33
II.6.1 Canal de transmission.	33
II.6.2 Supports de transmission	33
II.6.2.1 Les lignes spécialisées	33
II.6.2.2 Les ondes hertziennes	33
II.6.2.3 Le satellite.....	33
II.6.3 Ligne de transmission	34
II.6.3.1 Ligne bifilaire	34
II.6.3.2 La paire torsadée.....	34
II.6.3.3 Le câble coaxial	34
II.6.3.4 La fibre optique.....	34
Discussion	37

Chapitre III : Méthodes d'évaluation de la parole

Préambule.....	38
III.1 Qualité et intelligibilité de la parole	38
III.2 Les aspects déterminant la qualité de la voix	39
III.3 Les différents facteurs intervenant dans l'évaluation de la voix	40
III.3.1 Le délai du transit.....	40
III.3.2 Gigue de phase	41
III.3.3 La perte des données	41
III.4 Méthode d'évaluation de la qualité de la voix	41
III.4.1.1 MOS (Mean Opinion Score)... ..	42
III.4.1.2 CMOS (comparison).	42
III.4.1.3 DMOS (Degradation).....	43
III.4.2. Analyse et présentation des résultats.....	44
III.4.2.1 Analyse statique par t-test et ANOVA.	46

III.4.2.2.1 T-test ou test de student.....	47
III.4.2.2.2 ANOVA (analys of variance).	48
III.5 Critère objectifs.	49
III.5.1 SNR segmental (segSNR).	49
III.5.2 BSD et MBSD	50
III 5.3 PSQM	51
III 5.4 PESQ	51
III.5.5 Le modèle E.....	52
III.5.5.1 Equipement Impairement Factor.	52
III.5.5.2 L'échelle R	53
III.6 Conversion R/MOS.	54
Discussion	55

Chapitre IV :

Préambule	56
IV.1 Echelles de quantification de la qualité de la parole	56
IV.1.1 ACR type MOS	56
IV.1.2 Echelle R du modèle E	57
IV.1.3 Echelle nominale	58
IV.2 Conditions choisies.	59
IV.3 Application des tests.	61
IV.3.1 test MOS.....	61
IV.3.2 Test MOS 2.	65
IV.3.3 Comparaison entre MOS1 et MOS2.	69
IV.3.4 Test avec l'échelle R..	69
Discussion	74
Conclusion	
Bibliographie	
Annexe	

INTRODUCTION

Introduction

Les Technologies de l'information et de la communication (TIC) représentent aujourd'hui un facteur non négligeable du développement, parmi elles la téléphonie qui a été initialement prévue pour transmettre la voix humaine entre deux lieux distants l'un de l'autre, est devenu avec l'évolution des moyens de transports qui a induit l'augmentation des besoins en moyens de communication et de besoins de mobilité, un moyen essentiel de développement économique et de confort civile.

Vu son importance et son intérêt, on s'intéresse beaucoup à la qualité de service fournie, plus précisément à la qualité de la parole perçue. Durant la transmission, plusieurs facteurs interviennent et affecte la qualité de transmission du signal de parole, parmi ces facteurs, la bande passante. Dans le cas de la bande étroite, les fréquences présentes dans un signal de parole ne sont pas intégralement transmises. En effet, les fréquences utilisées par la parole humaine peuvent être comprise entre 110 et 7000 Hz, ce qui à première vu semble loin de la bande-étroite de la téléphonie. Comme autre facteur, les CoDecs qui réduisent le débit de transmission et causent des pertes de données lors du codage et décodage du signal. Aussi, le bruit qui est un signal nuisible qui se superpose au signal utile dégrade considérablement la qualité de la parole.

A cet effet, des traitements sont nécessaires pour améliorer la qualité de la parole. Plusieurs méthodes ont été donc utilisées pour évaluer la qualité dans le but de l'améliorer (ITU-Rec. G.107, 2005 ; Raake, 2005). Ces méthodes sont divisées en deux classes, les méthodes subjectives et les méthodes objectives. Les méthodes subjectives sont largement utilisées dans la littérature (Krebber, 1995 ; Raake, 2005) et donnent en général des résultats concluants. Des méthodes objectives sont aussi utilisée pour évaluer la qualité du signal de parole (Yang, 1999 ; Wang, 1997). Toutefois, ces méthodes nécessitent un ensemble de paramètres issus d'un matériel plus sophistiqué. Pour cette raison, nous avons opté pour une méthode subjective dans notre technique d'évaluation. Cette méthode consiste à appliquer des tests (MOS, modèle E) sur une base de données (signal de parole) acquise sous différentes conditions afin d'évaluer sa qualité en utilisant des échelles (ACR, DCR, CCR, R).

Notre mémoire est structuré en quatre chapitres.

Dans le premier chapitre, nous abordons des généralités sur le signal de parole et sur la téléphonie et nous présentons le réseau téléphonique commuté (RTC).

Le deuxième chapitre est consacré à la présentation des différentes techniques et modes de transmission de données.

Le troisième chapitre fera l'objet des méthodes d'évaluation de la qualité du signal de parole ainsi que les différents facteurs intervenant dans l'évaluation de cette qualité.

Dans le quatrième chapitre, nous présentons les différents tests appliqués à notre base de données, et nous donnons les résultats et leurs interprétations.

Nous terminons notre mémoire par une conclusion, deux annexes et un ensemble de références bibliographiques

CHAPITRE I

Généralités sur la téléphonie

Préambule

Le but du chapitre est de présenter des généralités sur la téléphonie, le signal de parole ainsi que le réseau téléphonique commuté (RTC) qui permet la mise en relation de deux abonnés afin de transmettre le signal de parole.

I.1. La téléphonie :

La téléphonie est considérée comme un système de communication permettant de transmettre la parole, elle peut en premier lieu être vue comme un système créant une liaison entre deux personnes et permettant la transmission d'un message. Son but comme celui de tout système comportant de la parole devrait être de maximiser la compréhension de ce message, le message étant défini comme un signal de parole transmis de la bouche d'un locuteur jusqu' à l'oreille d'un auditeur, le rôle des deux personnes changeant au cours de la conversation.

Dans le but de transmettre ces messages de paroles plusieurs systèmes de téléphonie ont été créés.

I.1.1. téléphonie analogique :

Dans un tel système, le signal vocal est transformé en signal électrique par le microphone. Les variations d'amplitudes et de fréquence qui composent le son de la voix sont transmises sous forme de variation de fréquence et d'amplitude du courant électrique.

Les transmissions analogiques sont très vulnérables aux interférences et aux divers bruits qui parasitent les communications. La qualité des communications téléphonique n'est donc pas toujours parfaite. Pour une utilisation vocale, ces perturbations ne sont pas gênantes. Par contre si l'on souhaite utiliser la ligne téléphonique pour transporter des données informatiques elle pose des problèmes, c'est pour remédier que l'on utilise des techniques de correction d'erreurs ou des moyens de transmission numériques.

I.1.2.téléphonie digitale :

La téléphonie digitale ou téléphonie numérique désigne les équipements capables de travailler avec un signal numérique. Les appareils numériques sont plus coûteux que les appareils analogiques. Ils procurent néanmoins une meilleure qualité de transmission et offrent des services annexes liés ou non au réseau téléphonique. En principe ces appareils sont utilisés sur le réseau numérique à intégration de service (RNIS).

I.2.Quelques généralités :**I.2.1.CoDec :**

les CoDecs sont des codeurs de données, qui permettent un transport plus facile sur certains canaux de transmission, en réduisant la taille des données. un décodeur de l'autre côté de cette chaîne de transmission, permet de retrouver les données envoyées avec parfois certaines dégradations. Il existe différents types de CoDec suivant le canal utilisé :

- RNIS : PCM fonctionnant à un débit fixe (G.711 à 64 Kbits/s)
- DCME : Adaptive Differential Pulse Code Modulation (ADPCM) fonctionnant à des débits variables suivant la disponibilité du réseau (G.726 à 40, 32, 24 ou 16 Kbits/s).

Le type de codage effectué sur la voix est souvent adapté au canal de transmission (en fonction du débit disponible). Mais l'utilisation de CoDec introduit une dégradation de la qualité vocale, des distorsions. Il est aujourd'hui possible de quantifier la dégradation introduite par un CoDec sur la qualité du signal de parole grâce notamment à des modèles objectifs de mesure [ITU-T Rec. G.107, 2005], ou des modèles subjectifs.

I.2.2. Bande passante :

La bande passante est la largeur mesurée en Hertz, d'une plage de fréquence $F_2 - F_1$, elle représente la capacité maximale de débit sur une liaison donnée déterminée par les technologies de transmission mises en œuvre à l'aide des équipements situés à chaque extrémité de cette liaison.

I.2.3.Wideband :

Le codage large-bande n'est pas un domaine d'étude récent, le premier CoDec utilisant une bande passante élargie, le G.722 (ITU-T Rec. G.722, 1998), fut développé dans les années 1980 pour être utilisé sur le réseau RNIS. Mais le codage de la parole dans la large-bande entraîne des techniques différentes de la bande-étroite. En effet, ya une plus forte dynamique spectrale pour la parole en bande large. Plusieurs études, efforts de développement et standardisation ont permis de créer quelques CoDecs de meilleurs qualité et moins couteux en débit.

I.3.évolution de la téléphonie par rapport à la bande passante :

La bande passante traditionnelle utilisée en téléphonie fixe, appelée « bande-étroite », est 300-3400 Hz. Elle résulte d'un filtre standardisé par l'ITU (ITU-T, Rec. G.712, 1992). Or, le réseau national de téléphonie, à l'origine analogique a subi un premier changement en passant à l'ère numérique : passant du Réseau Téléphonique Commuté (RTC) au réseau Numérique à Intégration de Service (RNIS). Pour autant ce réseau peut être défini par certaines caractéristiques comme la connexion physique entre deux interlocuteurs qui reste active durant toute la durée de la conversation. Ceci n'est plus vrai avec l'apparition de la « Voix sur IP » ou « Voice Over Internet Protocol » (VOIP) qui permet la transmission de données sous forme de paquets.

I.4.La parole et l'importance de la bande passante :

Le message à transmettre est un signal de parole. Or il est nécessaire d'adapter le système de communication au type du message transmis. Donc il est nécessaire d'augmenter la bande-étroite de la téléphonie au profit de la bande élargie. De plus, il faut prendre en compte la qualité du système dans son intégralité (facteurs de communication). la communication étant basée sur la création du message «production » et sa reception «perception ».

I.4.1.Production de la parole :

Un son à une forme physique qui se propage dans un milieu par le biais d'ondes. Ces ondes sont liées au canal de transmission (air, câble) mais surtout au producteur de ce son. Le signal de parole a des caractéristiques temporelles mais aussi fréquentielles.

La parole humaine est considérée comme un signal source (cordes vocales, glotte), filtré par des tuyaux formé par les conduits vocaux. La production vocale est composée de son ayant des composantes fréquentielles très spécifiques, dont on donne souvent comme valeurs la fréquence fondamentale F_0 correspondant au signal porteur et les premiers formants F_i correspondants aux piques dont l'amplitude spectrale dus aux résonances du conduit vocale. Certains phonèmes produisent de l'énergie dans les hautes fréquences ainsi que dans un formant très bas, autour de 150 Hz. les fréquences utilisées par la parole humaine peuvent être comprise entre 110 et 7000 Hz. La bande passante 300-3400 Hz utilisé par la téléphonie permet de faire passer les 3 formants, mais ne permet pas de transmettre l'intégralité des fréquences présentes dans un signal de parole. Par exemple il est difficile de différencier un « s » d'un « f » prononcé seul, lors d'une conversation téléphonique.

I.4.2. Perception de la parole :

Il est nécessaire de rappeler que les fréquences audible par une oreille humaine sont habituellement situées entre 20Hz et 20KHz, ce qui en première vu semble loin de la bande étroite de la téléphonie. De plus suite à de nombreuses utilisations, le cerveau humain, a créé une référence de la qualité « sonore » de la voix humaine, transmise à travers un système de téléphonie (Duncanson, 1969). L'évaluation de la qualité est alors biaisée par la référence de la téléphonie fixe, fortement liée aux fréquences de la bande-étroite.

I.5. Le réseau téléphonique commuté RTC :

Historiquement, le transport de la voie est à l'origine des premiers réseaux de transmission. Le réseau téléphonique public commuté (RTPC, ou simplement RTC) est un réseau analogique conçu pour assurer les relations de télécommunication. il permet de mettre en relation deux abonnés. Bien que destiné au transfert de la voix, le RTC offre également d'autres services comme la transmission et la réception de Fax et l'accès à internet.

I.5.1. Principe du réseau RTC :

Le réseau commuté a essentiellement pour rôle le transfert de la voix. Utilisant le principe de commutation de circuit il met en relation deux abonnés à travers une liaison dédiée pendant tout l'échange (Rogelio, 2008 ; Zhijun, 2015).

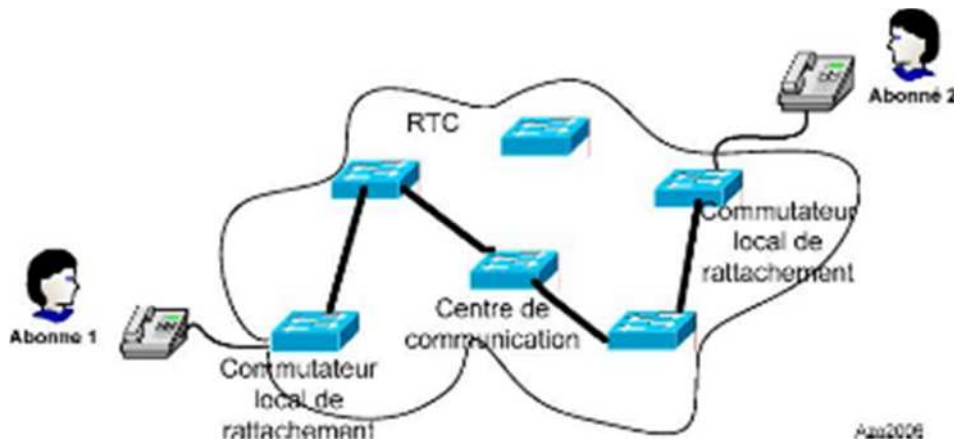


Figure I.1:principe du RTC

La commutation de circuit ou commutation spatiale consiste à juxtaposer bout à bout des voies physiques de communication (Manelaos et al, 2001), la liaison étant maintenue durant tout l'échange. A l'origine, la mise en relation était réalisée manuellement par des opérateurs. Avec la numérisation de la voix, la commutation spatiale a été remplacée par la commutation temporelle ou automatique.

On distingue deux grandes parties dans le RTC :

- ✓ Le réseau capillaire ou de distribution : c'est le raccordement de l'abonné à un point d'entrée du réseau.
- ✓ Le réseau de transit : qui effectue pour sa part le transport des communications entre les nœuds de transit concentrateurs/commutateurs.

I.5.2. Gestion du réseau téléphonique commuté (Jun-ich et Minoru, 1993):

La gestion générale du réseau décline trois fonctions : la commutation, la transmission, et la distribution.

- **Distribution** : organisation technique mise en œuvre pour relier les abonnés au commutateur le plus proche (commutateur de rattachement). l'ensemble des dispositifs permettant cette liaison est le réseau de distribution.

- **Commutation** : c'est la partie centrale du réseau, assure le raccordement des abonnés aux réseaux, gère la signalisation, assure la concentration, le multiplexage et l'aiguillage des communications dans le réseau. Il existe deux types de commutateurs :
 - ✓ les centres d'abonnés: permettent le rattachement des abonnés.
 - ✓ Les centres de transit : secondaires et principaux permettent de connecter les commutateurs qui n'ont pas de liaison entre eux. Ceci permet d'avoir un réseau étoilé plus facile à gérer et moins onéreux.
- **Transmission** : c'est l'ensemble des techniques mis en œuvre pour relier les commutateurs entre eux. L'ensemble des dispositifs et des supports de transmission entre les commutateurs sont appelés réseau de transmission.

I.5.3 : Hiérarchie du réseau téléphonique commuté :

Le réseau téléphonique est organisé en trois niveaux (zones) comme le montre la figure suivante :

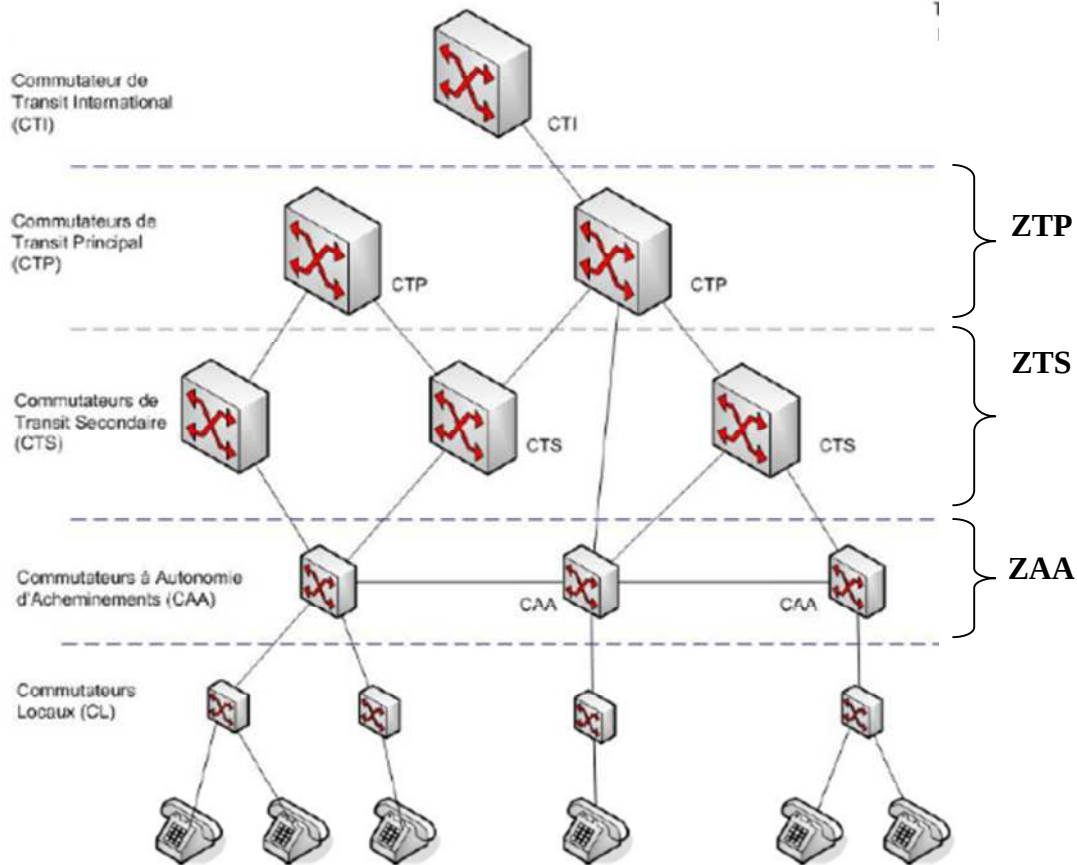


Figure I.2: architecture du RTC

• **Zone à Autonomie d'Acheminement (ZAA) :** cette zone la plus basse de la hiérarchie comporte un ou plusieurs commutateurs à Autonomie d'Acheminement(CAA) qui accueillent les abonnés et peuvent établir différents types de communication.

• **Zone de Transit Secondaire (ZTS) :** cette zone comporte des commutateurs de transit secondaire (CTS). Il n' y a pas d'abonnés reliés aux CTS. Ils assurent les brassages des circuits lorsqu'un CAA ne peut pas atteindre le CAA destinataire directement.

• **Zone de Transit principale (ZTP) :** cette zone assure la commutation des liaisons longues distance. Chaque ZTP comprend un commutateur de transit principal (CTP), l'un des CTP est relié au Commutateur International de Transit (CIT).

Dans les zones à faible densité, les abonnés sont rattachés à des commutateurs locaux (CL).

I.5.4. Phase d'établissement d'une communication téléphonique

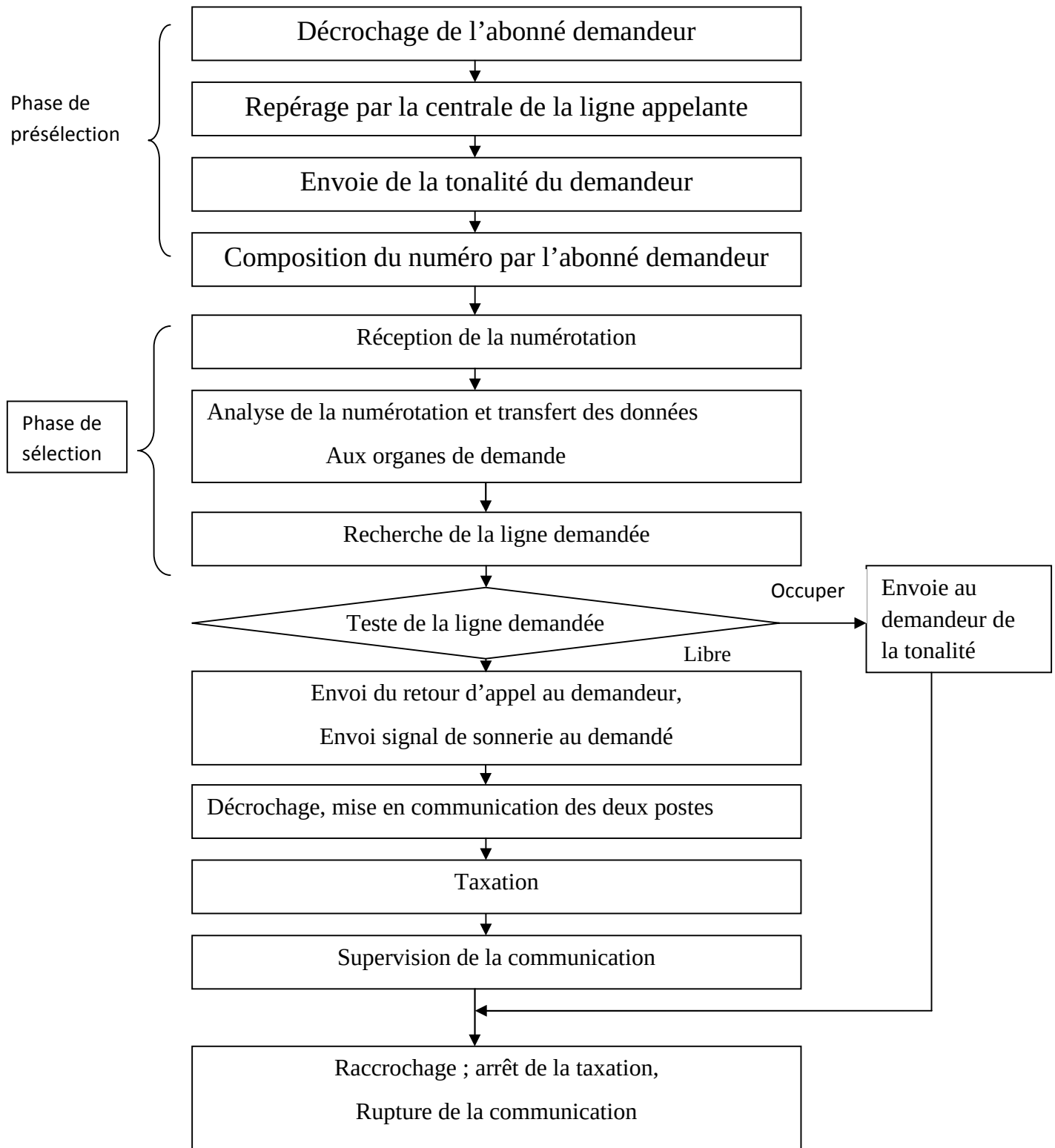


Figure I.3. Phase d'établissement d'un appel

• Ligne téléphonique au repos:

Le combiné est raccroché, Le centrale téléphonique produit en permanence une tension continue de 48V entre les deux conducteurs de la ligne. La consommation du poste doit rester inférieure à 50 micros Ampère. La consommation du poste au repos est négligée.

• Phase de présélection (prise de ligne):

Lors du décrochage, le courant appelé sur la ligne ne doit pas excéder 150mA, et se stabilise à une valeur comprise entre 20 et 60mA en moins de 2 seconde. La tension de ligne chute alors à une valeur comprise entre 10 et 22V.

Le centrale au quel est relié le demandeur détecte l'apparition du courant de ligne (il considère la ligne comme demandé dès que le courant appelé dépasse 5mA) et sait ainsi que l'abonné souhaite entrer en communication avec un autre abonné.

Une fois la prise de ligne détectée, le central envoie la tonalité d'invitation à numérote (dial tone). Selon les normes la tonalité est une tension sinusoïdale de fréquence 440 ± 35 Hz (voir figure I.4).

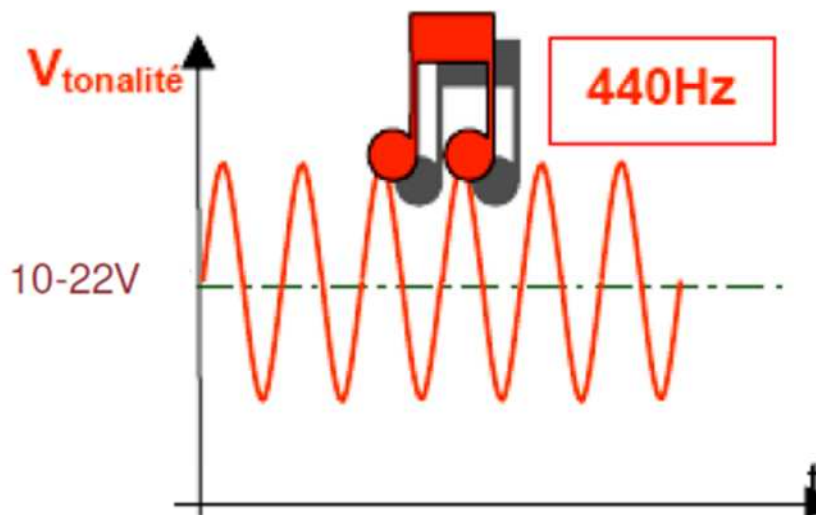
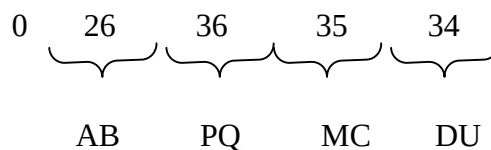


Figure I.4: la tension sinusoïdale de la tonalité

• La numérotation

▪ Le numéro de téléphone :

La numérotation en Algérie est une numérotation fermée c'est-à-dire que tous les abonnés ont un numéro qui se compose de quatre chiffres dans le centre (Millier Centaine Dizaine Unité), les centraux sont numérotés à deux chiffres appelés (indicatif ou préfix du centre PQ). La numérotation est sur 9 chiffres ce qui permet d'atteindre une capacité théorique de 640000 postes.



AB: numéro de zone.

PQ: indicatif du centre.

MC: }
 DU: } Numéro de l'abonné

Il existe deux procédés de numérotation :

▪ La numérotation décimale :

Ce procédé de numérotation est utilisé par les téléphones à cadran rotatif (poste à disque). Dans ce cas la numérotation est interrompue un nombre de fois correspondant au numéro envoyé. Une impulsion pour le chiffre 1, deux impulsions pour le chiffre 2 et ainsi de suite jusqu'à dix impulsions pour le chiffre 0.

Chaque impulsion dure 100mS, soit $33 \pm 3\text{mS}$ pour la ligne fermée (présence de courant) et $64 \pm 4\text{mS}$ pour la ligne ouverte. Pour composer le 1 il faut 100mS, pour le 2 il faut 200mS ainsi de suite jusqu'à 0 ou il faut 1S (voir figure I.4). Un intervalle de temps d'au moins 200mS doit séparer deux trains d'impulsions. Ce principe en soi est long mais il est toujours utilisé.

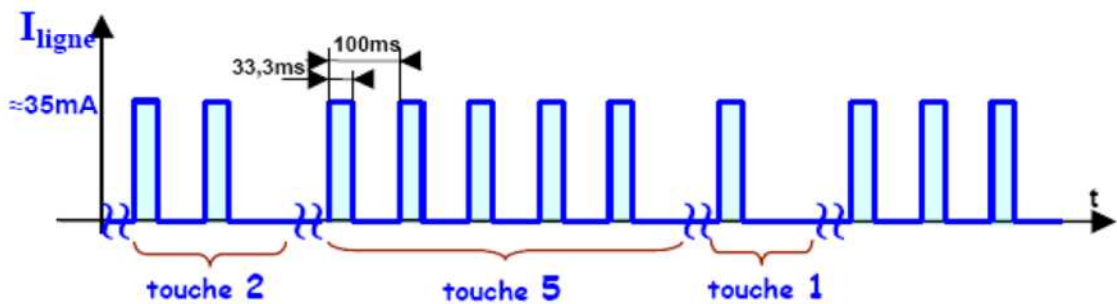


Figure I.5: chronogramme de numérotation décimale

▪ **La numérotation par fréquence vocale (DTMF : dual time multi frequency) :**

La DTMF est un procédé de numérotation qui génère des sonorités codées, il doit émettre des fréquences spécifiques dans la gamme 300Hz- 3400Hz. Mais si l'on attribue une fréquence simple pour chaque chiffre, un sifflement (ou un son propre et fort) peut provoquer une erreur de numérotation (voir figure I.6).

Ce problème d'interférence est résolu simplement par l'émission de deux fréquences simultanées par chiffre. Sur le clavier du téléphone, en appuyant sur une touche, on émet les deux tonalités correspondant à l'intersection de l'axe horizontal et de l'axe vertical. (Une fréquence « ligne » et une fréquence « colonne »)

Pour éviter les problèmes de distorsion et de génération d'harmonique, aucune fréquence n'est harmonique d'une autre fréquence. De plus il est impossible de synthétiser une fréquence à partir de la somme ou la différence de deux autres fréquences, pour éviter les erreurs de numérotations dues aux produits d'intermodulation.

Actuellement la durée d'enfoncement d'une touche et l'intervalle de temps entre deux manœuvres successives ne peuvent être inférieure à 40ms

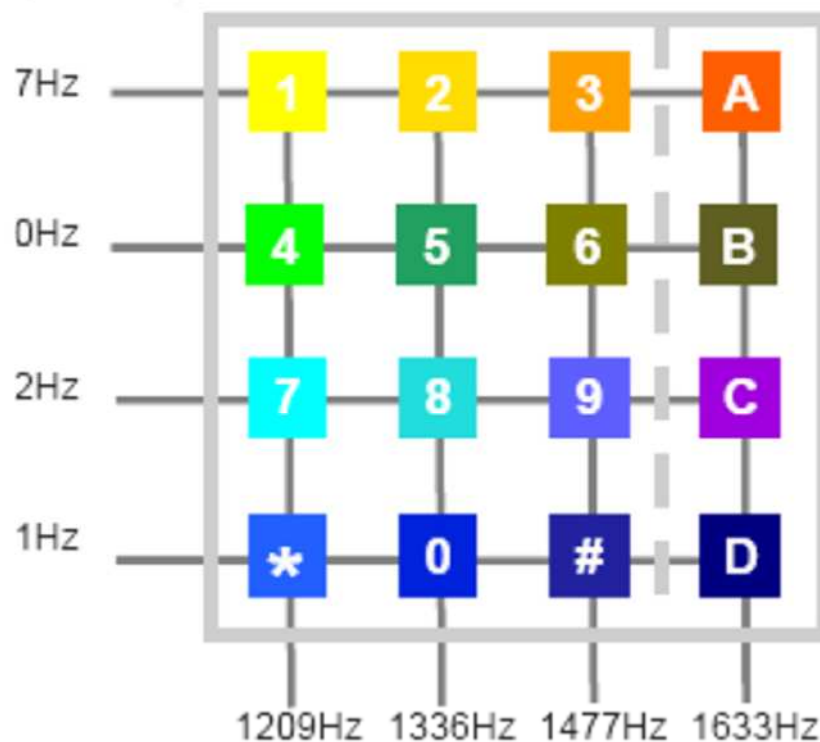


Figure I.6: numérotation par DTMF

- **Test de ligne de l'appelé :**

Après la numérotation, le central du demandeur est en relation avec le central de l'appelé. Ce second central procède tout d'abord au test de la ligne du correspondant à joindre :

- Soit l'appelé est déjà en ligne : le demandeur reçoit une tonalité d'occupation.
- Soit la ligne de l'appelé est libre : le central de l'appelé lui fait parvenir un signal de sonnerie pour lui signifier que quelqu'un cherche à le joindre. Dans le même temps l'appelant reçoit de la part de son central une tonalité de retour d'appel, simulant la sonnerie chez l'appelé.

• activation de la sonnerie et décrochage de l'appelé :

Pour activer la sonnerie, le central envoie vers le poste de l'appelé un signal sinusoïdale de fréquence environ 50Hz et de valeur efficace comprise entre 25 et 90V cette tension est envoyée par rafales, activée pendant 2 secondes et désactivée pendant environ 4 secondes. Ce signal est superposé à la tension continue de 48V (voir figure I.7).

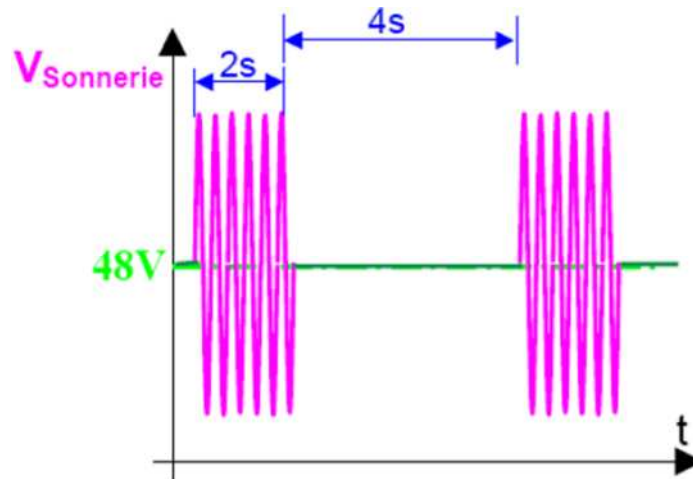


Figure I.7: le signal sinusoïdal de la sonnerie

Le décrochement du poste de l'appelé établit un courant continu d'environ 40mA dans la ligne, alors le central supprime la sonnerie et met en liaison les deux correspondants.

• phase de conversation :

La transmission de la voix s'effectue dans la bande passante 300 à 3400Hz. Avec un niveau moyen d'émission inférieur à -10dBm : la variation de la tension audio est de 5 à 500mV crête. La communication se déroule en duplex intégral : elle est bidirectionnelle, le courant sur la ligne est la somme du courant continu et des deux courants variables émis par chaque poste.

A l'intérieur de chaque appareil un dispositif « antilocal » évite que le signal émis par le microphone ne soit transmis à l'écouteur.

• fonction de taxation :

Les services rendus par les réseaux de télécommunication doivent être rémunérés, en exploitation normale c'est l'abonné demandeur qui supporte le coût de la communication. Le prix d'une communication peut être proportionnellement à sa durée mesuré à partir de l'instant où l'abonné a répondu.

• supervision et libération:

La communication est supervisée pour détecter le raccrochage. Lorsqu'un des abonnés raccroche, le circuit établi entre eux est libéré.

I.5.5 : Signalisation :

La signalisation au sein d'un réseau de télécommunication fait référence à l'ensemble des échanges d'informations entre les équipements du réseau, nécessaire pour fournir et maintenir le service. Les informations de signalisation sont acheminées sous forme de messages caractérisés par des paquets de données à débit élevé.

La signalisation peut être transmise de deux manières :

- Signalisation en bande de base : les signaux permettant d'établir un appel entre deux commutateurs s'effectuent toujours dans le même canal que le transport de la voix. La signalisation prend la forme d'une série d'impulsions multifréquences un peu comme la tonalité de composition du numéro entre les commutateurs.
- Signalisation hors bande : la conversation et la signalisation ne prennent pas le même canal. Ce type de signalisation nécessite l'établissement d'un canal numérique pour l'échange des informations de signalisation appelé lien de signalisation ou canal sémaphore. Les canaux sémaphores de signalisation véhiculent les informations avec des débits de 56Kb/s aux États-Unis ou 64Kb/s pour les autres pays.

I.5.6 : l'acheminement :

L'acheminement détermine le chemin physique et logique entre deux points extrêmes du réseau, le principe est de choisir le chemin disponible le plus court.

Discussion

Le réseau RTC présente des avantages d'être disponible partout et avec des couts d'installation très faible. Cependant, en général, le débit est faible, la fiabilité de la communication sensible à la qualité de la ligne dans certaines conditions. Aussi, le nombre de service qui peut fournir est limité et présente aussi l'inconvénient de monopoliser la ligne téléphonique.

CHAPITRE II

Généralités sur la transmission

Préambule :

Dans ce chapitre, nous présentons les techniques de transmission de données. Cette transmission de données entre émetteur et récepteur est établie par une liaison en utilisant un support de transmission, munie d'équipement à ses extrémités. Il existe plusieurs modes de transmission de données, les technologies se sont imposées au fil des années selon les besoins des consommateurs, certaines plus efficaces et plus commercialisées que d'autres. L'objectif du chapitre est de décrire les différentes techniques et modes de transmission de données en général, et de transmission de la voix en particulier.

II.1.Chaine de transmission :

La structure conventionnelle d'une chaine de transmission comprend différents éléments repris à la figure (II.1). On part d'un message, d'une information, transmis sous forme d'un signal analogique (physique). Il convient alors de convertir ce signalé physique en un signal électrique par le biais d'un transducteur d'émission. La chaine de transmission de transmission se termine de même par un transducteur de réception, dont le but est de transformer le signal électrique en une grandeur physique adaptée au correspondant destinataire.

La figure suivante illustre le schéma de transmission de donnée

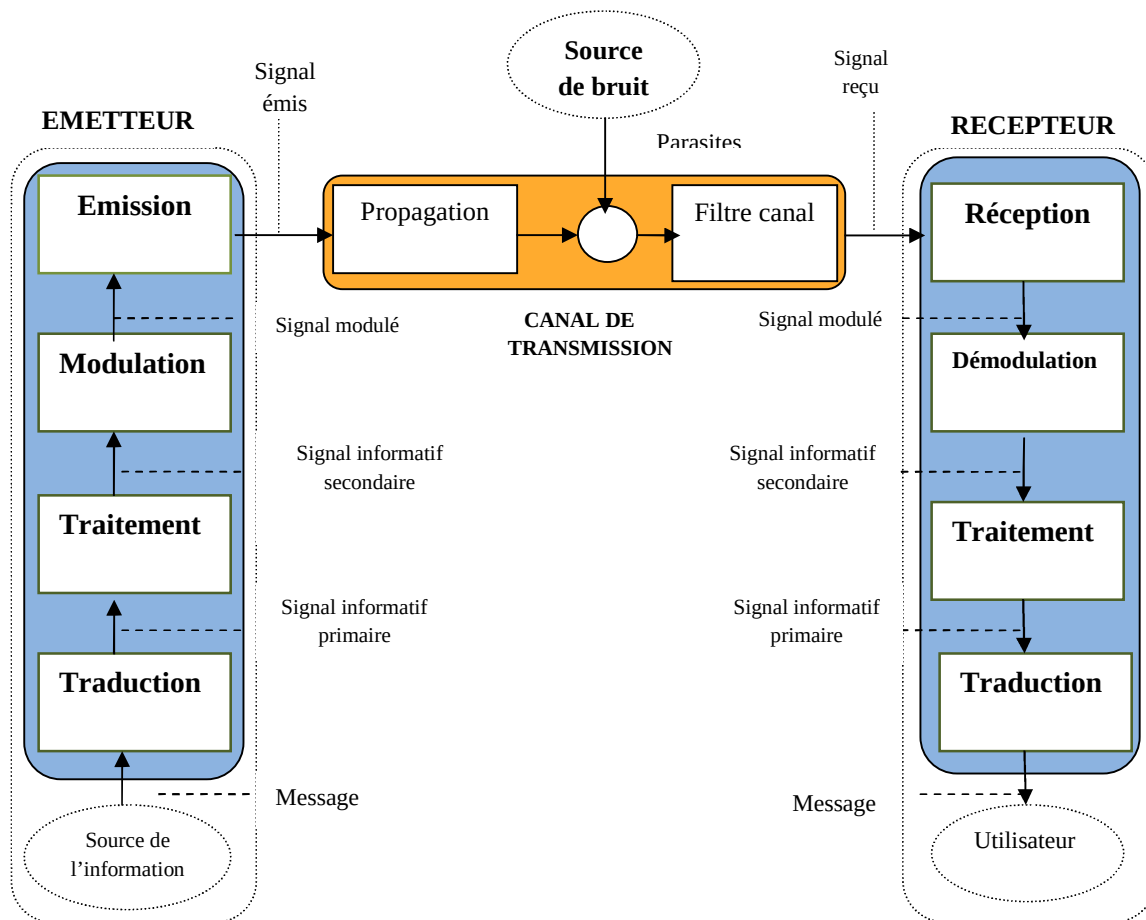


Figure II.1.schématisation d'une chaine de transmission

II.2 : description des fonctions utilisées :

II.2.1 : traduction : consiste à convertir un signal (voix par exemple) en un signal électrique.

II.2.2 : Traitement : la phase de traitement peut se réduire à une opération comme elle peut s'agir de diverses applications à savoir : la numérisation, la compression ...

II.2.3 : la modulation : la modulation consiste à adapter l'information à transmettre à un canal de communication, il existe plusieurs méthodes de modulation, celles qui sont analogiques et celles qui sont numériques.

II.2.4 : Amplification: cette étape intervient dans les systèmes de télécommunication à divers endroits de la chaîne de transmission. Les structures utilisées dépendent essentiellement de trois facteurs : la puissance de sortie, la gamme de fréquence des signaux d'entrées et de la bande passante.

II.2.5 : Canal : un canal indique le milieu dans lequel se propage le signal, mais il indique aussi la partie plus ou moins large du spectre occupé. Il peut être désigné par un code (en télévision par exemple), soit par sa fréquence centrale (en radio diffusion FM).sa largeur dépend du type d'application.

II.2.6 : Démodulation : c'est l'opération qui, à partir du signal modulé reçu du canal de transmission, permet de reconstituer le signal informatif (après traitement).

II.2.7 : Filtre : il a pour rôle de séparer les signaux utiles des autres, ils sont omniprésents, et à base de condensateurs et de bobines.

II.3: Les techniques de transmission de la voix :

II.3.1 : transmission analogique:

La transmission analogique ou en bande transposée consiste à faire circuler des informations sur un support physique de transmission sous la forme d'une onde. La transmission de données se fait par l'intermédiaire d'une onde porteuse, une onde simple dont le seul but est de transporter les données par modification de l'une de ces caractéristiques (amplitude, fréquence ou phase), c'est la raison pour laquelle la transmission analogique est généralement appelée transmission par modulation d'onde porteuse (Rajarshi et al, 2009 ; Stuart et Razavi, 2008). La modulation utilise deux signaux :

- Un signal modulant basse fréquence qui contient l'information et qui peut être analogique (voix) ou numérique (données informatiques).
- Un signal porteur de haute fréquence dont l'un des paramètres (amplitude, fréquence, phase) varie en fonction des évolutions du signal modulant.

Selon le paramètre de l'onde porteuse que l'on fait varier, on distinguera trois types de transmission analogiques :

II.3.1.1 : Transmission par modulation d'amplitude de la porteuse :

Le principe de la modulation d'amplitude est de faire varier l'amplitude d'un signal haute fréquence (la porteuse) au rythme d'un signal basse fréquence (le modulant) (voir figure II.2).

La porteuse est : $p(t) = v \cos(\omega_0 t)$ avec $\omega_0 = 2\pi f_0$

Le signal modulant est : $s(t)$

Le modulateur crée un signal modulé en amplitude ou un signal AM :

$$v(t) = v(1 + k s(t)) \cos(\omega_0 t) \dots\dots\dots (II.1)$$

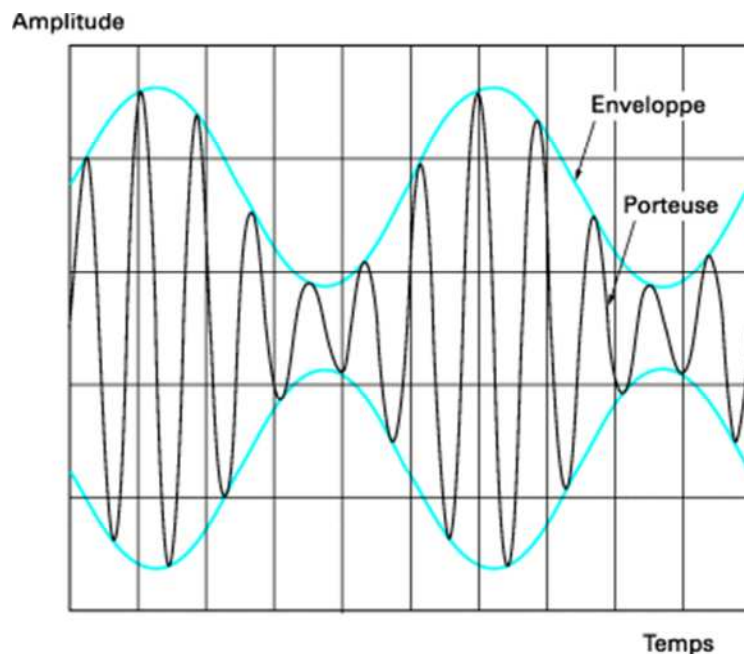


Figure II.2 : modulation d'amplitude

II.3.1.2 : Transmission par modulation de fréquence :

Elle consiste à faire varier la fréquence instantanée de la porteuse au rythme de celle du signal modulant.

Un signal peut s'écrire sous la forme :

$$p(t) = v \cos(\theta(t)) = v \cos(\omega_0 t + \varphi(t))$$

Où $\theta(t)$ est la phase instantanée et $\varphi(t)$ le déphasage du signal $p(t)$.

La fréquence instantanée du signal $p(t)$ est définie par la relation :

$$f(t) = \frac{1}{2\pi} \frac{d\theta}{dt} = f_0 + \frac{1}{2\pi} \frac{d\varphi}{dt}$$

Si : le signal modulant est : $s(t)$

Le signal porteuse est : $p(t) = v \cos(\omega_0 t)$ avec $\omega_0 \gg \Omega$ (la fréquence du signal modulant).

La fréquence instantanée s'écrit : $f(t) = f_0 + ks(t)$

Cette fréquence est donc modulée autour de f_0 par le signal $s(t)$. On a réalisé une modulation de fréquence. (Voir figure II.3)

On peut écrire alors : $\varphi(t) = 2\pi k \int_0^t s(u) du$

Le signal modulé en fréquence s'écrit alors :

$$v(t) = v \cos \left[\omega_0 t + 2\pi k \int_0^t s(u) du \right] \quad (\text{II.2})$$

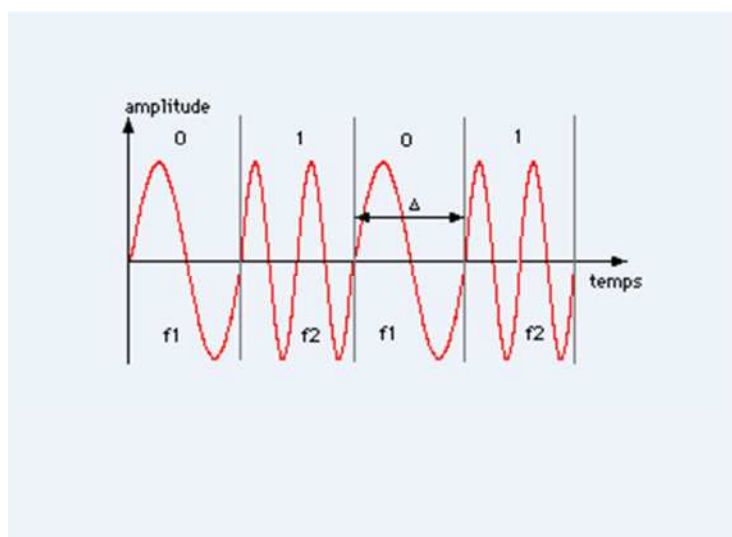


Figure II.3 : modulation de fréquence

II.3.1.3 : Transmission par modulation de phase :

La modulation de phase consiste à faire varier la phase du signal porteur en fonction de celle du signal modulant (voir figure II.4).

Si : le signal le signal modulant est $s(t)$, et la porteuse :

$$p(t) = v \cos(\omega_0 t) \quad \text{Avec } \omega_0 \gg \Omega$$

La phase instantanée du signal modulé en phase s'écrit :

$$\theta(t) = \omega_0 t + ks(t)$$

On peut alors écrire: $\varphi(t) = ks(t)$

La fréquence instantanée s'écrit donc :

$$f(t) = f_0 + \frac{k}{2\pi} \frac{ds}{dt}$$

Le signal modulé en phase s'écrit alors :

$$v(t) = v \cos(\omega_0 t + ks(t)) \quad (\text{II.3})$$

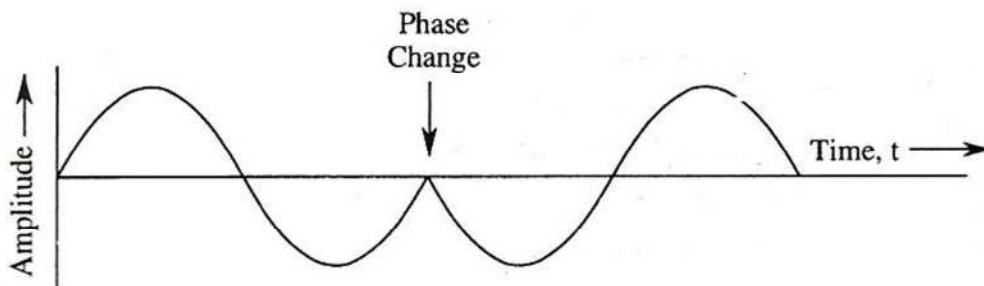


Figure II.4: modulation de phase

II.4 : Avantages et inconvénients de la transmission par modulation :

L'usage de la modulation permet ainsi de :

- diminuer les longueurs d'onde de signaux transmis (exemple : une fréquence de 100 MHz correspond à une longueur d'onde de 3,00 mètres) ;
- diminuer la présence de bruit lors de la transmission ;
- transmettre de signaux par voie hertzienne (exemple : la radio) ;
- transmettre simultanément sur le même support plusieurs informations par multiplexage fréquentiel.
- Transmettre l'information à des distances importantes.

Mais sa mise en œuvre est toutefois :

- plus complexe : risque l'augmentation de la dégradation du signal due aux équipements ;
- plus consommatrice de bande passante que pour le message d'origine. La bande de fréquences à l'émission est plus importante que celle du message d'origine (transposition vers les hautes fréquences).

II.3.2 : Transmission numérique :

La transmission numérique consiste à faire transiter les informations sur le support de communication sous forme de signaux numériques. Ainsi des données analogiques devront préalablement être numérisées avant d'être transmises. Ce mode de transmission est peu sensible aux parasites mais il nécessite une très grande bande passante (Gopal et al, 1984 ; Alireza et Siamak, 2004).

La numérisation du signal se fait grâce à un équipement appelé codec qui délivre une suite de données binaires. Le signal numérique est transmis et à l'arrivée, l'opération inverse est effectuée. Il s'agit de reconstituer correctement le signal sachant que la transformation analogique/numérique engendre toujours des pertes. La numérisation consiste en trois étapes :

II.3.2.1:l'échantillonnage :

- **Principe :**

Le signal analogique est un signal continu qui par définition contient un nombre infini d'éléments. L'échantillonnage consiste à prélever un nombre déterminé d'éléments (échantillons) qui seront suffisants pour reconstituer à l'arrivée un signal analogique de qualité.

Le principe de l'échantillonnage est de remplacer un signal analogique $X_p(t)$ par une suite périodique d'échantillons, c'est-à-dire des signaux identiques à $X_p(t)$ pendant une durée T , (voir figure II.5, II.6).

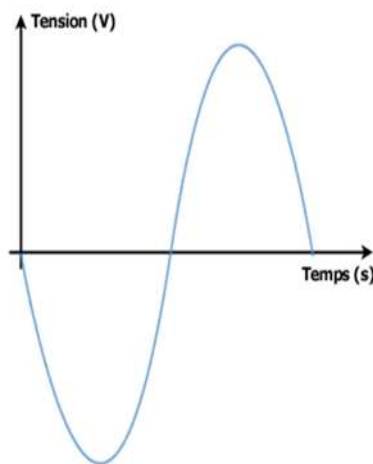


Figure II.5: signal analogique

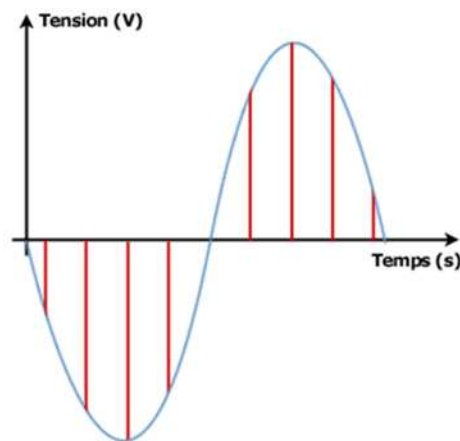


Figure II.6 : signal échantillonné

Prélevé à une fréquence $F_e(t)$ (fréquence d'échantillonnage), on peut écrire :

$$X_s(t) = X_p(t) \cdot X_e(t).$$

En vertu du théorème de Shannon, le signal original $X_p(t)$ peut être reproduit par le signal échantillonné $X_s(t)$ si $F_e > 2F_{\text{Max}}$, avec F_e la fréquence d'échantillonnage et F_{Max} : la fréquence maximum contenue dans le spectre du signal.

Pour la bande de fréquence de 300 à 3400Hz utilisée en téléphonie, pour une communication téléphonique $F_{\text{Max}} = 4\text{KHz}$, donc la fréquence d'échantillonnage $F_e = 8\text{KHz}$.

Ceci signifie qu'un signal de fréquence vocale est exploré 8000 fois par seconde. L'intervalle entre deux échantillons successifs se calcule comme suit : $T_e = 1/F_e = 125\text{ms}$

II.3.2.2 : La quantification :

- **Principe :**

Elle consiste à associer à chaque valeur de la tension retenue après l'échantillonnage un nombre binaire constitué de « 1 » et de « 0 » (voir figure II.7, II.8). C'est le convertisseur analogique-numérique ou CAN qui fait cette opération.

Le rôle de la quantification est de donner une image binaire d'un signal analogique : Passage analogique-numérique, Signal continu-signal discret, Tension-chiffre.

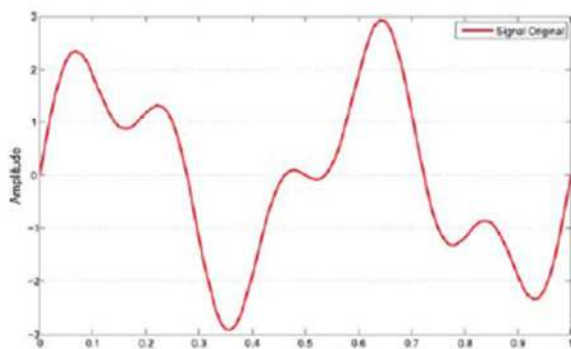


Figure II.7 : signal analogique

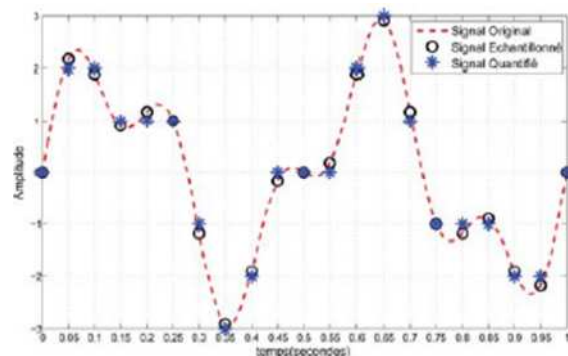


Figure II.8 : signal quantifié

- **Pas de quantification :**

Chaque niveau de tension est codé sur N bits, chaque bit peut prendre deux valeurs (0 ou 1), donc un convertisseur à N bits possède 2^N niveaux de quantification repartis de $-V_{\text{max}}$ à $+V_{\text{min}}$, ou V_{max} est la valeur maximale et V_{min} la valeur minimal du signal analogique ($V_{\text{min}} = -V_{\text{max}}$). On a ainsi un pas de quantification « Δ ou q » appelé aussi quantum (voir figure II.9), correspondant à la plus petite variation de tension que le convertisseur peut coder.

$$\Delta = \frac{V_{max} - V_{min}}{2^N} = \frac{2V_{max}}{2^N} \quad (II.4)$$

Plus le nombre de bit N est grand, (le pas de quantification q est petit) meilleure est la qualité de quantification (on obtient une meilleure précision).

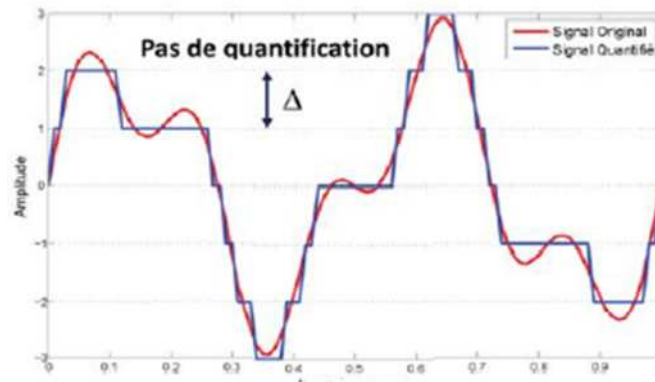


Figure II.9 : Pas de quantification

- **Bruit de quantification :**

Lors de la quantification une erreur de quantification " ε " (perte d'information) appelé aussi bruit de quantification se produit entre le signal échantillonné et la valeur du code correspondant à un niveau de tension (voir figure II.10), ce niveau étant la moyenne des tensions correspondantes à ce code, Gamme de tension $\Leftrightarrow$ code unique

Pour réduire ces pertes d'information, lors de la numérisation l'utilisateur doit choisir parmi celles possibles, la gamme $[V_{min} ; V_{max}]$ encadrant au plus près l'évolution du signal analogique.

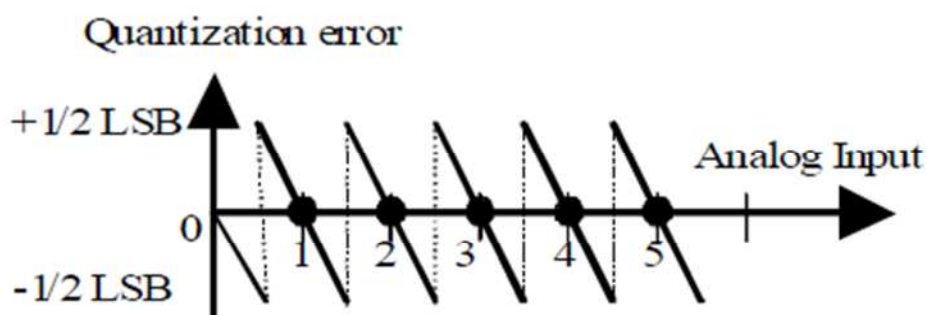


Figure II.10 : L'allure du bruit de quantification

L'évolution du bruit de quantification est une évolution en dent de scie avec une amplitude égale au quantum. En principe de quantification utilisé, les caractéristiques du bruit de quantification varient.

- **Les caractéristiques du bruit de quantification :**
- **Quantification linéaire par défaut :**

Dans ce cas le signal varie de 0 à E, (voir figure II.11, II.12), on code les 2^N niveaux de tension avec un pas de quantification :

$$q = \frac{E}{2^N}$$

L'erreur de quantification évolue entre 0 et q. $0 < \varepsilon < q$

La puissance du bruit de quantification est calculée comme suit :

$$P_\varepsilon = (\varepsilon^2(t)) = \frac{q^2}{3} \quad (\text{II.5})$$

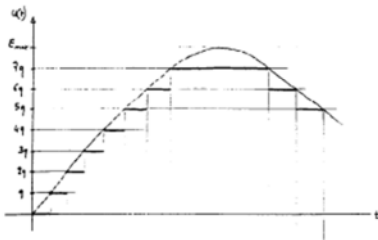


Figure II.11 : Quantification défaut

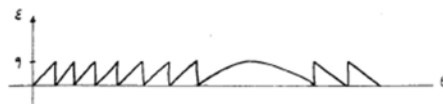


Figure II.12 : erreur de quantification

- **Quantification linéaire centrée :**

Dans la pratique on préfère effectuer une quantification centrée (voir figures : II.13, II.14). Dans ce cas le bruit de quantification évolue entre $-q/2$ et $+q/2$

$$-\frac{q}{2} < \varepsilon < +\frac{q}{2}$$

La puissance du bruit de quantification que l'on obtient est :

$$P_{\varepsilon} = (\varepsilon^2(t)) = \frac{q^2}{12} \quad (\text{II.6})$$

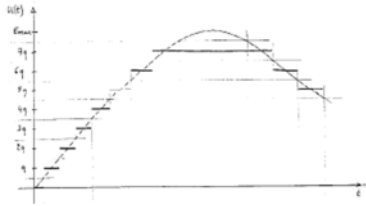
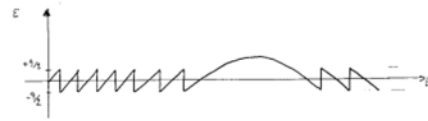


Figure II.13 : Quantification centrée



II.14 : Bruit de quantification

L'avantage de la quantification centrée par rapport à la quantification par défaut est que le bruit de quantification est plus faible.

- **Quantification non linéaire :**

Dans le cadre de l'audio (cadre de la téléphonie) plutôt que d'utiliser une loi de quantification linéaire on fait appel à une quantification logarithmique.

L'oreille est un capteur sensible de manière logarithmique au son (une amplitude sonore dix fois plus importante induit une sensation de volume sonore double). On adapte l'amplitude du bruit de quantification à l'amplitude du signal d'entrée de manière à avoir un rapport signal sur bruit ou SNR constant (ou presque) et meilleur qu'une quantification linéaire (voir figures II.15, II.16).

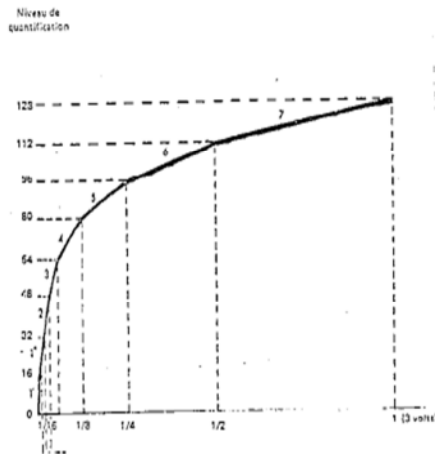


Figure II.15 : quantification logarithmique

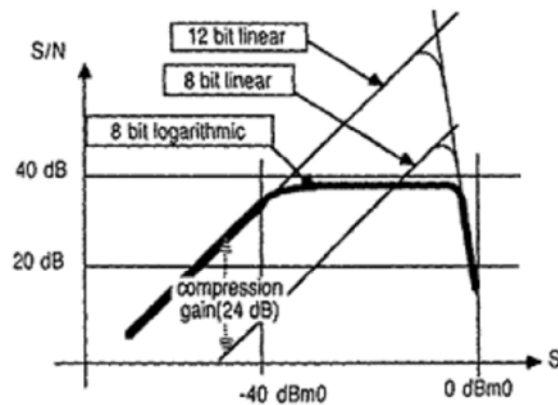


Figure II.16 : Evolution du SNR

II.3.2.3 : Le codage :

Lors d'une transmission d'information numérique, on est souvent amené à coder les mots à transmettre, soit pour minimiser la longueur moyenne des mots, soit pour se prémunir contre d'éventuelles erreurs de transmission (code correcteur d'erreurs). Chaque échantillon sera codé sur un ensemble de bits.

Pour que la transmission soit optimale, il est nécessaire que le signal soit codé de façon à faciliter sa transmission sur le support physique. Il existe pour cela différents systèmes de codage.

- **le codage à deux niveaux :**

Le signal peut prendre uniquement une valeur strictement négative ou strictement positive ($-X$ ou $+X$, X représentant une valeur de la grandeur physique permettant de transporter le signal).

Exemple : le codage NRZ (No Return to Zero ou Non Retour à Zéro) est le plus simple système de codage. Il consiste tout simplement à transformer les zéros en $-X$ et les 1 en $+X$ (voir figures II.17, II.18), de cette façon on a un codage bipolaire dans le quel le signal n'est jamais nul. Par conséquent le receveur peut déterminer la présence ou non d'un signal.

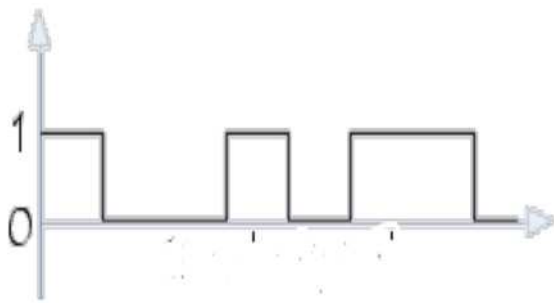


Figure II.17 : Signal original

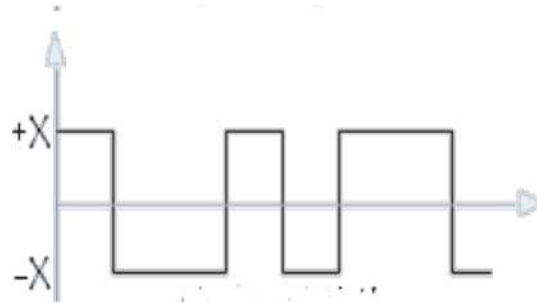


Figure II.18 : Codage NRZ

- **Codages à trois niveaux :**

Le signal peut prendre une valeur strictement négative, nulle ou strictement positive ($-X$, 0 ou $+X$).

Exemple : codage bipolaire simple : le codage bipolaire simple est un codage sur trois niveaux (voir figure II.19, II.20). Il propose donc trois états de la grandeur transportée sur le support physique :

- La valeur 0 lorsque le bit est 0.
- alternativement $-X$ et $+X$ lorsque le bit est à 1.

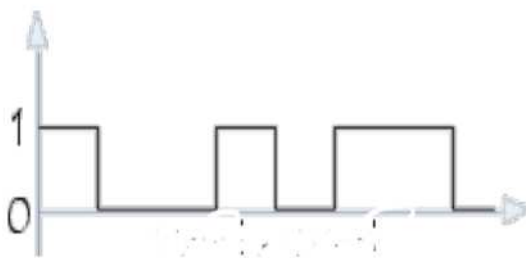


Figure II.19 : Signal original

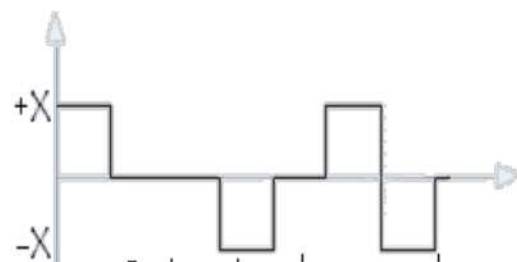


Figure II.20 : Codage bipolaire simple

II.3.3 : Transmission en bande de base :

II.3.3.1 : principe :

La transmission en bande de base consiste à transmettre directement le signal sur le support sans transposition de fréquence.

Cela peut se faire par exemple pour des signaux numériques en utilisant un modem dit à bande de base. Celui-ci utilise directement des supports physiques de types métallique (paires torsadées ou câble coaxiaux) ou fibre optique.

Dans la plupart des cas, les harmoniques supérieures à un certain rang peuvent ne pas être transmises sans qu'on note une altération inacceptable du signal. Les harmoniques d'un signal transmis sur une ligne sont diversement atténués, suivant leur fréquence, par la bande passante de la ligne. Si l'ensemble des harmoniques utiles du signal à transmettre se situent dans la bande passante de la ligne que l'on souhaite utiliser, on peut appliquer ce signal directement à l'entrée de la ligne. Il sera transmis sans atténuation notable à l'autre extrémité.

II.3.3.2 : Les inconvénients de la transmission en bande de base :

Les inconvénients majeurs de ce mode de transmission sont :

- Sensibilité aux parasites (bruits en $1/f$).
- Coût élevé pour la transmission sur fibre optique ou câble coaxial.
- Impossibilité de partage direct d'un même canal par plusieurs sources (on ne peut pas suivre plusieurs conversations à la fois !)

II.4 : Modes de transmission :

II.4.1 : Transmission série et parallèle

II.4.1.1 : liaison parallèle :

On désigne par liaison parallèle la transmission simultanée de N bits. Ces bits sont envoyés simultanément sur N voies différentes.

II.4.1.2 : liaison en série :

Dans une liaison en série, les données sont envoyées bit par bit sur la voie de transmission. Toutefois, étant donné que la plupart des processeurs traitent les informations de façon parallèle, il s'agit de transformer des données arrivant de façon parallèle en données en série au niveau de l'émetteur, et inversement au niveau du récepteur.

II.4.2 : transmission synchrone et asynchrone :**II.4.2.1 : Liaison synchrone :**

Dans laquelle émetteur et récepteur sont cadencés à la même horloge. Le récepteur reçoit de façon continue (même lorsque aucun bit n'est transmis) les informations au rythme ou l'émetteur les envoie. C'est pourquoi il est nécessaire qu'émetteur et récepteur soient cadencés à la même vitesse. De plus, des informations supplémentaires sont insérées afin de garantir l'absence d'erreurs lors de la transmission. Lors d'une transmission synchrone, les bits sont envoyés de façon successive sans séparation entre chaque caractère, il est donc nécessaire d'insérer des éléments de synchronisation, on parle alors de synchronisation au niveau caractère.

II.4.2.2 : Liaison asynchrone :

Dans ce cas l'émetteur et le récepteur ne sont pas cadencés à la même horloge. Chaque caractère est émis de façon irrégulière dans le temps (par exemple un utilisateur envoyant en temps réel des caractères saisis au clavier). Chaque caractère précédé d'une information indiquant le début de la transmission du caractère (l'information de début d'émission est appelée bit START) et termine par l'envoi d'une information de fin de transmission (appelée bit STOP, il peut éventuellement y avoir plusieurs bits STOP).

II.5 : sens de transmission :**II.5.1 : Mode simplex :**

La transmission ne peut se faire que dans un seul sens (de A vers B, ex : radio, télévision) voir figure (II.21).



Figure II.21. Transmission en mode simplex

II.5.2 : Mode half-duplex :

La transmission peut se faire dans les deux sens, mais pas en même temps. Bidirectionnel à alternatif (voir figure II.22), exemple : talkie-walkie.

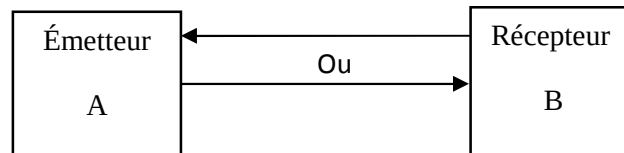


Figure II.22. Transmission en mode half duplex

II.5.3 : Mode full duplex :

La transmission peut se faire dans les deux sens simultanément, bidirectionnel à intégral (voir figure II.23), exemple : téléphone.

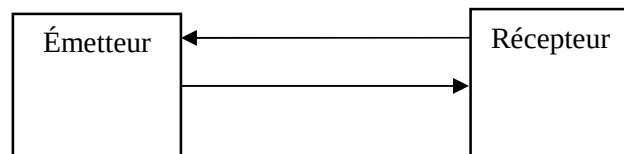


Figure II.23. Transmission en mode full duplex

II.6. Supports physiques de transmission (Stuart et Razavi, 2008):**II.6.1. Canal de transmission :**

La ligne de transmission, appelé aussi parfois canal de transmission ou voie de transmission, n'est pas forcément constituée d'un seul support physique de transmission, c'est pourquoi les machines d'extrémités (par opposition aux machines intermédiaires), appelées ETTD (Equipement Terminal de Traitement de Données) ou, en anglais DTE, (Data Terminal Equipment) possèdent chacune un équipement relatif au support physique auxquelles elles sont reliées, appelé ETCD (Equipement Terminal de Circuit de Données) ou, en anglais DCE

(Data Communication Equipment). On nomme circuit de données l'ensemble constitué des ETCD de chaque machine et de la ligne de données.

II.6.2. Supports de transmission :

II.6.2.1. Les lignes spécialisées :

Ce sont des lignes louées qui permettent la transmission de données à moyens et hauts débits (2,4 Kb/s à 140 Mb/s) en liaison point à point ou multipoints. Les 3 lignes les plus répandues sont les T1 (1.5 Mb/s), les T2 (6 Mb/s), et les T3 (45 Mb/s). Il existe aussi des lignes nettement plus rapides : ce sont les E1 (2 Mb/s), E2 (8 Mb/s), E3 (34 Mb/s), et E4 (140 Mb/s) qui sont inaccessibles pour les particuliers.

II.6.2.2. Les ondes hertziennes :

Elles supportent de grandes distances et de grandes capacités pour une propagation en visibilité directe (entre 50 et 80 Km). Les débits sont très élevés mais les transmissions sont sensibles aux perturbations et les possibilités d'écoute sont nombreuses.

II.6.2.3. Le satellite :

Effectivement les liaisons satellites sont très rapides, cependant elles sont dans un seul sens (réception), ainsi lorsque l'on veut aller sur un site on ne peut pas en donner l'ordre. La solution est simple, un modem sur une ligne téléphonique suffit pour envoyer ces informations. Canal Satellite et TPS sont déjà sur les rangs pour un Internet destiné au grand public. Le téléchargement de données s'effectue actuellement à un taux de transfert de 400Kb/s.

II.6.3. Ligne de transmission :

II.6.3.1. Ligne bifilaire :

Est le moyen de transmission le plus simple. On peut l'utiliser pour des connexions n'excédant pas 50 m, pour relier un DTE à un DCE, un modem par exemple.

Ce type de câble possède l'inconvénient d'être sensible d'une part à la diaphonie, à cause des capacités parasites, d'autre part aux perturbations électromagnétiques.

II.6.3.2. La paire torsadée :

Est un câble téléphonique constitué à l'origine de deux fils de cuivre isolés et enroulés l'un sur l'autre (d'où le nom). Actuellement on utilise plutôt des câbles constitués de 2 ou 4 paires torsadées. Elle est très répandue, d'une connexion facile et d'un faible coût mais elle possède une faible immunité aux bruits.

Il en existe les deux types suivants ; blindées (STP : Shielded Twisted Pair), non blindées (UTP : Unshielded Twisted Pair).

II.6.3.3. Le câble coaxial :

Il s'agit de 2 conducteurs concentriques isolés par une gaine diélectrique. Une gaine extérieure, parfois blindée, entoure l'ensemble. Cette construction assure au câble une bonne résistance aux signaux parasites.

Des débits de 10 Mb/s peuvent transiter par de tels câbles, dont la longueur n'excède pas quelques centaines de mètres. Il existe deux types de câbles coaxiaux, le câble épais (jaune) et le câble fin (noir).

II.6.3.4. La fibre optique :

La fibre optique peut être utilisée pour véhiculer des signaux de données sous forme de signaux optiques modulés. Elle est constituée d'un cylindre de verre extrêmement fin (le cœur) entouré d'une couche concentrique de verre (le revêtement) et joue le rôle d'un guide d'ondes lumineuses pour les longueurs d'ondes dans la gamme infrarouges : 850 nm, 1300 nm, 1500 nm. L'indice de réfraction de revêtement étant plus faible que celui du cœur, les faisceaux lumineux est réfléchi vers le cœur dès qu'il heurte le revêtement. Une onde optique guidée par des réflexions successives peut être représentée par des faisceaux de rayons que l'on appelle « modes ».

Les fibres optiques offrent des bandes passantes très importantes, de l'ordre du GHz. Leur atténuation est très faible et les signaux véhiculés n'étant, bien entendu, pas sujets aux interférences électriques, le taux d'erreur est également très faible (de l'ordre de 10^{-9}). De plus le câble optique est plus léger et moins encombrant qu'un support à cuivre. Néanmoins, le prix d'une liaison en fibre optique reste élevé, en raison des coupleurs optoélectroniques d'une

part, et de l'installation de la fibre proprement dite d'autre part. Notons également que les raccordements restent délicats à effectuer et qu'ils posent des problèmes d'affaiblissement. Il existe plusieurs types de fibre optique :

- la fibre multi mode à saut d'indice (diamètres : 50- 125 μm) dont la bande passante est de 40 MHz sur 1 Km.
- La fibre multi mode à gradient d'indice (mêmes diamètres) dont la bande passante atteint 500 MHz sur 1 km.
- La fibre monomode (diamètres : 2-8 μm) qui est la plus fine. Elle ne transmet qu'un seul mode est présente le plus grand potentiel de bande passante, de l'ordre de 100 GHz /km. Toutefois, sa mise en œuvre délicate est son coût élevé font qu'elle n'est pratiquement employée que par les opérateurs de télécommunications pour les très grandes distances.

Le tableau II.1. Représente les différentes caractéristiques des principaux câbles de transmission.

	Paire torsadée	Câble coaxial	Fibre optique
Coût	Faible	Moyen	Assez élevé
Bande passante	Moyenne	Large	Très large
Longueur maximale	Moyenne	Elevée	Elevée
Immunité aux interférences	Basse (UTP) Moyenne (STP)	Moyenne à élevée	Très élevée
Facilité de connexion	Simple	Variable	Difficile
Facilité d'installation	Variable	Variable	Difficile
Fiabilité	Bonne	Bonne	Très bonne

Tab II.1: *Caractéristiques physiques des différents câbles.*

Discussion :

Ce chapitre a fait l'objet d'une présentation des différentes méthodes de transmission de données, dans lequel nous avons décrit les différents modes et méthodes de transmissions ainsi que les supports de transmission. Le prochain chapitre exposera les différentes méthodes d'évaluation de la qualité du signal de parole.

CHAPITRE III

Méthodes d'évaluation de la qualité de parole

CHAPITRE III

Méthodes d'évaluation de la qualité de parole

Préambule :

L'objectif de ce chapitre est de présenter les méthodes utilisées pour l'évaluation de la qualité du signal de parole. Ces méthodes sont classées selon deux critères, à savoir : critère subjectif et critère objectif. Pour les méthodes objectives, un ensemble de paramètres issus d'un matériels plus sophistiqué et indispensable dans la majorité des cas doivent être utilisées. A défauts, les méthodes subjectives sont largement utilisées dans la littérature (Krebber, 1995 ; Raake, 2005) et donnent des résultats concluants. Ces méthodes permettent de prendre en comptes la perception de l'auditeur, et l'évaluation de la qualité se fait toujours sur les paramètres physiques d'une connexion. A cet effet, dans ce chapitre, nous nous intéressons particulièrement à ce type de méthodes et qui sera utilisé dans notre travail.

L'étude de la qualité de parole est un domaine important de la psycho acoustique pour ses applications dans le système sonore, la médecine ou comme ici, pour la téléphonie. Les applications de téléphonie ont évolué parallèlement avec l'apparition des nouvelles modalités de communication ; l'apparition des téléphones mobiles ou la transmission de la voix par internet conduit à l'adaptation (amélioration, création pour certains) de modèles permettant la prédiction de la qualité d'un signal de parole sur une ligne téléphonique.

III.1. Qualité et intelligibilité de la parole :

L'intelligibilité de la parole correspond à la capacité de comprendre un message linguistique contenu dans un signal de parole (Virole, 2001). L'intelligibilité est donc une mesure objective définie par le nombre de mots prononcé correctement identifiés par l'auditeur. Le meilleur moyen de juger l'intelligibilité est d'effectuer des tests d'écoute avec des sujets, dont la capacité d'écoute est normale, en utilisant par exemple la méthode du test de rime DRT (Diagnostic Rhyme Test). Celui-ci permet d'évaluer la transparence du message reçu à travers une mesure du degré de dégradation des caractéristiques élémentaires des consonnes (Buniet, 1997 ; Quackenbush, 1988). Une version plus générale du test DRT permet de tester tout aussi bien les voyelles que les consonnes (Fairbanks, 1958). Il existe cependant des moyens objectifs qui permettent d'estimer l'intelligibilité de la parole et qui sont largement utilisés dans la littérature, à savoir le test STI (Speech Transmission Index) (IEC-standard.60268-16 98), le test SII (Speech Intelligibility Index).

La qualité de parole est la capacité de véhiculer dans de bonnes conditions le message de parole sur une ligne téléphonique, en termes de disponibilité, délais de transmission, gigue, distorsion ... etc.

La qualité d'un signal de parole permet de prendre en compte la présence d'agents extérieurs « perturbateurs » (environnements bruyant, distorsions,...) la clarté du message peut en effet être affectée ce qui nuit au confort d'écoute. C'est donc une mesure subjective liée à l'aspect agréable du signal de parole par l'auditeur. Les tests de jugement de la qualité par les auditeurs sont les seuls moyens d'évaluation valables et sûrs. Mais comme l'intelligibilité, il existe des critères objectifs d'évaluation tels que le PESQ, MBSD, ...

III.2. Les aspects déterminant la qualité de la voix :

Les aspects déterminant la qualité de la voix sur un réseau sont le traitement de la voix, la clarté, le délai de bout en bout de l'écho, ils dépendent des différents composants de la chaîne de transmission, de leurs paramétrage, de l'architecture générale de la chaîne de transmission, ces aspects sont les suivants :

- Traitement de la voix : lors de l'émission du signal, la voix est traitée, c'est-à-dire codée et éventuellement compressée, avant d'être transmise.
- La clarté est la mesure de fidélité de la voix reçue par rapport à la voix émise.
- Le délai de bout en bout est le temps de propagation de la voix à travers le réseau de l'émetteur vers le récepteur.
- L'écho est le son émis par l'émetteur qui lui revient.

La problématique de qualité de voix est particulière car la voix attend de son transporteur autre chose que les données. La transmission de données classique (fichiers, messages,...) ne supporte aucune perte en ligne sous peine de graves conséquences pour l'interprétation et « l'utilisation de ces données par l'équipement récepteur, mais elle supporte en revanche une dérive importante en terme de durée d'acheminement. Peu importe qu'un paquet arrive avec 100ms de retard. Le comportement attendu pour la voix est exactement inverse : 1% ou 2% de perte de données de voix en ligne ne sont pas trop gênants pour la qualité de service mais

en revanche une variation fréquente de 100 millisecondes sur le délai de transit est catastrophique et rend le service inutilisable.

III.3. Les différents facteurs intervenant dans l'évaluation de la qualité de la voix :

Pendant la transmission de la voix ya des facteurs qui interviennent et qui infectent la qualité de la voix (du message transmis), ici on va citer quelques uns :

III.3.1.Le délai du transit :

Le délai de transi (ou end to end delay) est un des paramètres critique influençant fortement la qualité de service (QOS). C'est le temps que va mettre en moyenne un échantillon de voix pour traverser l'infrastructure entre deux interlocuteurs. Ce temps de transit comporte quatre composants :

- Délai d'échantillonnage.
- Délai de propagation.
- Délai de tranport.
- Délai de buffers de gigue.

Le délai d'échantillonnage est la durée de numérisaion de la voix à l'émission puis de conversion en signal voix à la reception.ce temps depend du type de codage choisi.le délai varie de quelques millisecondes à plus de 50millisecondes.

Le délai de propagation est la durée de transmission en ligne des données numérisées. Cette durée est normalment très faible par rapport aux autres composantes du délai de transit, de l'ordre de quelques millisecondes.

Le délai de transport est la durée passée à, traverser les routeurs, les commutateurs et les autres composants du réseau de l'infrastructure de téléphonie. L'ordre de grandeur est de plusieurs dizaines de millisecondes, voir certaines de millisecondes.

Le délai des buffers de gigue est le retard introduit à la reception en vue de laisser la variation de temps de transit, et donc de reduire la gigue de phase. L'ordre de grandeur est de 50 millisecondes.

III.3.2.Gigue de phase :

Variation de temps de transit ou gigue de phase, est la conséquence du fait que tous les échantillons de voix ne vont pas traverser le réseau à la même vitesse. Cela crée une déformation de la voix ou un hachage.

La gigue de phase est indépendante du délai de transit. Le délai peut être court et la gigue importante ou inversement. La gigue est une conséquence de congestion passagère sur le réseau, ce dernier ne pouvant plus transporter les données de manière constante dans le temps. La valeur de la gigue va de quelques millisecondes à quelques dizaines de millisecondes.

III.3.3.La perte des données :

La perte de données peut être entraînée par le codage et la compression, et aussi lors de conversion numérique/ analogique du signal de la voix (reconstitution du signal voix).

III.4.Méthodes d'évaluation de la qualité de la voix :

Les recherches en psychoacoustique sur la qualité de la parole ont entre autre, comme but de créer des outils d'évaluation. Ceux-ci donnent une valeur de la qualité globale de la parole perçue à travers un système de communications. Il existe des modèles permettant la prédiction de la qualité de la voix.

III.4.1.Critères subjectifs :

Les mesures de qualité subjective les plus fréquemment utilisés sont le MOS (Mean Opinion Score), le DMOS (Degradation Mean Opinion Score) et le CMOS (Comparaison Mean Opinion Score) (Keagy 00) et le modèle E.

III.4.1.1. MOS (Mean Opinion Score)

Le MOS est le résultat de l'analyse par catégories absolues ACR (Absolute Category Rating) dans laquelle un groupe d'auditeurs écoute un ensemble de fichiers audio et les évalue indépendamment, un à un, selon une échelle de notation sur la qualité perçue donnée dans le tableau suivant :

MOS	Qualité de voix
5	Excellente
4	Bonne
3	Juste
2	Pauvre
1	Mauvaise

Tab III.1 : Échelle MOS

III.4.1.2. CMOS (Comparaison Mean Opinion Score) :

Le CMOS est le résultat de l'analyse par catégories de comparaisons CCR (Comparaison Category Rating) dans laquelle on fournit à un groupe d'auditeurs des signaux par paires. L'auditeur compare les deux signaux de chaque paire en terme de qualité en précisant lequel est le meilleur et évalue la différence selon une échelle de notation bien définie. L'échelle est donnée dans le tableau qui suit :

Echelle	Qualité de la voix
3	Bien meilleure
2	Meilleure
1	Légèrement meilleure
0	A peu près équivalente
-1	Un peu moins bonne
-2	Moins bonne
-3	Nettement médiocre

Tab III.2 : Echelle CMOS

III.4.1.3. DMOS (Degradation Mean Opinion Score) :

Le DMOS il résulte de l'analyse par catégories de dégradations DCR (Degradation Category Rating) dans laquelle on fournit à un groupe d'auditeurs des paires de signaux pour comparer cette fois-ci la qualité en terme de dégradation. Contrairement au CMOS, les auditeurs savent a priori que la qualité du second signal est moins bonne que celle du premier. Ils doivent donc indiquer à quel point le second est justement moins bon suivant une échelle donnée dans le tableau si dessous :

Echelle	Qualité de la voix
5	Dépourvu de dégradation
4	Dégradation audible mais pas gênante
3	Dégradation un peu gênante
2	Dégradation gênante
1	Dégradation très gênante

Tab III.3. Echelle DMOS

D'une manière générale, lors de ces trois types de tests, la qualité du signal de parole dépend de la personne qui la juge et l'évalue. Sa façon de percevoir met en jeu l'expérience passé, l'environnement dans lequel elle s'est déroulée, son humeur et ses attentes. Ainsi, afin de diminuer l'effet subjectif sur l'évaluation de la qualité vocale, les notes des participants pour une condition de test donnée sont moyennées pour obtenir la note moyenne d'opinion.

On désigne par MOS, comme terme général les trois tests subjectifs déjà définis. Par définition donc, le MOS est un sondage auprès d'un échantillon de personnes représentatives du reste de la population. Lors de ce sondage, les auditeurs sont invités à écouter et à juger. Le jugement se fait à travers l'attribution d'une note sanctionnant la qualité perçue du signal de parole qu'ils ont écouté. La moyenne des notes attribuées constitue donc le MOS. L'avantage du MOS est qu'il quantifie la qualité perçue par les auditeurs participants aux tests. C'est donc une évaluation réelle, fiable et correcte de la qualité des signaux mis en jeu.

III.4.2. Analyse et présentation des résultats :

La distinction entre qualité de parole et qualité de bruit de fond est faite avant d'attribuer une dernière note à la qualité globale (annexe A). Les auditeurs sont alors invités à donner trois notes pour chaque séquence de test. Celle-ci est composée de trois phrases séparées par des phases de silence de durées suffisantes pour permettre à l'auditeur de donner une note. Pour l'évaluation du signal vocal, l'auditeur se concentre uniquement sur le signal vocal. Pour l'évaluation du bruit de fond, il se concentre uniquement sur le bruit de fond. Pour l'évaluation de la qualité globale, il donne ses appréciations en fonction du tout. Seules les notes relatives au critère de la qualité globale seront utilisées pour l'évaluation finale. Les autres critères vont juste permettre d'analyser la façon d'attribuer les notes par les auditeurs et de vérifier la cohérence. Il est à signaler aussi qu'avant le calcul de la moyenne des notes sur l'ensemble des sujets (auditeurs), une étude des notes individuelles est indispensable afin d'écarter les sujets dont les notes sont aberrantes. Pour dire qu'une note est aberrante ya des critères à suivre.

Soit u_{ijk} la note d'évaluation de l'échantillon k par l'observateur i dans la condition de test j et N le nombre d'observateurs. La quantité u_{ijk} présente des variations dues aux différences dans les jugements entre les N sujets. La moyenne des notes ou le MOS est le premier calcul à effectuer pour analyser les résultats du test subjectif :

$$\bar{u}_{jk} = \frac{1}{N} \sum_{i=1}^N u_{ijk} \quad (\text{III.1})$$

Pour évaluer la fiabilité des résultats, on associe à chaque MOS un intervalle de confiance à 95% des notes moyennes de chaque échantillon. Cet intervalle est calculé à partir de l'écart type empirique δ_{jk} et de la note moyenne $\bar{u}_{jk}$, est donné par :

$$[\bar{u}_{jk} - \Delta_{jk}, \bar{u}_{jk} + \Delta_{jk}] \quad (\text{III.2})$$

Ou

$$\Delta_{jk} = 1.96 \frac{\delta_{jk}}{\sqrt{N}} \quad \text{Et} \quad \delta_{jk} = \sqrt{\sum_{i=1}^N \frac{(u_{ijk} - \bar{u}_{jk})^2}{(N-1)}}$$

Il est possible de rejeter un observateur qui n'a pas fait le test correctement (faute d'attention, de mauvaise compréhension, ...). Le critère est basé sur le calcul du kurtosis qui permet d'éliminer les notes non cohérentes. On définit pour cette raison les paramètres suivant : β_{2jk} , m_4 et m_2 , qui sont le kurtosis et les moments empiriques d'ordre 4 et 2 tels que :

$$\beta_{2jk} = \frac{m_4}{(m_2)^2} \quad \text{Et} \quad m_l = \sum_{i=1}^N (u_{ijk} - \bar{u}_{jk})^l, \quad l = 2, 4.$$

Selon cette recommandation, si $2 \leq \beta_{2jk} \leq 4$, la distribution des notes par présentation est considérée comme normale. On calcule ainsi L_i et R_i , deux compteurs associés à l'observateurs i , le test j et l'échantillon k tel que :

$$L_i = R_i = 0 ;$$

Pour $j = 1, 2, \dots, j$, et $k = 1, 2, \dots, k$

Si $2 \leq \beta_{2jk} \leq 4$ alors :

$$\text{Si } u_{ijk} \geq \bar{u}_{jk} + 2\delta_{jk}, \quad \text{alors } R_i = R_i + 1$$

$$\text{Si } u_{ijk} \leq \bar{u}_{jk} - 2\delta_{jk}, \quad \text{alors } L_i = L_i + 1$$

Sinon

$$\text{Si } u_{ijk} \geq \bar{u}_{jk} + \sqrt{20\delta_{jk}}, \quad \text{alors } R_i = R_i + 1$$

$$\text{Si } u_{ijk} \leq \bar{u}_{jk} - \sqrt{20\delta_{jk}}, \quad \text{alors } L_i = L_i + 1$$

FinSi

FinPour

J : le nombre total de conditions de tests.

K : le nombre total d'échantillons

A partir de L_i et R_i , les relations suivantes permettent de rejeter l'auditeur dont les résultats sont aberrants ou $j \times k$ est le nombre total de notes attribuées par chaque utilisateur. Alors :

Si $\frac{L_i+R_i}{j \times k} > 0.05$ et $\left| \frac{L_i-R_i}{L_i+R_i} \right| < 0.3$, les notes du sujet i doivent être supprimées. A signaler que le rapport $\frac{L_i+R_i}{j \times k}$ prend en compte les écarts importants par rapport à la moyenne et le rapport $\left| \frac{L_i-R_i}{L_i+R_i} \right|$ modélise les écarts autour de la moyenne. Après l'élimination des sujets dont les notes ne pourraient pas conduire à des estimations cohérentes, on passe à une analyse statistique plus poussée des résultats et qui s'avère indispensable pour une représentation plus significative des résultats et qui s'avère indispensable pour une représentation plus significative des notes MOS. Cette analyse statistique est effectuée par le test de student t-test et l'analyse de la variance ANOVA.

III.4.2.1. Analyse statistique par t-test et ANOVA :

Le but de l'analyse statistique est de pouvoir tirer des conclusions ayant une signification statistique. L'hypothèse de départ est que les notes sont conçues à partir de distributions de moyennes différentes et de variances identiques. Pour étudier la validité de cette hypothèse, on définit les deux hypothèses de test suivantes :

$$H_0: V_1 = V_2$$

$$H_1: V_1 \neq V_2$$

Pour tester les deux hypothèses on fait appel au test statistique t désigné souvent dans la littérature par le terme t-test.

III.4.2.2.1. T-test ou test de student

C'est un test paramétrique qui repose sur des comparaisons de moyennes issues de deux échantillons indépendants (deux algorithmes de test ou de groupes). Il s'applique à tout test statistique dont la distribution est de type student quand l'hypothèse nulle est vraie. Une variable T suit une distribution de student ou une loi de student si elle est définie par le quotient entre une variable aléatoire Z centrée et réduite et une variable U indépendante de Z et distribuée suivant la loi X^2 , tel que :

$$T = \frac{Z}{\sqrt{U/K}} \quad (\text{III.3})$$

K est le degré de liberté de la loi de student. Le t-test est calculé ainsi,

$$t = \frac{V_1 - V_2}{S_{v_1 - v_2}} \quad (\text{III.4})$$

Avec

$$S_{v_1 - v_2} = \sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}$$

Où S_i^2 et n_i sont respectivement la variance et le nombre d'échantillons (nombre de participant par groupe) du groupe i et t c'est la valeur critique du test de student.

La valeur de t est physiquement, un rapport entre la différence des moyennes des groupes et la dispersion ou la variabilité des notes entre groupes. A partir de la valeur de t , on décide de retenir ou de rejeter l'hypothèse H_0 , mais avant on se fixe un certain seuil de confiance α qui est la probabilité de rejeter H_0 alors qu'elle est vraie. On choisit souvent $\alpha = 0.05$ autrement dit, avec cette valeur de α on rejette correctement H_0 avec une probabilité de 0.95. On détermine aussi le degré de liberté qui est égale à $(n_1 + n_2 - 2)$. Donc suivant la valeur de α et le degré de liberté, on obtient une certaine valeur critique issue de la table de la loi de student (Annexe C). si la valeur de t qu'on vient de calculer est supérieure à cette valeur critique, on rejette H_0 et considère que les moyennes des deux échantillons sont différentes.

III.4.2.2.2. ANOVA (Analysis Of Variance)

L'analyse de variance ANOVA est une généralisation du t-test à plus de deux groupes. Les deux hypothèses de test sont :

$$H_0 : V_1 = V_2 = \dots V_M$$

H_1 : au moins une seule moyenne diffère des autres.

On compare la variance inter-groupe à la variance intra-groupe. La variance intra-groupe est la variance à l'intérieur de chaque groupe et la variance intra-groupe reflète les variances entre groupes. Techniquement, l'ANOVA est un test unilatéral ; on cherche seulement à savoir si la variance inter-groupe est plus grande que la variance intra-groupe en calculant la statistique de Fisher :

$$F = \frac{\text{variance inter-groupe}}{\text{variance intra-groupe}} \quad (\text{III.5})$$

Cette statistique a une distribution F de degré de liberté (dl_1, dl_2) tel que $dl_1 = M-1$ et $dl_2 = M(n-1)$, avec M le nombre de populations ou de groupes et n le nombre d'échantillons par groupe. La valeur critique du test de statistique F issue de la table de Fisher (Annexe B) dépend des degrés de liberté dl_1 et dl_2 ainsi que du paramètre α . On compare cette valeur critique avec la valeur expérimentale obtenue par l'équation (III.5). Si la valeur expérimentale est supérieure à la valeur critique, on rejette l'hypothèse H_0 tout en concluant qu'au moins une des moyenne V_i diffère des autres. Dans le cas contraire on accepte l'hypothèse H_0 et on conclut que les moyennes de toutes les populations sont égales.

III.5.Critères objectifs

Les mesures objectives de qualité des signaux vocaux les plus communément utilisées sont citées et classées dans le tableau suivant :

Mesure dans le domaine temporel	Mesure dans le domaine fréquentiel	Mesure dans le domaine perceptuel
SNR segSNR	IS CD WSS LLR	BSD, MBSD PSQM PESQ

Tab III.4 : Les mesures objectives de qualité des signaux les plus utilisées

Les critères temporels et fréquentiels se basent essentiellement sur l'évaluation de la qualité en termes de comparaison de distorsion de forme signal de référence et signal débruité, sans tenir compte de l'aspect perceptif. Certes, c'est une condition nécessaire mais non suffisante dans la mesure où deux signaux sont pratiquement de même forme peuvent être perçus différemment (Wang 92), d'où l'intérêt d'introduire le facteur psychoacoustique pour tout système ayant pour objectif de conserver la qualité de la parole. Diverses mesures objectives perceptuelles sont élaborées, elles sont essentiellement dédiées au codage de la parole.

III.5.1.SNR segmental (segSNR) :

Le SNR (Signal to Noise Ratio) segmental segSNR est la mesure de qualité objective la plus utilisée dans le domaine temporel. Il définit la moyenne des SNRs issus de plusieurs segments de courte durée (15 à 20 ms) :

$$SNR_{seg} = \frac{10}{M} \sum_{m=0}^{M-1} \log_{10} \frac{\sum_{i=mN}^{mN+N-1} s^2(i)}{\sum_{i=mN}^{mN+N-1} (s(i) - \hat{s}(i))^2} \quad (\text{III.6})$$

Avec

$s(i)$: Le signal de référence.

$\hat{s}(i)$: Le signal débruité.

N : la longueur d'un segment.

M : le nombre total de segments.

Le SNR segmental souffre de deux limitations : d'abord si le signal de parole contient des segments de silence, ce qui est très probable, le $s(i)$ sera nul et n'importe quelle quantité de bruit entraînera un SNR en dB négatif pour ce segment, du coup le SNR total sera biaisé par cette quantité. Ce problème peut être résolu partiellement en choisissant un seuil d'énergie au delà duquel le SNR segmental sera calculé. En suite il faut nécessairement que les deux signaux comparés soient alignés temporellement car ce critère est très sensible au déphasage.

III.5.2.BSD et MBSD :

La mesure BSD (Bark Spectral Distorsion) (Wang, 1992) est parmi les premiers critères à avoir incorporé des notions de relation avec notre système d'audition dans l'évaluation de la qualité de la parole (Wang, 1997). Le BSD a pour objectif de mesurer la distorsion entre le signal de référence et celui codé, dans le domaine de Bark. La sensation de force sonore connue sous le nom de sonie est mise en jeu pour calculer cette distorsion. En effet, la distorsion totale est

la moyenne de la distance euclidienne entre la sonie du signal de référence et celle du signal débruité.

Le MBSD (Modified Bark Spectral Distorsion) (Yang, 1999) introduit le seuil de masquage du bruit pour calculer la distorsion dans le BSD ; l'idée est de ne tenir compte que de la distorsion audible. Effectivement, tout ce qui est au dessous du seuil de masquage du bruit est imperceptible à l'oreille humaine. Par conséquent, la distorsion totale est la moyenne de la différence entre les sonies du signal de référence et du signal débruité pondérée par un paramètre s'annulant lorsque la distorsion est inaudible.

III.5.3.PSQM :

Le PSQM (Perceptual Speech Quality Measure) est une version typique aux signaux de parole décrite par la norme P.861 (P.861 98). Elle constitue un cas particulier du critère PAQM (Perceptual Audio Quality Measure) (Beerends, 1992) dédié aux signaux audio en général. L'intérêt de concevoir une mesure unique pour la parole revient aux différences de caractéristiques existant entre la parole et la musique. Le PSQM exploite à son tour les propriétés de la perception auditive humaine pour évaluer la qualité de la parole. La moyenne

de la différence en sonie , désignés dans la norme par le terme bruit perturbateur, constitue la note PSQM attribuée à la qualité du signal codé.

III.5.4.PESQ :

Le PESQ (Perceptual Evaluation of Speech Quality) est l'évaluation de la qualité vocale perçue désignée dans la norme P.862 (P862 00) comme moyen adapté aux codes vocaux et aux mesures de bout en bout. De ce fait, d'autres facteurs supplémentaires sont pris en considération pour mieux simuler les conditions réelles, à savoir le temps de propagation, les distorsions dues aux erreurs de transmission, les pertes de paquets. Néanmoins, il existe d'autres facteurs techniques et applications (P862 00) pour les quels la méthode d'évaluation PESQ n'a pas été encore validée. Notamment les artefactes causés par les algorithmes de réduction de bruit ainsi que les dégradations liées à l'interaction bidirectionnelle lors de la transmission comme l'effet d'écho.

Ce critère se base sur un calcul de distance perceptuelle (différence audible entre la représentation perceptuelle du signal de référence et celle de test) suivie d'un modèle cognitif qui permet de prendre en compte le fait qu'une dégradation n'a pas le même impact selon son contexte (segment de parole ou non) et sa distribution (localisé ou non). La note d'évaluation PESQ finale est une combinaison linéaire de la valeur de perturbation moyenne et de la valeur de perturbation asymétrique moyenne.

Le PESQ permet d'évaluer la qualité d'écoute dans de nombreuses conditions de dégradation (perte de paquets, distorsion due au codage et bruit du côté émission ...), aboutissant à une corrélation proche des notes subjectives.

il existe d'autres mesures de qualités qui ne requièrent pas de disposer d'un signal de référence . elles sont de deux types paramétriques et non paramétriques, et elles ont pour objectif d'évaluer en continu la qualité de la parole dans les réseaux de télécommunication.

Les mesures non paramétriques, basées sur le signal, comparent ce dernier à un signal artificiel issu d'une base de référence (codebook). Les mesures paramétriques telles que le modèle E sont basées sur les mesures physiques (délai, écho, bruit, ...) du système à évaluer et exploitent les propriétés du conduit vocal pour déterminer la distorsion.

III.5.5.Le modèle E :

Le but de ce genre de modèle est donc bien de connaître tous les paramètres liés à une connexion, de la bouche de locuteur jusqu'à l'oreille de l'auditeur. Ainsi le modèle peut prédire la qualité globale perçue par un utilisateur lors d'une communication. Le modèle E est principalement basé sur des résultats de jugement de qualité (test de perception).

Ce type de modèle permet d'étudier toutes les dégradations perçues par les utilisateurs, mais n'utilise qu'une seule dimension. La prédiction est obtenue sur une échelle « psychologique » d'évaluation.

Pour planifier un réseau complet il est difficile de connaître la qualité globale avant la mise en place de ce système. Le modèle E permet la prédiction de la qualité vocale transmise par une ligne téléphonique (Moller et Raak, 2002). Ce type de modèle estime la qualité globale du système grâce à des mesures instrumentales et une description du système par de nombreux paramètres scalaires le modèle E est la compilation des différents modèles de l'« European Telecommunications Standards Institute » (ETSI ETR 250, 1996). Il est également le modèle recommandé par l'ITU-T pour la conception de réseaux de qualité (ITU-T 2005).

Le paramètre le plus important pour cette étude est « Impairment Factor » I_e

III.5.5.1.Equipement Impairment Factor :

Le facteur de dégradation I_e permet de quantifier les dégradations dues à la phase de « codage-décodage » donc dues au codec (ITU-T, Appendice I, 2003), il est défini pour quantifier les dégradations se produisant pendant la transmission. Celles –ci sont typiquement non linéaires et peuvent varier dans le temps chaque I_e est spécifique aux caractéristiques perceptives d'un codec, il est calculé en fonction de la dégradation affecté au signal.

L'hypothèse fondamentale du modèle E postule que toutes les catégories de dégradation sont additives sur une échelle de qualité appropriée. Selon ce modèle, les différentes dégradations individuelles lors de l'utilisation de plusieurs codecs, peuvent être additionnées sur l'échelle R du modèle E de la manière suivante :

$$I_{e_{tandem}} = \sum_i I_{e_{individual}} \quad (III.7)$$

III.5.5.2. L'échelle R :

La valeur de sortie du modèle E, se situe sur cette échelle perceptive. Elle permet de situer certains codecs dont l'évaluation est standardisée. Elle permet d'exprimer la qualité grâce à 5 valeurs.

Le principe de la formule est de partir d'un certain capital de QOS égale à R_0 et lui imputer les dommages causés par les différents aspects de la transmission.

$$R = R_0 - I_s - I_d - I_e + A \quad (\text{III.8})$$

R_0 : coefficient initial signal/bruit « capitale initial QOS ».

I_s : coefficient de dommage simultanés avec l'émission de la voix (bruit de fond...).

I_d : coefficient de dommages dus au délai de transmission et de transport.

I_e : coefficient de dommages de distorsion causés par les équipements.

A : coefficient d'amélioration.

L'échelle R est comprise entre 0 et 100, ou 0 correspond à la plus mauvaise qualité et 100 la meilleure. La valeur sur l'échelle R de la téléphonie bande étroite utilisé sur un canal RNIS propre de bruit vaut $R=93.2$.

Echelle	Qualité de la voix
90-100	Excellente
80-90	Bonne
80-70	Juste
70-60	Pauvre
60-0	Mauvaise

Tab III.5. Echelle MOS

Le facteur R ainsi calculé de 0 à 100 permet de déduire directement un coefficient MOS de 0 à 5. Dans la réalité les valeurs de R oscillent entre 50 et 93.2. La moyenne de la valeur R sur les services en production se situe dans la fourchette 70 à 80.

III.6. Conversion R/MOS

Les valeurs MOS peuvent être converties en valeurs sur l'échelle R, du modèle E (ITU-T rec. G.107, 2005), grâce à la formule suivante :

$$\text{MOS} = \begin{cases} 1 \\ 1+0.035*R+R(R-60)(100-R)*7.10^{-6} \\ 4.5 \end{cases} \begin{cases} \text{pour : } R < 0 \\ \text{pour : } 0 < R < 100 \\ \text{Pour : } R > 100 \end{cases} \quad (\text{III.9})$$

Le graphique suivant présente la courbe MOS= f (R). Le facteur MOS est la prédiction humaine tirée d'un protocole de test précis. L'inflexion vers le haut indique qu'à partir d'un certain niveau de qualité, l'augmentation de celle-ci reflétée par le facteur R, est moins perçue par l'utilisateur. On observe le même phénomène en bas de la courbe, une diminution de R moins perçue par l'utilisateur quand la qualité est déjà très dégradée.

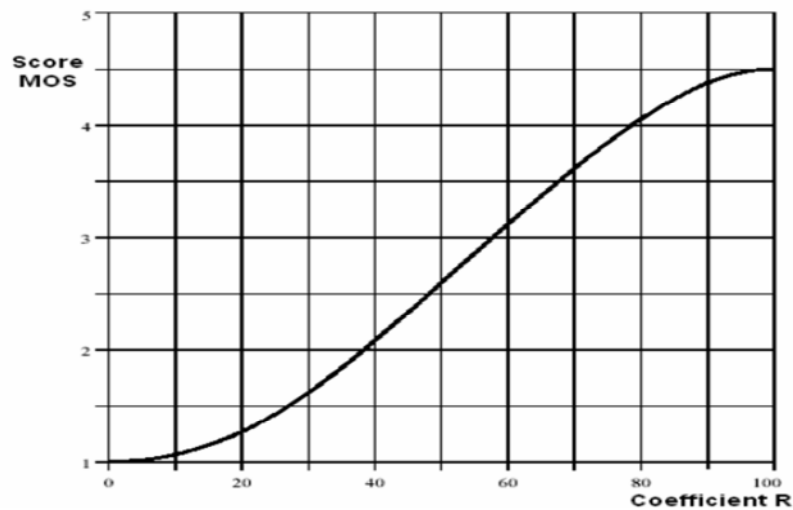


Figure III.1. Conversion MOS/R

Discussion :

Deux sortes de méthodes ont été présentées dans ce chapitre. Les méthodes subjectives restent de loin plus intéressantes en fonction des moyens disponibles. De plus, les résultats obtenus par ce type de méthodes sont significatifs. En effet, les paramètres utilisés sont référencés dans les bases de données de test du signal de parole. Dans le prochain chapitre, nous appliquons la méthode MOS qui est une méthode subjective basée sur un ensemble de tests pour évaluer la qualité du signal de parole.

Préambule :

L'objectif de ce chapitre est de présenter les méthodes utilisées pour l'évaluation de la qualité du signal de parole. Ces méthodes sont classées selon deux critères, à savoir : critère subjectif et critère objectif. Pour les méthodes objectives, un ensemble de paramètres issus d'un matériels plus sophistiqué et indispensable dans la majorité des cas doivent être utilisées. A défauts, les méthodes subjectives sont largement utilisées dans la littérature (Krebber, 1995 ; Raake, 2005) et donnent des résultats concluants. Ces méthodes permettent de prendre en comptes la perception de l'auditeur, et l'évaluation de la qualité se fait toujours sur les paramètres physiques d'une connexion. A cet effet, dans ce chapitre, nous nous intéressons particulièrement à ce type de méthodes et qui sera utilisé dans notre travail.

L'étude de la qualité de parole est un domaine important de la psycho acoustique pour ses applications dans le système sonore, la médecine ou comme ici, pour la téléphonie. Les applications de téléphonie ont évolué parallèlement avec l'apparition des nouvelles modalités de communication ; l'apparition des téléphones mobiles ou la transmission de la voix par internet conduit à l'adaptation (amélioration, création pour certains) de modèles permettant la prédiction de la qualité d'un signal de parole sur une ligne téléphonique.

III.1. Qualité et intelligibilité de la parole :

L'intelligibilité de la parole correspond à la capacité de comprendre un message linguistique contenu dans un signal de parole (Virole, 2001). L'intelligibilité est donc une mesure objective définie par le nombre de mots prononcé correctement identifiés par l'auditeur. Le meilleur moyen de juger l'intelligibilité est d'effectuer des tests d'écoute avec des sujets, dont la capacité d'écoute est normale, en utilisant par exemple la méthode du test de rime DRT (Diagnostic Rhyme Test). Celui-ci permet d'évaluer la transparence du message reçu à travers une mesure du degré de dégradation des caractéristiques élémentaires des consonnes (Buniet, 1997 ; Quackenbush, 1988). Une version plus générale du test DRT a permis de tester tout aussi bien les voyelles que les consonnes (Fairbanks, 1958). Il existe cependant des moyens objectifs qui permettent d'estimer l'intelligibilité de la parole et qui sont largement utilisés dans la littérature, à savoir le test STI (Speech Transmission Index) (IEC-standard.60268-16 98), le test SII (Speech Intelligibility Index).

La qualité de parole est la capacité de véhiculer dans de bonnes conditions le message de parole sur une ligne téléphonique, en termes de disponibilité, délais de transmission, gigue, distorsion ... etc.

La qualité d'un signal de parole permet de prendre en compte la présence d'agents extérieurs « perturbateurs » (environnements bruyant, distorsions,...) la clarté du message peut en effet être affectée ce qui nuit au confort d'écoute. C'est donc une mesure subjective liée à l'aspect agréable du signal de parole par l'auditeur. Les tests de jugement de la qualité par les auditeurs sont les seuls moyens d'évaluation valables et sûrs. Mais comme l'intelligibilité, il existe des critères objectifs d'évaluation tels que le PESQ, MBSD, ...

III.2. Les aspects déterminant la qualité de la voix :

Les aspects déterminant la qualité de la voix sur un réseau sont le traitement de la voix, la clarté, le délai de bout en bout de l'écho, ils dépendent des différents composants de la chaîne de transmission, de leurs paramétrage, de l'architecture générale de la chaîne de transmission, ces aspects sont les suivants :

- Traitement de la voix : lors de l'émission du signal, la voix est traitée, c'est-à-dire codée et éventuellement compressée, avant d'être transmise.
- La clarté est la mesure de fidélité de la voix reçue par rapport à la voix émise.
- Le délai de bout en bout est le temps de propagation de la voix à travers le réseau de l'émetteur vers le récepteur.
- L'écho est le son émis par l'émetteur qui lui revient.

La problématique de qualité de voix est particulière car la voix attend de son transporteur autre chose que les données. La transmission de données classique (fichiers, messages,...) ne supporte aucune perte en ligne sous peine de graves conséquences pour l'interprétation et « l'utilisation de ces données par l'équipement récepteur, mais elle supporte en revanche une dérive importante en terme de durée d'acheminement. Peu importe qu'un paquet arrive avec 100ms de retard. Le comportement attendu pour la voix est exactement inverse : 1% ou 2% de perte de données de voix en ligne ne sont pas trop gênants pour la qualité de service mais

en revanche une variation fréquente de 100 millisecondes sur le délai de transit est catastrophique et rend le service inutilisable.

III.3. Les différents facteurs intervenant dans l'évaluation de la qualité de la voix :

Pendant la transmission de la voix ya des facteurs qui interviennent et qui infectent la qualité de la voix (du message transmis), ici on va citer quelques uns :

III.3.1.Le délai du transit :

Le délai de transi (ou end to end delay) est un des paramètres critique influençant fortement la qualité de service (QOS). C'est le temps que va mettre en moyenne un échantillon de voix pour traverser l'infrastructure entre deux interlocuteurs. Ce temps de transit comporte quatre composants :

- Délai d'échantillonnage.
- Délai de propagation.
- Délai de tranport.
- Délai de buffers de gigue.

Le délai d'échantillonnage est la durée de numérisaion de la voix à l'émission puis de conversion en signal voix à la reception.ce temps depend du type de codage choisi.le délai varie de quelques millisecondes à plus de 50millisecondes.

Le délai de propagation est la durée de transmission en ligne des données numérisées. Cette durée est normalment très faible par rapport aux autres composantes du délai de transit, de l'ordre de quelques millisecondes.

Le délai de transport est la durée passée à, traverser les routeurs, les commutateurs et les autres composants du réseau de l'infrastructure de téléphonie. L'ordre de grandeur est de plusieurs dizaines de millisecondes, voir certaines de millisecondes.

Le délai des buffers de gigue est le retard introduit à la reception en vue de laisser la variation de temps de transit, et donc de reduire la gigue de phase. L'ordre de grandeur est de 50 millisecondes.

III.3.2.Gigue de phase :

Variation de temps de transit ou gigue de phase, est la conséquence du fait que tous les échantillons de voix ne vont pas traverser le réseau à la même vitesse. Cela crée une déformation de la voix ou un hachage.

La gigue de phase est indépendante du délai de transit. Le délai peut être court et la gigue importante ou inversement. La gigue est une conséquence de congestion passagère sur le réseau, ce dernier ne pouvant plus transporter les données de manière constante dans le temps. La valeur de la gigue va de quelques millisecondes à quelques dizaines de millisecondes.

III.3.3.La perte des données :

La perte de données peut être entraînée par le codage et la compression, et aussi lors de conversion numérique/ analogique du signal de la voix (reconstitution du signal voix).

III.4.Méthodes d'évaluation de la qualité de la voix :

Les recherches en psychoacoustique sur la qualité de la parole ont entre autre, comme but de créer des outils d'évaluation. Ceux-ci donnent une valeur de la qualité globale de la parole perçue à travers un système de communications. Il existe des modèles permettant la prédiction de la qualité de la voix.

III.4.1.Critères subjectifs :

Les mesures de qualité subjective les plus fréquemment utilisés sont le MOS (Mean Opinion Score), le DMOS (Degradation Mean Opinion Score) et le CMOS (Comparaison Mean Opinion Score) (Keagy 00) et le modèle E.

III.4.1.1. MOS (Mean Opinion Score)

Le MOS est le résultat de l'analyse par catégories absolues ACR (Absolute Category Rating) dans laquelle un groupe d'auditeurs écoute un ensemble de fichiers audio et les évalue indépendamment, un à un, selon une échelle de notation sur la qualité perçue donnée dans le tableau suivant :

MOS	Qualité de voix
5	Excellente
4	Bonne
3	Juste
2	Pauvre
1	Mauvaise

Tab III.1 : Échelle MOS

III.4.1.2. CMOS (Comparaison Mean Opinion Score) :

Le CMOS est le résultat de l'analyse par catégories de comparaisons CCR (Comparaison Category Rating) dans laquelle on fournit à un groupe d'auditeurs des signaux par paires. L'auditeur compare les deux signaux de chaque paire en terme de qualité en précisant lequel est le meilleur et évalue la différence selon une échelle de notation bien définie. L'échelle est donnée dans le tableau qui suit :

Echelle	Qualité de la voix
3	Bien meilleure
2	Meilleure
1	Légèrement meilleure
0	A peu près équivalente
-1	Un peu moins bonne
-2	Moins bonne
-3	Nettement médiocre

Tab III.2 : Echelle CMOS

III.4.1.3. DMOS (Degradation Mean Opinion Score) :

Le DMOS il résulte de l'analyse par catégories de dégradations DCR (Degradation Category Rating) dans laquelle on fournit à un groupe d'auditeurs des paires de signaux pour comparer cette fois-ci la qualité en terme de dégradation. Contrairement au CMOS, les auditeurs savent a priori que la qualité du second signal est moins bonne que celle du premier. Ils doivent donc indiquer à quel point le second est justement moins bon suivant une échelle donnée dans le tableau si dessous :

Echelle	Qualité de la voix
5	Dépourvu de dégradation
4	Dégradation audible mais pas gênante
3	Dégradation un peu gênante
2	Dégradation gênante
1	Dégradation très gênante

Tab III.3. Echelle DMOS

D'une manière générale, lors de ces trois types de tests, la qualité du signal de parole dépend de la personne qui la juge et l'évalue. Sa façon de percevoir met en jeu l'expérience passé, l'environnement dans lequel elle s'est déroulée, son humeur et ses attentes. Ainsi, afin de diminuer l'effet subjectif sur l'évaluation de la qualité vocale, les notes des participants pour une condition de test donnée sont moyennées pour obtenir la note moyenne d'opinion.

On désigne par MOS, comme terme général les trois tests subjectifs déjà définis. Par définition donc, le MOS est un sondage auprès d'un échantillon de personnes représentatives du reste de la population. Lors de ce sondage, les auditeurs sont invités à écouter et à juger. Le jugement se fait à travers l'attribution d'une note sanctionnant la qualité perçue du signal de parole qu'ils ont écouté. La moyenne des notes attribuées constitue donc le MOS. L'avantage du MOS est qu'il quantifie la qualité perçue par les auditeurs participants aux tests. C'est donc une évaluation réelle, fiable et correcte de la qualité des signaux mis en jeu.

III.4.2. Analyse et présentation des résultats :

La distinction entre qualité de parole et qualité de bruit de fond est faite avant d'attribuer une dernière note à la qualité globale (annexe A). Les auditeurs sont alors invités à donner trois notes pour chaque séquence de test. Celle-ci est composée de trois phrases séparées par des phases de silence de durées suffisantes pour permettre à l'auditeur de donner une note. Pour l'évaluation du signal vocal, l'auditeur se concentre uniquement sur le signal vocal. Pour l'évaluation du bruit de fond, il se concentre uniquement sur le bruit de fond. Pour l'évaluation de la qualité globale, il donne ses appréciations en fonction du tout. Seules les notes relatives au critère de la qualité globale seront utilisées pour l'évaluation finale. Les autres critères vont juste permettre d'analyser la façon d'attribuer les notes par les auditeurs et de vérifier la cohérence. Il est à signaler aussi qu'avant le calcul de la moyenne des notes sur l'ensemble des sujets (auditeurs), une étude des notes individuelles est indispensable afin d'écarter les sujets dont les notes sont aberrantes. Pour dire qu'une note est aberrante ya des critères à suivre.

Soit u_{ijk} la note d'évaluation de l'échantillon k par l'observateur i dans la condition de test j et N le nombre d'observateurs. La quantité u_{ijk} présente des variations dues aux différences dans les jugements entre les N sujets. La moyenne des notes ou le MOS est le premier calcul à effectuer pour analyser les résultats du test subjectif :

$$\bar{u}_{jk} = \frac{1}{N} \sum_{i=1}^N u_{ijk} \quad (\text{III.1})$$

Pour évaluer la fiabilité des résultats, on associe à chaque MOS un intervalle de confiance à 95% des notes moyennes de chaque échantillon. Cet intervalle est calculé à partir de l'écart type empirique δ_{jk} et de la note moyenne $\bar{u}_{jk}$, est donné par :

$$[\bar{u}_{jk} - \Delta_{jk}, \bar{u}_{jk} + \Delta_{jk}] \quad (\text{III.2})$$

Ou

$$\Delta_{jk} = 1.96 \frac{\delta_{jk}}{\sqrt{N}} \quad \text{Et} \quad \delta_{jk} = \sqrt{\sum_{i=1}^N \frac{(u_{ijk} - \bar{u}_{jk})^2}{(N-1)}}$$

Il est possible de rejeter un observateur qui n'a pas fait le test correctement (faute d'attention, de mauvaise compréhension, ...). Le critère est basé sur le calcul du kurtosis qui permet d'éliminer les notes non cohérentes. On définit pour cette raison les paramètres suivant : β_{2jk} , m_4 et m_2 , qui sont le kurtosis et les moments empiriques d'ordre 4 et 2 tels que :

$$\beta_{2jk} = \frac{m_4}{(m_2)^2} \quad \text{Et} \quad m_l = \sum_{i=1}^N (u_{ijk} - \bar{u}_{jk})^l, \quad l = 2, 4.$$

Selon cette recommandation, si $2 \leq \beta_{2jk} \leq 4$, la distribution des notes par présentation est considérée comme normale. On calcule ainsi L_i et R_i , deux compteurs associés à l'observateurs i , le test j et l'échantillon k tel que :

$$L_i = R_i = 0 ;$$

Pour $j = 1, 2, \dots, j$, et $k = 1, 2, \dots, k$

Si $2 \leq \beta_{2jk} \leq 4$ alors :

$$\text{Si } u_{ijk} \geq \bar{u}_{jk} + 2\delta_{jk}, \quad \text{alors } R_i = R_i + 1$$

$$\text{Si } u_{ijk} \leq \bar{u}_{jk} - 2\delta_{jk}, \quad \text{alors } L_i = L_i + 1$$

Sinon

$$\text{Si } u_{ijk} \geq \bar{u}_{jk} + \sqrt{20\delta_{jk}}, \quad \text{alors } R_i = R_i + 1$$

$$\text{Si } u_{ijk} \leq \bar{u}_{jk} - \sqrt{20\delta_{jk}}, \quad \text{alors } L_i = L_i + 1$$

FinSi

FinPour

J : le nombre total de conditions de tests.

K : le nombre total d'échantillons

A partir de L_i et R_i , les relations suivantes permettent de rejeter l'auditeur dont les résultats sont aberrants ou $j \times k$ est le nombre total de notes attribuées par chaque utilisateur. Alors :

Si $\frac{L_i+R_i}{j \times k} > 0.05$ et $\left| \frac{L_i-R_i}{L_i+R_i} \right| < 0.3$, les notes du sujet i doivent être supprimées. A signaler que le rapport $\frac{L_i+R_i}{j \times k}$ prend en compte les écarts importants par rapport à la moyenne et le rapport $\left| \frac{L_i-R_i}{L_i+R_i} \right|$ modélise les écarts autour de la moyenne. Après l'élimination des sujets dont les notes ne pourraient pas conduire à des estimations cohérentes, on passe à une analyse statistique plus poussée des résultats et qui s'avère indispensable pour une représentation plus significative des résultats et qui s'avère indispensable pour une représentation plus significative des notes MOS. Cette analyse statistique est effectuée par le test de student t-test et l'analyse de la variance ANOVA.

III.4.2.1. Analyse statistique par t-test et ANOVA :

Le but de l'analyse statistique est de pouvoir tirer des conclusions ayant une signification statistique. L'hypothèse de départ est que les notes sont conçues à partir de distributions de moyennes différentes et de variances identiques. Pour étudier la validité de cette hypothèse, on définit les deux hypothèses de test suivantes :

$$H_0: V_1 = V_2$$

$$H_1: V_1 \neq V_2$$

Pour tester les deux hypothèses on fait appel au test statistique t désigné souvent dans la littérature par le terme t-test.

III.4.2.2.1.T-test ou test de student

C'est un test paramétrique qui repose sur des comparaisons de moyennes issues de deux échantillons indépendants (deux algorithmes de test ou de groupes). Il s'applique à tout test statistique dont la distribution est de type student quand l'hypothèse nulle est vraie. Une variable T suit une distribution de student ou une loi de student si elle est définie par le quotient entre une variable aléatoire Z centrée et réduite et une variable U indépendante de Z et distribuée suivant la loi X^2 , tel que :

$$T = \frac{Z}{\sqrt{U/K}} \quad (\text{III.3})$$

K est le degré de liberté de la loi de student. Le t-test est calculé ainsi,

$$t = \frac{V_1 - V_2}{S_{v_1 - v_2}} \quad (\text{III.4})$$

Avec

$$S_{v_1 - v_2} = \sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}$$

Où S_i^2 et n_i sont respectivement la variance et le nombre d'échantillons (nombre de participant par groupe) du groupe i et t c'est la valeur critique du test de student.

La valeur de t est physiquement, un rapport entre la différence des moyennes des groupes et la dispersion ou la variabilité des notes entre groupes. A partir de la valeur de t , on décide de retenir ou de rejeter l'hypothèse H_0 , mais avant on se fixe un certain seuil de confiance α qui est la probabilité de rejeter H_0 alors qu'elle est vraie. On choisit souvent $\alpha = 0.05$ autrement dit, avec cette valeur de α on rejette correctement H_0 avec une probabilité de 0.95. On détermine aussi le degré de liberté qui est égale à $(n_1 + n_2 - 2)$. Donc suivant la valeur de α et le degré de liberté, on obtient une certaine valeur critique issue de la table de la loi de student (Annexe C). si la valeur de t qu'on vient de calculer est supérieure à cette valeur critique, on rejette H_0 et considère que les moyennes des deux échantillons sont différentes.

III.4.2.2.2. ANOVA (Analysis Of Variance)

L'analyse de variance ANOVA est une généralisation du t-test à plus de deux groupes. Les deux hypothèses de test sont :

$$H_0 : V_1 = V_2 = \dots V_M$$

H_1 : au moins une seule moyenne diffère des autres.

On compare la variance inter-groupe à la variance intra-groupe. La variance intra-groupe est la variance à l'intérieur de chaque groupe et la variance intra-groupe reflète les variances entre groupes. Techniquement, l'ANOVA est un test unilatéral ; on cherche seulement à savoir si la variance inter-groupe est plus grande que la variance intra-groupe en calculant la statistique de Fisher :

$$F = \frac{\text{variance inter-groupe}}{\text{variance intra-groupe}} \quad (\text{III.5})$$

Cette statistique a une distribution F de degré de liberté (dl_1, dl_2) tel que $dl_1 = M-1$ et $dl_2 = M(n-1)$, avec M le nombre de populations ou de groupes et n le nombre d'échantillons par groupe. La valeur critique du test de statistique F issue de la table de Fisher (Annexe B) dépend des degrés de liberté dl_1 et dl_2 ainsi que du paramètre α . On compare cette valeur critique avec la valeur expérimentale obtenue par l'équation (III.5). Si la valeur expérimentale est supérieure à la valeur critique, on rejette l'hypothèse H_0 tout en concluant qu'au moins une des moyenne V_i diffère des autres. Dans le cas contraire on accepte l'hypothèse H_0 et on conclut que les moyennes de toutes les populations sont égales.

III.5.Critères objectifs

Les mesures objectives de qualité des signaux vocaux les plus communément utilisées sont citées et classées dans le tableau suivant :

Mesure dans le domaine temporel	Mesure dans le domaine fréquentiel	Mesure dans le domaine perceptuel
SNR segSNR	IS CD WSS LLR	BSD, MBSD PSQM PESQ

Tab III.4 : Les mesures objectives de qualité des signaux les plus utilisées

Les critères temporels et fréquentiels se basent essentiellement sur l'évaluation de la qualité en termes de comparaison de distorsion de forme signal de référence et signal débruité, sans tenir compte de l'aspect perceptif. Certes, c'est une condition nécessaire mais non suffisante dans la mesure où deux signaux sont pratiquement de même forme peuvent être perçus différemment (Wang 92), d'où l'intérêt d'introduire le facteur psychoacoustique pour tout système ayant pour objectif de conserver la qualité de la parole. Diverses mesures objectives perceptuelles sont élaborées, elles sont essentiellement dédiées au codage de la parole.

III.5.1.SNR segmental (segSNR) :

Le SNR (Signal to Noise Ratio) segmental segSNR est la mesure de qualité objective la plus utilisée dans le domaine temporel. Il définit la moyenne des SNRs issus de plusieurs segments de courte durée (15 à 20 ms) :

$$SNR_{seg} = \frac{10}{M} \sum_{m=0}^{M-1} \log_{10} \frac{\sum_{i=mN}^{mN+N-1} s^2(i)}{\sum_{i=mN}^{mN+N-1} (s(i) - \hat{s}(i))^2} \quad (\text{III.6})$$

Avec

$s(i)$: Le signal de référence.

$\hat{s}(i)$: Le signal débruité.

N : la longueur d'un segment.

M : le nombre total de segments.

Le SNR segmental souffre de deux limitations : d'abord si le signal de parole contient des segments de silence, ce qui est très probable, le $s(i)$ sera nul et n'importe quelle quantité de bruit entraînera un SNR en dB négatif pour ce segment, du coup le SNR total sera biaisé par cette quantité. Ce problème peut être résolu partiellement en choisissant un seuil d'énergie au delà duquel le SNR segmental sera calculé. En suite il faut nécessairement que les deux signaux comparés soient alignés temporellement car ce critère est très sensible au déphasage.

III.5.2.BSD et MBSD :

La mesure BSD (Bark Spectral Distorsion) (Wang, 1992) est parmi les premiers critères à avoir incorporé des notions de relation avec notre système d'audition dans l'évaluation de la qualité de la parole (Wang, 1997). Le BSD a pour objectif de mesurer la distorsion entre le signal de référence et celui codé, dans le domaine de Bark. La sensation de force sonore connue sous le nom de sonie est mise en jeu pour calculer cette distorsion. En effet, la distorsion totale est

la moyenne de la distance euclidienne entre la sonie du signal de référence et celle du signal débruité.

Le MBSD (Modified Bark Spectral Distorsion) (Yang, 1999) introduit le seuil de masquage du bruit pour calculer la distorsion dans le BSD ; l'idée est de ne tenir compte que de la distorsion audible. Effectivement, tout ce qui est au dessous du seuil de masquage du bruit est imperceptible à l'oreille humaine. Par conséquent, la distorsion totale est la moyenne de la différence entre les sonies du signal de référence et du signal débruité pondérée par un paramètre s'annulant lorsque la distorsion est inaudible.

III.5.3.PSQM :

Le PSQM (Perceptual Speech Quality Measure) est une version typique aux signaux de parole décrite par la norme P.861 (P.861 98). Elle constitue un cas particulier du critère PAQM (Perceptual Audio Quality Measure) (Beerends, 1992) dédié aux signaux audio en général. L'intérêt de concevoir une mesure unique pour la parole revient aux différences de caractéristiques existant entre la parole et la musique. Le PSQM exploite à son tour les propriétés de la perception auditive humaine pour évaluer la qualité de la parole. La moyenne

de la différence en sonie , désignés dans la norme par le terme bruit perturbateur, constitue la note PSQM attribuée à la qualité du signal codé.

III.5.4.PESQ :

Le PESQ (Perceptual Evaluation of Speech Quality) est l'évaluation de la qualité vocale perçue désignée dans la norme P.862 (P862 00) comme moyen adapté aux codes vocaux et aux mesures de bout en bout. De ce fait, d'autres facteurs supplémentaires sont pris en considération pour mieux simuler les conditions réelles, à savoir le temps de propagation, les distorsions dues aux erreurs de transmission, les pertes de paquets. Néanmoins, il existe d'autres facteurs techniques et applications (P862 00) pour les quels la méthode d'évaluation PESQ n'a pas été encore validée. Notamment les artefactes causés par les algorithmes de réduction de bruit ainsi que les dégradations liées à l'interaction bidirectionnelle lors de la transmission comme l'effet d'écho.

Ce critère se base sur un calcul de distance perceptuelle (différence audible entre la représentation perceptuelle du signal de référence et celle de test) suivie d'un modèle cognitif qui permet de prendre en compte le fait qu'une dégradation n'a pas le même impact selon son contexte (segment de parole ou non) et sa distribution (localisé ou non). La note d'évaluation PESQ finale est une combinaison linéaire de la valeur de perturbation moyenne et de la valeur de perturbation asymétrique moyenne.

Le PESQ permet d'évaluer la qualité d'écoute dans de nombreuses conditions de dégradation (perte de paquets, distorsion due au codage et bruit du côté émission ...), aboutissant à une corrélation proche des notes subjectives.

il existe d'autres mesures de qualités qui ne requièrent pas de disposer d'un signal de référence . elles sont de deux types paramétriques et non paramétriques, et elles ont pour objectif d'évaluer en continu la qualité de la parole dans les réseaux de télécommunication.

Les mesures non paramétriques, basées sur le signal, comparent ce dernier à un signal artificiel issu d'une base de référence (codebook). Les mesures paramétriques telles que le modèle E sont basées sur les mesures physiques (délai, écho, bruit, ...) du système à évaluer et exploitent les propriétés du conduit vocal pour déterminer la distorsion.

III.5.5.Le modèle E :

Le but de ce genre de modèle est donc bien de connaître tous les paramètres liés à une connexion, de la bouche de locuteur jusqu'à l'oreille de l'auditeur. Ainsi le modèle peut prédire la qualité globale perçue par un utilisateur lors d'une communication. Le modèle E est principalement basé sur des résultats de jugement de qualité (test de perception).

Ce type de modèle permet d'étudier toutes les dégradations perçues par les utilisateurs, mais n'utilise qu'une seule dimension. La prédiction est obtenue sur une échelle « psychologique » d'évaluation.

Pour planifier un réseau complet il est difficile de connaître la qualité globale avant la mise en place de ce système. Le modèle E permet la prédiction de la qualité vocale transmise par une ligne téléphonique (Moller et Raak, 2002). Ce type de modèle estime la qualité globale du système grâce à des mesures instrumentales et une description du système par de nombreux paramètres scalaires le modèle E est la compilation des différents modèles de l'« European Telecommunications Standards Institute » (ETSI ETR 250, 1996). Il est également le modèle recommandé par l'ITU-T pour la conception de réseaux de qualité (ITU-T 2005).

Le paramètre le plus important pour cette étude est « Impairment Factor » I_e

III.5.5.1.Equipement Impairment Factor :

Le facteur de dégradation I_e permet de quantifier les dégradations dues à la phase de « codage-décodage » donc dues au codec (ITU-T, Appendice I, 2003), il est défini pour quantifier les dégradations se produisant pendant la transmission. Celles –ci sont typiquement non linéaires et peuvent varier dans le temps chaque I_e est spécifique aux caractéristiques perceptives d'un codec, il est calculé en fonction de la dégradation affecté au signal.

L'hypothèse fondamentale du modèle E postule que toutes les catégories de dégradation sont additives sur une échelle de qualité appropriée. Selon ce modèle, les différentes dégradations individuelles lors de l'utilisation de plusieurs codecs, peuvent être additionnées sur l'échelle R du modèle E de la manière suivante :

$$I_{e_{tandem}} = \sum_i I_{e_{individual}} \quad (III.7)$$

III.5.5.2. L'échelle R :

La valeur de sortie du modèle E, se situe sur cette échelle perceptive. Elle permet de situer certains codecs dont l'évaluation est standardisée. Elle permet d'exprimer la qualité grâce à 5 valeurs.

Le principe de la formule est de partir d'un certain capital de QOS égale à R_0 et lui imputer les dommages causés par les différents aspects de la transmission.

$$R = R_0 - I_s - I_d - I_e + A \quad (\text{III.8})$$

R_0 : coefficient initial signal/bruit « capitale initial QOS ».

I_s : coefficient de dommage simultanés avec l'émission de la voix (bruit de fond...).

I_d : coefficient de dommages dus au délai de transmission et de transport.

I_e : coefficient de dommages de distorsion causés par les équipements.

A : coefficient d'amélioration.

L'échelle R est comprise entre 0 et 100, ou 0 correspond à la plus mauvaise qualité et 100 la meilleure. La valeur sur l'échelle R de la téléphonie bande étroite utilisé sur un canal RNIS propre de bruit vaut $R=93.2$.

Echelle	Qualité de la voix
90-100	Excellente
80-90	Bonne
80-70	Juste
70-60	Pauvre
60-0	Mauvaise

Tab III.5. Echelle MOS

Le facteur R ainsi calculé de 0 à 100 permet de déduire directement un coefficient MOS de 0 à 5. Dans la réalité les valeurs de R oscillent entre 50 et 93.2. La moyenne de la valeur R sur les services en production se situe dans la fourchette 70 à 80.

III.6. Conversion R/MOS

Les valeurs MOS peuvent être converties en valeurs sur l'échelle R, du modèle E (ITU-T rec. G.107, 2005), grâce à la formule suivante :

$$\text{MOS} = \begin{cases} 1 \\ 1+0.035*R+R(R-60)(100-R)*7.10^{-6} \\ 4.5 \end{cases} \begin{cases} \text{pour : } R < 0 \\ \text{pour : } 0 < R < 100 \\ \text{Pour : } R > 100 \end{cases} \quad (\text{III.9})$$

Le graphique suivant présente la courbe MOS= f (R). Le facteur MOS est la prédiction humaine tirée d'un protocole de test précis. L'inflexion vers le haut indique qu'à partir d'un certain niveau de qualité, l'augmentation de celle-ci reflétée par le facteur R, est moins perçue par l'utilisateur. On observe le même phénomène en bas de la courbe, une diminution de R moins perçue par l'utilisateur quand la qualité est déjà très dégradée.

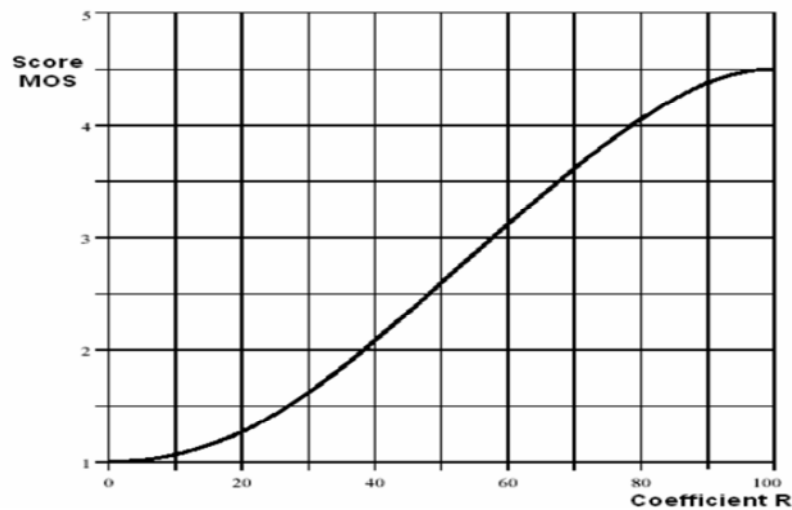


Figure III.1. Conversion MOS/R

Discussion :

Deux sortes de méthodes ont été présentées dans ce chapitre. Les méthodes subjectives restent de loin plus intéressantes en fonction des moyens disponibles. De plus, les résultats obtenus par ce type de méthodes sont significatifs. En effet, les paramètres utilisés sont référencés dans les bases de données de test du signal de parole. Dans le prochain chapitre, nous appliquons la méthode MOS qui est une méthode subjective basée sur un ensemble de tests pour évaluer la qualité du signal de parole.

CHAPITRE IV

Application des tests

CHAPITRE IV

Application des tests

Préambule :

Dans ce chapitre, nous allons présenter nos tests réalisés dans le but d'évaluer la qualité de parole et de voir la différence entre deux bandes de fréquence. Pour ce faire, des données ont été collectées en posant un ensemble de conditions et des traitements correspondants ont été effectués. Les données utilisées dans notre travail sont fournies par l'Institut für Kommunikationen Akoustik (IKA) de l'université de Borchum en Allemagne et le Laboratoire d'Informatique pour la Mécanique et les Sciences d'Ingénieur (LIMSI) de l'université d'Orsay en France. Initialement, ces données sont acquises sous conversations téléphoniques enregistrées avec une fréquence d'échantillonnage de 32Khz. Dans ce travail, nous avons exploité les notations données aux différents stimuli sur une échelle de 1 à 5 ou de 0 à 174 et vérifiées par des experts pour évaluer la qualité de la parole.

Notre chapitre présente dans un premier temps, les échelles utilisées pour l'évaluation de la qualité de la parole. Dans un second temps, nous allons appliquer des tests sur notre base de données et nous montrons les résultats obtenus.

IV.1.Echelles de quantification de la qualité de la parole

Trois échelles sont utilisées dans cette étude pour évaluer la qualité du signal de parole. Ce sont des échelles très largement utilisées (Raake, 2005 ; Moller, 2000) pour les méthodes subjectives. Ces échelles sont ACR type MOS, l'échelle R du modèle E, et DCR ou échelle nominal.

IV.1.1.ACR type MOS :

Pour obtenir une valeur globale sur la qualité, deux tests ont été utilisés : l'un utilisant toutes les conditions bande étroite, et un autre utilisant toutes les conditions, large-bande et bande-étroite. Pour ces deux tests quatre locuteurs ont été utilisés (2 locuteurs et 2 locutrices).

L'auditeur écoute une seule fois chaque stimulus, puis estime la quantité du stimulus en appuyant sur l'une des valeurs de l'échelle MOS (voir figure 4.1). Cette valeur est validée lorsque le sujet appuie sur « suivant », ce qui joue automatiquement le prochain stimulus.



Figure IV. 1. Interface graphique du test utilisant l'échelle MOS

IV.1.2.Echelle R du modèle E :

L'échelle R du modèle E est efficace pour un test large bande, l'échelle étant très large de 0 à 100 (voir figure 4.2), pour évaluer les conditions en fonction des références bandes étroites, plusieurs Codecs dont le I_e est connu, ont été placés sur cette échelle :

- La référence du modèle E, le G.711 ($I_e=0$), cette référence est positionnée à 100, maximum de l'échelle R pour la bande étroite, ainsi tout résultat supérieur à cette valeur correspond à une amélioration de la bande étroite.
- Le CoDec G.729 ($I_e=10$), positionné à 90 sur l'échelle.
- Le CoDec IS-54 ($I_e=20$), positionné à 80.
- Le tandem G.726 (24 kbits/s)*G.726 (24 Kbits/s) ($I_e=50$), positionné à 50.

Ce test est effectué à chaque fois, avec deux locuteurs différents, un homme et une femme, prononçant la même phrase. Ce choix a été fait pour éviter les différences entre les conditions, liées à un contenu phonétiquement différent entre différentes phrases.



Figure IV.2. Interface graphique du test utilisant l'échelle R du modèle E

L'auditeur écoute le son à évaluer grâce au bouton PLAY, et les références en appuyant sur le bouton correspondant. Chaque son peut être écouté plusieurs fois, ce qui permet une meilleure évaluation de chaque stimulus. L'échelle R est agrandie jusqu'à 174 pour éviter tout problème de saturation comme sur l'échelle MOS. Enfin, l'évaluation de qualité du stimulus est choisie en positionnant le curseur sur la valeur souhaitée entre 1 et 174.

IV.1.3.Echelle nominale :

Un test de comparaison entre les conditions, a été réalisé, il évalue le choix des sujets entre les conditions dégradées large bande et les conditions non dégradées bande-étroite.

Le sujet devait écouter les deux sons avec les boutons « play 1 » et « play 2 » (voir figure IV.3), puis évaluer le stimulus ayant la meilleure qualité en appuyant sur le bouton « 1 » ou « 2 ». Chaque stimulus peut être écouté plusieurs fois.



Figure IV.3. Interface graphique du test utilisant l'échelle de comparaison par paires

IV.2. Conditions choisies :

Les conditions représentant les différents systèmes à évaluer, doivent être choisis de manière à étudier les différentes étapes de la compréhension du message. Certaines conditions permettent une meilleure intelligibilité du message. D'autres améliorent la sensation naturelle. En outre l'utilisation avec différentes bandes passantes permet de vérifier le modèle de dégradation.

Plusieurs types de conditions ont été choisis :

- Conditions introduisant des dégradations dues aux codeurs.
- Des conditions sans utilisation de codeurs mais avec des bandes passantes différentes.

18 conditions au total ont été choisies, réparties en 9 conditions concernant la bande étroite, et 9 conditions pour la bande élargie.

Ce choix de 18 conditions comprend principalement :

- Des codeurs large-bande, dont un l'AMR-WB (noté : G.722.2), utilisé à différents débits suivant la recommandation P.833 de l'ITU. En effet , nous observons que la

qualité globale varie avec le débit. La dégradation augmente lorsque le débit diminue.

- Une référence large-bande PCM. Cette condition représente notre référence pour l'extension de l'échelle R du midél E.
- Des conditions de référence bande-étroite suivant la recommandation P.833 de l'ITU. Ces références utilisent un filtre représentant les caractéristiques électroacoustiques d'un téléphone fixe habituel, appelé filtre « Intermediate Reference System » (IRS_R ou IRS_S pour receive ou send), il permet de simuler la réponse fréquentielle des téléphones du commerce.
- Plusieurs bandes passantes entre 600-2000 Hz et 50-7000 Hz

le tableau suivant présente les 18 conditions choisies :

Numéro de condition	Nom du CoDec	Type de CoDec	débit	Bandes passantes	Impairment Factor Ie	Fréquence centrale Hz	ΔZ (Barks)
1	G.711	PCM	64	G.712, IRS	0.0	-	-
2	G.726	ADPCM	32	G.712, IRS	7.0	-	-
3	G.729	CS-ACELP	8	G.712, IRS	11.0	-	-
4	IS-54	VSELP	8	G.712, IRS	20.0	-	-
5	G.726	ADPCM	16	G.712, IRS	50.0	-	-
6	-	-	-	300-3400 Hz	-	1010	13.4
7	-	-	-	600-3400 Hz	-	1428	10.7
8	-	-	-	300-2000 Hz	-	775	10.2
9	-	-	-	600-2000 Hz	-	1095	7.5
10	G.722	ADPCM	64	Plate		-	-
11	G.722.2	ACELP	23.05	Plate		-	-
12	G.722.2	ACELP	12.65	Plate		-	-
13	G.722.2	ACELP	6.6	Plate		-	-
14	-	-	-	50-7000 Hz	-	592	20
15	-	-	-	200-7000 Hz	-	1183	18.5
16	-	-	-	600-7000 Hz	-	2049	14.9
17	-	-	-	100-5000 Hz	-	707	17.6
18	-	-	-	50-3400 Hz	-	412	15.8

Tab IV.1. Caractéristiques des conditions

Dans le tableau suivant, nous allons représenter la répartition des conditions dans les différents tests :

N ⁰	Type	Test MOS1	Test MOS 2	Test échelle R	Test par comparaison
1	G.711	X	X	X	X
2	G.726 (32kbits/s)	X	X	X	X
3	G.729A	X	X	X	X
4	IS-54	X	X	X	X
5	G.726 (24kbits/s)*2	X	X	X	-
6	300-3400 Hz	X	X	X	-
7	600-3600 Hz	X	X	X	X
8	300-2000 Hz	X	X	X	X
9	600-2000 Hz	X	X	X	-
10	G.722	-	X	X	X
11	G.722.2 (23.05 Kbits/s)	-	X	X	X
12	G.722.2 (12.65 Kbits/s)	-	X	X	X
13	G.722.2 (6.6 Kbits/s)	-	X	X	-
14	50-7000 Hz	-	X	X	X
15	200-7000 Hz	-	X	X	X
16	600-7000 Hz	-	X	X	X
17	100-5000 Hz	-	X	X	-
18	50-3400 Hz	-	X	X	X

Tab.IV.2.Répartition des conditions dans chaque test

IV.3. Application des tests :

IV.3.1.Test MOS 1 :

Le test MOS permet d'évaluer les conditions décrites en haut, et de les confronter aux résultats de la littérature (Krebbber, 1995 ; Raake, 2005). Ce test consiste à faire une comparaison entre les locuteurs, moyenne globale.

- **Comparaison entre locuteurs, moyenne globale :**

Le tableau suivant présente les notes attribuées pour chaque locuteur utilisant l'échelle MOS.

Codec	Locutrice 1 Femme 	Locutrice 2 Femme 	Locuteur 3 Homme 	Locuteur 4 Homme 
G.711	4.13	4.13	4.13	4
G.726 (32kbts/s)	2.8	3	2.6	3.3
G.729A	3.4	3.5	3.52.5	2.7
IS-54	2.5	2.7	2.4	2.8
G.726 (24kbts/s)*2	1.6	2	1.5	1.8
300-3400 Hz	3.7	3.5	3.5	3.8
600-3600 Hz	3.2	2.9	3	3.4
300-2000 Hz	1.6	2	2.3	2.3
600-2000 Hz	1.8	1.8	1.6	1.5

Tab IV.3. Notes attribuées aux locuteurs du test MOS

Les notes montrent que pour certaines conditions, comme la condition 1 (G.711), les sujets ont donné le même niveau de qualité pour les stimuli quel que soit le locuteur. Un calcul d'analyse de variance (ANOVA), entre les locuteurs, montre que sur la globalité du test, il n'existe pas de différence significative entre les résultats des quatre locuteurs ($P=0.131$). Pourtant pour d'autres conditions, comme la condition 8, les résultats sont assez différents suivant le locuteur. Un calcul d'ANOVA pour chaque condition montre qu'il existe une différence entre les quatre locuteurs pour la condition 2 (G.726 à 32 Kbits/s, $P=0.033$) et 8 (300-2000hz) ($P=0.04$).

Un autre calcul d'ANOVA entre les sujets, montre qu'il existe toujours des différences dans les réponses des sujets à part pour la condition 1 ($P=0.381$), et dans une moindre mesure des conditions 3 (G.729) ($P=0.053$) et 8 ($P=0.059$). Cela nous indique que les différences entre les locuteurs pour la condition 8 sont voulues par l'ensemble des sujets. Il est intéressant de remarquer que pour cette condition 8, sur les 4 locuteurs, les 2 ayant la moins bonne qualité globale sont les locutrices. Cela montre que la fréquence de coupure haute à 2000Hz touche principalement les locutrices.

On a utilisé ces notes pour calculer la moyenne et l'écart type, les résultats sont présentés dans le tableau suivant :

Condition	Moyenne MOS	Ecart type MOS
G.711	4.10	0.695
G.726 (32kbits/s)	2.93	0.657
G.729A	3.52	0.711
IS-54	2.60	0.635
G.726 (24kbits/s)*2	1.72	0.642
300-3400 Hz	3.63	0.848
600-3600 Hz	3.13	0.800
300-2000 Hz	2.05	0.757
600-2000 Hz	1.68	0.653
Total	2.82	1.086

Tab IV.4. Moyenne pour chaque condition avec écart type

Les résultats en valeurs MOS du tableau montrent que la condition G.711, référence de la téléphonie bande-étroite, est bien évaluée comme la condition ayant la meilleure qualité. En effet en l'absence de compression de CoDec, la qualité sonore est meilleure que pour les autres conditions. De plus, la différence entre la condition G.726 avec un seul ou deux codages montre que le codage successif du même CoDec induit une dégradation plus importante.

On remarque aussi que la différence entre la condition G.711 et la bande passante 300-3400 Hz est simplement due à l'utilisation du filtre IRS. Ce filtre qui permet une meilleure intelligibilité de la voix.

- **Valeurs sur l'échelle R, I_e :**

La première étape consiste à utiliser la formule (III.9) pour obtenir des valeurs en échelle R, puis d'utiliser [ITU-T rec. Appendice C,2001] pour obtenir les I_e . Les valeurs obtenues $I_{e,sub}$ (subject), sont calculées à partir de la valeur sur l'échelle R de la condition de référence G.711.

$$I_{e,sub} = R(G.711) - R(\text{autres conditions}) \quad (IV.1)$$

Il est nécessaire par la suite de tracer les points $I_{e,sub}$ en fonction des valeurs attendues $I_{e,exp}$ (expected) puis d'obtenir la droite passant le plus proche des points tracés, pour ensuite

transformer toutes nos valeurs (références, bandes passantes) en valeurs normalisées sur l'échelle R.

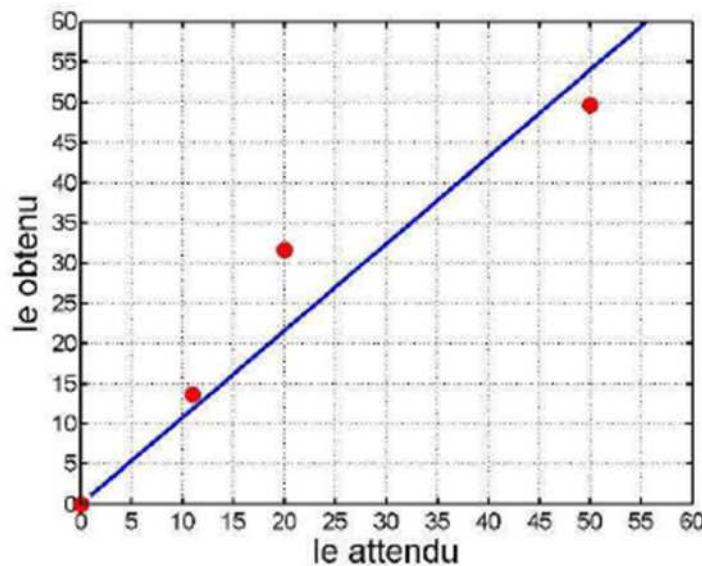


Figure IV.4. Courbe d'interpolation des Ie obtenus pour test MOS1

La courbe est calculée avec seulement quatre références, les valeurs de bruit de ligne l'ITU (-90 dB au lieu de -64 dB). Par conséquent, le CoDec G.726(24 bits/s)*G.726(24 bits/s) qui produit un bruit de quantification important a été évalué de moins bonne qualité que la valeur théorique. A part le niveau de bruit de ligne moins élevé, d'autres tests sur la qualité du CoDec G.726 avaient aussi montré une influence plus importante du CoDec (Moller 2000). De plus, une interface large bande a été utilisée, ce qui a augmenté les dégradations.

Le calcul des nouvelles valeurs R utilise le coefficient de la courbe obtenue ($a = 1.0836$) pour trouver les nouvelles valeurs de Ie

$$I_{e,sub} = a \cdot I_{e,exp}$$

Puis calculer les nouvelles valeurs normalisées sur l'échelle R en fonction de la référence G.711 :

$$R-833 = 93.2 - I_{e-833}$$

Les valeurs normalisées sur l'échelle R sont données dans le tableau suivant :

Condition	Moyenne R	Ie,sub R	Ie,exp R	Ie-833 R	R-833
G.711	82.1404	0	0	0	93.2000
G.726 (32 kbits/s)	56.7768	25.3636	7	23.4075	69.7925
G.729	68.4341	13.7063	10	12.6493	80.5507
IS-54	50.5191	31.6213	20	29.1826	64.0174
2*G.726(32 kbits/s)	32.5162	49.6242	50	45.9771	47.4029
300-3400 Hz	70.5997	11.5407	-	10.6506	82.2131
600-3400 Hz	60.4833	21.6571	-	19.9868	73.2131
300-2000 Hz	39.6207	42.5197	-	39.2405	53.9595
600-2000 Hz	31.7286	50.4118	-	46.5240	46.6760

Tab IV.5. Calcul des valeurs normalisées sur l'échelle R des conditions du test MOS1

Le tableau montre que les références sont évaluées comme ayant une plus grande dégradation que les valeurs théorique.

IV.3.2.Test MOS 2 :

Ce test alliant condition bande-étroite et large-bande va permettre de comparer les références dans ces deux tests et ainsi de trouver une première valeur sur l'échelle R d'une condition large-bande . Comme pour le test MOS , une comparaison entre les locuteurs est d'abord effectuée pour connaître l'influence du sexe du locuteur sur les différentes conditions. Par la suite une valeur globale pour l'ensemble des locuteurs est calculé.

Les notes données par chaque locuteurs sont données dans le tableau suivant :

Codec	Locutrice 1 Femme 	Locutrice 2 Femme 	Locuteur 3 Homme 	Locuteur 4 Homme 
G.711	2.8	3.4	3.3	3.2
G.726 (32kbits/s)	1.9	2.1	2	2.5
G.729A	2.7	2.7	2.6	2.8
IS-54	1.9	1.9	1.8	2.1
G.726 (24kbits/s)*2	1.4	1.7	1.3	1.6
300-3400 Hz	3.2	3.1	2.8	3
600-3600 Hz	2.5	3	2.5	3
300-2000 Hz	1.5	2.2	2	1.9
600-2000 Hz	1.6	1.6	1.9	1.4
G.722	2.6	2.9	2.7	2.7
G.722.2 (23.05 Kbits/s)	4.2	4	4	4.4
G.722.2 (12.65 Kbits/s)	3.7	3.8	4.1	4
G.722.2 (6.6 Kbits/s)	2.8	2.6	3.2	3.2
50-7000 Hz	4.2	4.4	4.3	4.1
200-7000 Hz	3.9	4.5	4.2	4.1
600-7000 Hz	3.2	3	3.4	3.2
100-5000 Hz	3.5	3.9	3.8	3.8
50-3400 Hz	2.8	2.3	2.9	2.9

Tab IV.6. Notes attribuées aux locuteurs du test MOS2

Ici les mêmes stimuli bande-étroite sont utilisés. Or ceux-ci sont bien évalués comme ayant une moins bonne qualité par rapport au test précédent. De plus, on remarque qu'il existe une plus grande différence entre les locuteurs sur les conditions de référence. Un calcul d'ANOVA montre que les conditions large-bande ont peu de différence entre les locuteurs. Par exemple un calcul pour la condition 11 (G.722.2 à 23.05 Kbits/s), montre que la moyenne est identique quelque soit le locuteur ($F=0.37$; $P=0.7768$), et qu'il ya peu de différence entre les réponses des sujets ($F=1.34$; $P=0.1818$). de plus on remarque que pour la condition 13 (G.722.2 à 6.06 Kbits/s), il existe des différences significatives entre les locuteurs ($F=4.49$; $P=0.063$), mais celle-ci est très défavorable aux locutrices.

Un calcul sur l'ensemble des conditions montre qu'il n'existe pas de différence significative entre les locuteurs ($F=1.623$; $P=0.182$).

le tableau montre la moyenne pour chaque condition en valeur MOS ainsi que leurs valeur R

Conditions	Moyenne MOS	Ecart type	Moyenne R	Moyenne MOS1
G.711	3.18	0.876	61.6446	4.10
G.726 (32kbits/s)	2.13	0.744	41.3458	2.93
G.729A	2.75	0.673	53.3235	3.52
IS-54	1.96	0.610	37.7768	2.60
G.726 (24kbits/s)*2	1.50	0.584	27.2688	1.72
300-3400 Hz	3.02	0.825	58.4948	3.63
600-3600 Hz	2.71	0.704	52.4989	3.13
300-2000 Hz	1.89	0.703	36.3646	2.05
600-2000 Hz	1.62	0.608	30.2550	1.68
G.722	2.74	0.797	53.1174	-
G.722.2 (23.05 Kbits/s)	4.16	0.730	83.9168	-
G.722.2 (12.65 Kbits/s)	3.91	0.608	77.1784	-
G.722.2 (6.6 Kbits/s)	2.96	0.847	57.2478	-
50-7000 Hz	.426	0.876	87.0731	-
200-7000 Hz	4.17	0.797	84.2474	-
600-7000 Hz	3.22	0.765	62.2819	-
100-5000 Hz	3.73	0.782	72.8634	-
50-3400 Hz	2.72	0.83	52.7051	-
Total	2.92	0.816	57.2019	2.82

Tab IV.7. Résultats du test MOS2

La figure suivante représente la moyenne pour chaque condition avec écart type :

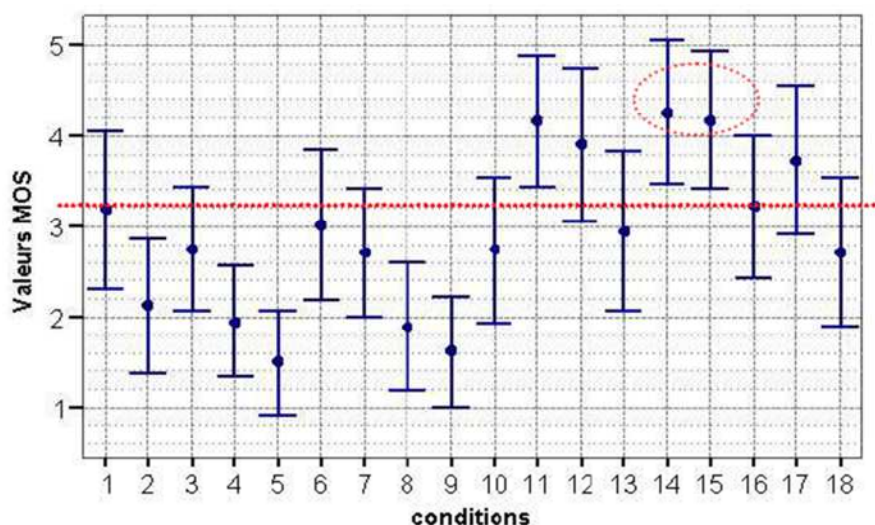


Figure IV.6. Moyenne pour chaque condition avec écart type

Une des premières informations données par la figure est la préférence de presque toutes les conditions large-bande sauf la 10, 13 et 18. Les valeurs de ces dernières montrent qu'une dégradation trop importante du signal liée à un CoDec est plus importante dans l'évaluation globale du stimulus que l'augmentation de la bande passante. Les trois conditions du CoDec AMR-WB montrent effectivement que la qualité de la transmission diminue quand le débit diminue. Enfin, nous observons que l'augmentation de la bande passante dans les fréquences basses sans changer la fréquence de coupure haute, induit une qualité inférieure du stimulus que la référence bande-étroite G.711. Ces fréquences peuvent donc jouer un rôle négatif dans la compréhension du message malgré une sensation de la voix plus proche de la réalité. Il faut aussi remarquer qu'une légère différence, non significative ($F=0.571$; $P=0.451$) de 0.1 point MOS, existe entre les deux conditions 50-7000 Hz et 200-7000 Hz, la première étant préférée sur la deuxième.

Un calcul d'amélioration entre certaines conditions large-bande et la condition de référence G.711, nous donne :

- Une amélioration de 1.08 points MOS pour la condition 50-7000 Hz (soit 34%).
- Une amélioration de 0.99 points MOS pour la condition 200-7000 Hz (soit 31%).

IV.3.3. Comparaison entre MOS1 et MOS2

le tableau suivant montre la répartition des résultats sur l'échelle MOS pour le premier et second test.

Valeur MOS	1	2	3	4	5	Total
% test MOS1	11.9	28..3	31.9	22.1	5.8	100.0
% test MOS2	10.3	27.9	31.5	19.9	10.4	100.0

Tab IV.8. Comparaison des résultats entre MOS1 et MOS2

On remarque que la répartition est différente entre les deux tests, on obtient le double pour la valeur 5. Ce la montre une augmentation globale de la qualité sur l'ensemble des stimuli

IV.3.4. Test avec l'échelle R :

Ce test utilise les 18 conditions bande-étroite et bande élargie. L'interface de jugement correspond à une échelle R du modèle agrandie au dessus de 100 (100 correspond au maximum de qualité pour la téléphonie bande-étroite) et sur laquelle quatre boutons permettant d'écouter le stimulus utilisant les quatre CoDec de référence, en effet la même conversation était jouée quelque soit la condition à évalué par le sujet. Ceci permettait au sujet de mémoriser la qualité des références afin de les comparer plus facilement aux stimuli. Les quatre références sont :

- La référence bande-étroite, le CoDec G.711, positionné à 100 sur l'échelle R.
- Le CoDec G.729, positionné à 90.
- Le CoDec IS-54, positionné à 80.
- Le tandem de CoDecs G.726 à 24 Kbits/s, positionné à 50.

Ce test a été effectué dans les deux pays France et Allemagne avec à chaque fois deux locuteurs, un masculin et un féminin. Il est important de préciser également que tous les stimuli correspondaient à la même conversation processé avec des conditions différentes.

La figure suivante montre les notes attribuées pour chaque locuteur

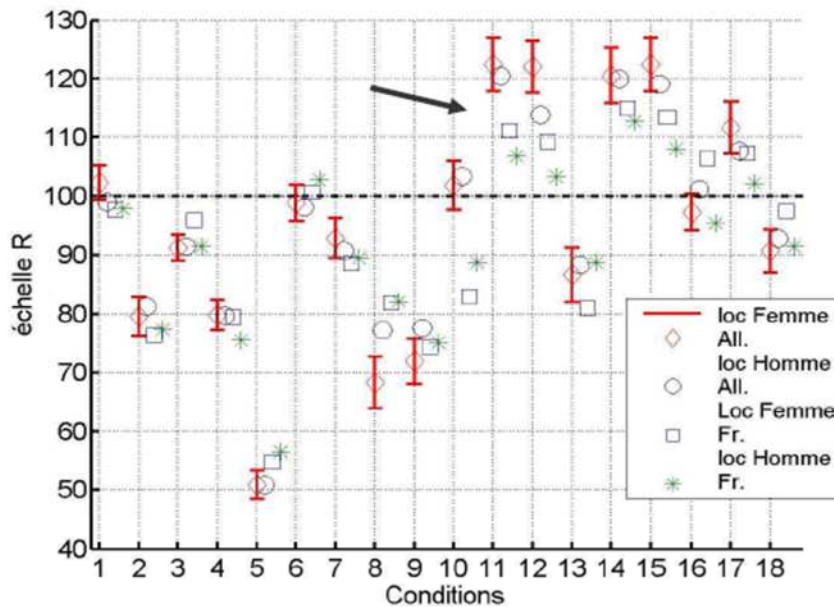


Figure IV.6. Résultats MOS2 pour chaque condition

On remarque des variations plus importantes entre les locuteurs pour chaque condition. Pour certaines conditions, il existe deux groupes distincts entre les deux locuteurs allemands et les deux Français.

Pour les conditions 10 à 18 (celles en bande élargie), on observe qu'une proportion plus importante de points se situe au dessus de la barre 100. Cette ligne représente la qualité théorique de la référence bande-étroite, le G.726.

- Comparaison entre locuteurs

-Test allemand

Les résultats obtenus avec les locuteurs allemands sont donnés dans le tableau qui suit :

Condition	Valeur MOS femme 	Ecart type	Valeur MOS homme 	Ecart type
G.711	102.29	8.615	99.05	5.084
G.726 (32kbts/s)	76.62	10.925	81.33	8.639
G.729	91.24	4.816	91.52	3.544
IS-54	79.86	6.560	79.76	1.513
G.726 (24kbts/s)*2	50.86	5.833	50.90	7.739
300-3400 Hz	98.86	9.345	98.14	11.573
600-3600 Hz	92.90	11.322	90.81	6.161
300-2000 Hz	68.29	19.499	77.29	14.217
600-2000 Hz	71.95	14.746	77.67	14.434
G.722	101.81	17.505	103.38	17.834
G.722.2 (23.05 Kbits/s)	122.43	20.289	120.48	18.597
G.722.2 (12.65 Kbits/s)	122.10	19.131	113.90	17.216
G.722.2 (6.6 Kbits/s)	86.67	21.355	88.48	19.884
50-7000 Hz	120.52	22.2	120.00	21.881
200-7000 Hz	122.48	20.868	119.19	20.532
600-7000 Hz	97.29	9.514	101.19	13.512
100-5000 Hz	111.67	19.417	107.67	12.959
50-3400 Hz	90.76	13.123	92.90	15.016
Total	95.20	24.832	95.09	22.370

Tab IV.9. Résultats du test avec l'échelle R pour les locuteurs allemands

Un calcul d'ANOVA entre les locuteurs, sur l'ensemble des conditions montre qu'il n'existe pas de différence significative entre les locuteurs ($F=0.005$; $P=0.946$). Le tableau montre les résultats pour les deux locuteurs. Un second calcul d'ANOVA pour chaque condition montre qu'il n'y a pas de différence entre les deux locuteurs quelle que soit la condition. Pourtant, la condition G.722.2 à 12.65 Kbits/s ($F=2.921$; $P=0.095$) semble plutôt désavantageux pour le locuteur homme. De même, la condition 300-2000 Hz, montre une différence entre les deux locuteurs ($F=2.127$; $P=0.153$), la fréquence de coupure haute à 2000 Hz semble toucher principalement la locutrice. De plus, la condition G.726 à 32 Kbits/s, semble également être évaluée de moins bonne qualité que sa valeur théorique (qui est normalement 93).

-Test français

Les résultats obtenus avec les locuteurs français sont données dans le tableau qui suit :

Condition	Valeur MOS femme □	Ecart type	Valeur MOS homme *	Ecart type
G.711	97.93	4.713	97.73	7.005
G.726 (32kbits/s)	77.33	11.667	76.4	13.200
G.729	91.47	9.270	95.93	5.489
IS-54	75.60	7.615	97.47	11.969
G.726 (24kbits/s)*2	56.47	10.557	54.80	12.194
300-3400 Hz	102.80	10.913	100.67	11.839
600-3600 Hz	89.53	13.285	88.73	8.271
300-2000 Hz	82.13	15.281	81.93	19.324
600-2000 Hz	75.13	16.561	74.4	13.092
G.722	88.8	17.714	83.07	13.203
G.722.2 (23.05 Kbits/s)	106.8	13.246	111.20	13.311
G.722.2 (12.65 Kbits/s)	103.27	9.225	109.33	10.032
G.722.2 (6.6 Kbits/s)	88.8	14.646	81.07	16.372
50-7000 Hz	112.80	13.177	114.93	16.032
200-7000 Hz	108.13	11.432	113.43	9.234
600-7000 Hz	95.47	21.189	106.53	11.103
100-5000 Hz	102.13	11.012	107.4	15.108
50-3400 Hz	91.47	14.667	97.53	13.851
Total	91.45	20.496	93.03	18.544

Tab IV.10. Résultats du test avec l'échelle R pour les locuteurs français

Pour les résultats de ce test, comme ceux du test précédent, les calculs d'ANOVA ont montré qu'il n'y a pas de différence significative entre les locuteurs sur l'ensemble des conditions ($F=0.888$; $P=0.346$) ni pour une condition particulière. Cependant certaines conditions montrent des différences entre les locuteurs. La condition G.722.2 à 6.6 Kbits/s ($F=1.859$; $P=0.184$), se révèle ici désavantageuse pour la locutrice. La bande passante 600-7000 Hz, montre également une différence entre les deux locuteurs ($F=3.210$; $P=0.084$). En effet, une fréquence de coupure basse semble très désavantageuse pour le locuteur. La condition 50-3400 Hz, montre à l'inverse un désavantage pour le locuteur en raison de l'accentuation trop prononcée des fréquences graves ($F=1.859$; $P=0.184$).

- **Comparaison entre conditions :**

Conditions	Moyenne Allemagne	Ecart type	Moyenne France	Ecart type
G.711	100.67	7.176	97.83	5.867
G.726 (32kbits/s)	80.48	9.766	76.87	12.250
G.729	91.38	4.179	93.7	7.822
IS-54	79.81	4.702	77.53	10.051
G.726 (24kbits/s)*2	50.88	6.769	55.63	11.239
300-3400 Hz	98.50	10.395	101.73	11.240
600-3600 Hz	91.86	9.065	89.13	10.881
300-2000 Hz	72.79	17.459	82.03	17.117
600-2000 Hz	74.81	14.699	74.77	14.673
G.722	102.60	17.473	85.93	15.625
G.722.2 (23.05 Kbits/s)	121.45	19.248	109.00	13.238
G.722.2 (12.65 Kbits/s)	118.00	18.447	106.30	9.959
G.722.2 (6.6 Kbits/s)	87.57	20.400	84.93	15.761
50-7000 Hz	120.26	21.773	113.87	14.460
200-7000 Hz	120.83	20.514	110.80	10.565
600-7000 Hz	99.24	11.710	101.00	17.548
100-5000 Hz	109.67	16.430	104.77	13.263
50-3400 Hz	91.83	13.972	94.50	14.352
Total	95.15	23.618	92.24	19.542

Tab IV.11. Résultats de la comparaison entre conditions

Une comparaison entre les conditions permet d'observer que la qualité des stimuli, pour les conditions G.722.2 baisse avec le débit. Puis un calcul d'ANOVA entre les conditions en comparant chaque condition à la référence G.711 montre que :

- Pour le test effectué en Allemagne, il existe une différence significative entre la référence bande-étroite et toutes les conditions large-bande sauf pour les conditions 300-3400 Hz, G.722 et 600-7000 Hz.
- De même, pour le test en France, il existe une différence significative sauf pour les conditions G.729, 300-3400 Hz, 600-7000 Hz, 100-5000 Hz et 50-3400 Hz.

Nous pouvons en conclure que la fréquence de coupure basse à 50 Hz, est considérée comme une dégradation pour les sujets allemands mais pas pour les Français. Aussi une augmentation de 2 Barks de la bande passante pour cette condition et une augmentation de 1.5

Barks pour la condition 600-7000 Hz ne semblent pas être suffisante pour améliorer la qualité de transmission.

- Les conditions large-bande semblent avoir une meilleure qualité que la référence G.711 sauf la condition G.722.2 à 6.6 Kbits/s. Ceci montre qu'une dégradation trop importante du signal liée à un CoDec ne permet pas une amélioration grâce à la bande élargie.

Discussion :

Ce chapitre a permis de mettre en évidence les différentes techniques utilisées pour évaluer la qualité de la parole. Ces techniques ont été appliquées suivant plusieurs conditions d'acquisition du signal de parole. Selon les résultats d'évaluation, la meilleure qualité est obtenue en utilisant la condition large bande.

Préambule :

Dans ce chapitre, nous allons présenter nos tests réalisés dans le but d'évaluer la qualité de parole et de voir la différence entre deux bandes de fréquence. Pour ce faire, des données ont été collectées en posant un ensemble de conditions et des traitements correspondants ont été effectués. Les données utilisées dans notre travail sont fournies par l'Institut für Kommunikations Akoustik (IKA) de l'université de Borchum en Allemagne et le Laboratoire d'Informatique pour la Mécanique et les Sciences d'Ingénieur (LIMSI) de l'université d'Orsay en France. Initialement, ces données sont acquises sous conversations téléphonique enregistrées avec une fréquence d'échantillonnage de 32Khz. Dans ce travail, nous avons exploité les notations données aux différents stimuli sur une échelle de 1 à 5 ou de 0 à 174 et vérifiées par des experts pour évaluer la qualité de la parole.

Notre chapitre présente dans un premier temps, les échelles utilisées pour l'évaluation de la qualité de la parole. Dans un second temps, nous allons appliquer des tests sur notre base de données et nous montrons les résultats obtenus.

IV.1.Echelles de quantification de la qualité de la parole

Trois échelles sont utilisées dans cette étude pour évaluer la qualité du signal de parole. Ce sont des échelles très largement utilisées (Raake, 2005 ; Moller, 2000) pour les méthodes subjectives. Ces échelles sont ACR type MOS, l'échelle R du modèle E, et DCR ou échelle nominal.

IV.1.1.ACR type MOS :

Pour obtenir une valeur globale sur la qualité, deux tests ont été utilisés : l'un utilisant toutes les conditions bande étroite, et un autre utilisant toutes les conditions, large-bande et bande-étroite. Pour ces deux tests quatre locuteurs ont été utilisés (2 locuteurs et 2 locutrices).

L'auditeur écoute une seule fois chaque stimulus, puis estime la quantité du stimulus en appuyant sur l'une des valeurs de l'échelle MOS (voir figure 4.1). Cette valeur est validée lorsque le sujet appuie sur « suivant », ce qui joue automatiquement le prochain stimulus.



Figure IV. 1. Interface graphique du test utilisant l'échelle MOS

IV.1.2.Echelle R du modèle E :

L'échelle R du modèle E est efficace pour un test large bande, l'échelle étant très large de 0 à 100 (voir figure 4.2), pour évaluer les conditions en fonction des références bandes étroites, plusieurs Codecs dont le I_e est connu, ont été placés sur cette échelle :

- La référence du modèle E, le G.711 ($I_e=0$), cette référence est positionnée à 100, maximum de l'échelle R pour la bande étroite, ainsi tout résultat supérieur à cette valeur correspond à une amélioration de la bande étroite.
- Le CoDec G.729 ($I_e=10$), positionné à 90 sur l'échelle.
- Le CoDec IS-54 ($I_e=20$), positionné à 80.
- Le tandem G.726 (24 kbits/s)*G.726 (24 Kbits/s) ($I_e=50$), positionné à 50.

Ce test est effectué à chaque fois, avec deux locuteurs différents, un homme et une femme, prononçant la même phrase. Ce choix a été fait pour éviter les différences entre les conditions, liées à un contenu phonétiquement différent entre différentes phrases.



Figure IV.2. Interface graphique du test utilisant l'échelle R du modèle E

L'auditeur écoute le son à évaluer grâce au bouton PLAY, et les références en appuyant sur le bouton correspondant. Chaque son peut être écouté plusieurs fois, ce qui permet une meilleure évaluation de chaque stimulus. L'échelle R est agrandie jusqu'à 174 pour éviter tout problème de saturation comme sur l'échelle MOS. Enfin, l'évaluation de qualité du stimulus est choisie en positionnant le curseur sur la valeur souhaitée entre 1 et 174.

IV.1.3.Echelle nominale :

Un test de comparaison entre les conditions, a été réalisé, il évalue le choix des sujets entre les conditions dégradées large bande et les conditions non dégradées bande-étroite.

Le sujet devait écouter les deux sons avec les boutons « play 1 » et « play 2 » (voir figure IV.3), puis évaluer le stimulus ayant la meilleure qualité en appuyant sur le bouton « 1 » ou « 2 ». Chaque stimulus peut être écouté plusieurs fois.



Figure IV.3. Interface graphique du test utilisant l'échelle de comparaison par paires

IV.2. Conditions choisies :

Les conditions représentant les différents systèmes à évaluer, doivent être choisis de manière à étudier les différentes étapes de la compréhension du message. Certaines conditions permettent une meilleure intelligibilité du message. D'autres améliorent la sensation naturelle. En outre l'utilisation avec différentes bandes passantes permet de vérifier le modèle de dégradation.

Plusieurs types de conditions ont été choisis :

- Conditions introduisant des dégradations dues aux codeurs.
- Des conditions sans utilisation de codeurs mais avec des bandes passantes différentes.

18 conditions au total ont été choisies, répartis en 9 conditions concernant la bande étroite, et 9 conditions pour la bande élargie.

Ce choix de 18 conditions comprend principalement :

- Des codeurs large-bande, dont un l'AMR-WB (noté : G.722.2), utilisé à différents débits suivant la recommandation P.833 de l'ITU. En effet , nous observons que la

qualité globale varie avec le débit. La dégradation augmente lorsque le débit diminue.

- Une référence large-bande PCM. Cette condition représente notre référence pour l'extension de l'échelle R du midél E.
- Des conditions de référence bande-étroite suivant la recommandation P.833 de l'ITU. Ces références utilisent un filtre représentant les caractéristiques électroacoustiques d'un téléphone fixe habituel, appelé filtre « Intermediate Reference System » (IRS_R ou IRS_S pour receive ou send), il permet de simuler la réponse fréquentielle des téléphones du commerce.
- Plusieurs bandes passantes entre 600-2000 Hz et 50-7000 Hz

le tableau suivant présente les 18 conditions choisies :

Numéro de condition	Nom du CoDec	Type de CoDec	débit	Bandes passantes	Impairment Factor Ie	Fréquence centrale Hz	ΔZ (Barks)
1	G.711	PCM	64	G.712, IRS	0.0	-	-
2	G.726	ADPCM	32	G.712, IRS	7.0	-	-
3	G.729	CS-ACELP	8	G.712, IRS	11.0	-	-
4	IS-54	VSELP	8	G.712, IRS	20.0	-	-
5	G.726	ADPCM	16	G.712, IRS	50.0	-	-
6	-	-	-	300-3400 Hz	-	1010	13.4
7	-	-	-	600-3400 Hz	-	1428	10.7
8	-	-	-	300-2000 Hz	-	775	10.2
9	-	-	-	600-2000 Hz	-	1095	7.5
10	G.722	ADPCM	64	Plate		-	-
11	G.722.2	ACELP	23.05	Plate		-	-
12	G.722.2	ACELP	12.65	Plate		-	-
13	G.722.2	ACELP	6.6	Plate		-	-
14	-	-	-	50-7000 Hz	-	592	20
15	-	-	-	200-7000 Hz	-	1183	18.5
16	-	-	-	600-7000 Hz	-	2049	14.9
17	-	-	-	100-5000 Hz	-	707	17.6
18	-	-	-	50-3400 Hz	-	412	15.8

Tab IV.1. Caractéristiques des conditions

Dans le tableau suivant, nous allons représenter la répartition des conditions dans les différents tests :

N ⁰	Type	Test MOS1	Test MOS 2	Test échelle R	Test par comparaison
1	G.711	X	X	X	X
2	G.726 (32kbits/s)	X	X	X	X
3	G.729A	X	X	X	X
4	IS-54	X	X	X	X
5	G.726 (24kbits/s)*2	X	X	X	-
6	300-3400 Hz	X	X	X	-
7	600-3600 Hz	X	X	X	X
8	300-2000 Hz	X	X	X	X
9	600-2000 Hz	X	X	X	-
10	G.722	-	X	X	X
11	G.722.2 (23.05 Kbits/s)	-	X	X	X
12	G.722.2 (12.65 Kbits/s)	-	X	X	X
13	G.722.2 (6.6 Kbits/s)	-	X	X	-
14	50-7000 Hz	-	X	X	X
15	200-7000 Hz	-	X	X	X
16	600-7000 Hz	-	X	X	X
17	100-5000 Hz	-	X	X	-
18	50-3400 Hz	-	X	X	X

Tab.IV.2.Répartition des conditions dans chaque test

IV.3. Application des tests :

IV.3.1.Test MOS 1 :

Le test MOS permet d'évaluer les conditions décrites en haut, et de les confronter aux résultats de la littérature (Krebbber, 1995 ; Raake, 2005). Ce test consiste à faire une comparaison entre les locuteurs, moyenne globale.

- **Comparaison entre locuteurs, moyenne globale :**

Le tableau suivant présente les notes attribuées pour chaque locuteur utilisant l'échelle MOS.

Codec	Locutrice 1 Femme 	Locutrice 2 Femme 	Locuteur 3 Homme 	Locuteur 4 Homme 
G.711	4.13	4.13	4.13	4
G.726 (32kbits/s)	2.8	3	2.6	3.3
G.729A	3.4	3.5	3.52.5	2.7
IS-54	2.5	2.7	2.4	2.8
G.726 (24kbits/s)*2	1.6	2	1.5	1.8
300-3400 Hz	3.7	3.5	3.5	3.8
600-3600 Hz	3.2	2.9	3	3.4
300-2000 Hz	1.6	2	2.3	2.3
600-2000 Hz	1.8	1.8	1.6	1.5

Tab IV.3. Notes attribuées aux locuteurs du test MOS

Les notes montrent que pour certaines conditions, comme la condition 1 (G.711), les sujets ont donné le même niveau de qualité pour les stimuli quel que soit le locuteur. Un calcul d'analyse de variance (ANOVA), entre les locuteurs, montre que sur la globalité du test, il n'existe pas de différence significative entre les résultats des quatre locuteurs ($P=0.131$). Pourtant pour d'autres conditions, comme la condition 8, les résultats sont assez différents suivant le locuteur. Un calcul d'ANOVA pour chaque condition montre qu'il existe une différence entre les quatre locuteurs pour la condition 2 (G.726 à 32 Kbits/s, $P=0.033$) et 8 (300-2000hz) ($P=0.04$).

Un autre calcul d'ANOVA entre les sujets, montre qu'il existe toujours des différences dans les réponses des sujets à part pour la condition 1 ($P=0.381$), et dans une moindre mesure des conditions 3 (G.729) ($P=0.053$) et 8 ($P=0.059$). Cela nous indique que les différences entre les locuteurs pour la condition 8 sont voulues par l'ensemble des sujets. Il est intéressant de remarquer que pour cette condition 8, sur les 4 locuteurs, les 2 ayant la moins bonne qualité globale sont les locutrices. Cela montre que la fréquence de coupure haute à 2000Hz touche principalement les locutrices.

On a utilisé ces notes pour calculer la moyenne et l'écart type, les résultats sont présentés dans le tableau suivant :

Condition	Moyenne MOS	Ecart type MOS
G.711	4.10	0.695
G.726 (32kbits/s)	2.93	0.657
G.729A	3.52	0.711
IS-54	2.60	0.635
G.726 (24kbits/s)*2	1.72	0.642
300-3400 Hz	3.63	0.848
600-3600 Hz	3.13	0.800
300-2000 Hz	2.05	0.757
600-2000 Hz	1.68	0.653
Total	2.82	1.086

Tab IV.4. Moyenne pour chaque condition avec écart type

Les résultats en valeurs MOS du tableau montrent que la condition G.711, référence de la téléphonie bande-étroite, est bien évaluée comme la condition ayant la meilleure qualité. En effet en l'absence de compression de CoDec, la qualité sonore est meilleure que pour les autres conditions. De plus, la différence entre la condition G.726 avec un seul ou deux codages montre que le codage successif du même CoDec induit une dégradation plus importante.

On remarque aussi que la différence entre la condition G.711 et la bande passante 300-3400 Hz est simplement due à l'utilisation du filtre IRS. Ce filtre qui permet une meilleure intelligibilité de la voix.

- **Valeurs sur l'échelle R, I_e :**

La première étape consiste à utiliser la formule (III.9) pour obtenir des valeurs en échelle R, puis d'utiliser [ITU-T rec. Appendice C,2001] pour obtenir les I_e . Les valeurs obtenues $I_{e,sub}$ (subject), sont calculées à partir de la valeur sur l'échelle R de la condition de référence G.711.

$$I_{e,sub} = R(G.711) - R(\text{autres conditions}) \quad (IV.1)$$

Il est nécessaire par la suite de tracer les points $I_{e,sub}$ en fonction des valeurs attendues $I_{e,exp}$ (expected) puis d'obtenir la droite passant le plus proche des points tracés, pour ensuite

transformer toutes nos valeurs (références, bandes passantes) en valeurs normalisées sur l'échelle R.

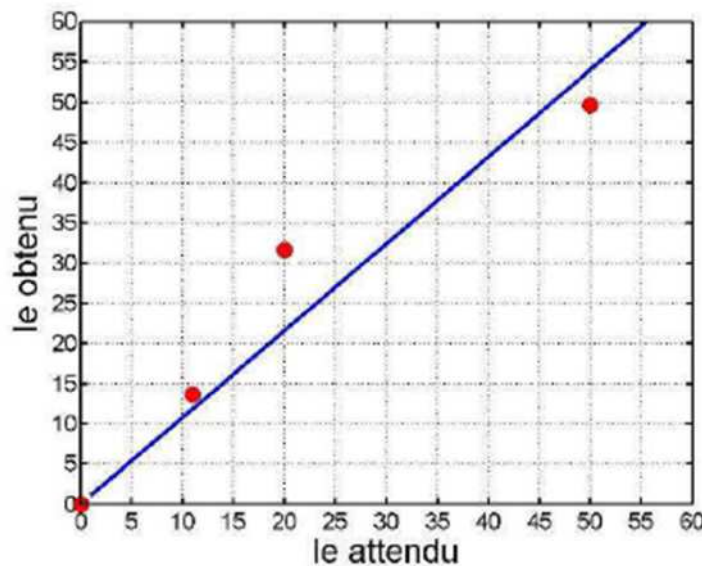


Figure IV.4. Courbe d'interpolation des Ie obtenus pour test MOS1

La courbe est calculée avec seulement quatre références, les valeurs de bruit de ligne l'ITU (-90 dB au lieu de -64 dB). Par conséquent, le CoDec G.726(24 bits/s)*G.726(24 bits/s) qui produit un bruit de quantification important a été évalué de moins bonne qualité que la valeur théorique. A part le niveau de bruit de ligne moins élevé, d'autres tests sur la qualité du CoDec G.726 avaient aussi montré une influence plus importante du CoDec (Moller 2000). De plus, une interface large bande a été utilisée, ce qui a augmenté les dégradations.

Le calcul des nouvelles valeurs R utilise le coefficient de la courbe obtenue ($a = 1.0836$) pour trouver les nouvelles valeurs de Ie

$$I_{e,sub} = a \cdot I_{e,exp}$$

Puis calculer les nouvelles valeurs normalisées sur l'échelle R en fonction de la référence G.711 :

$$R-833 = 93.2 - I_{e-833}$$

Les valeurs normalisées sur l'échelle R sont données dans le tableau suivant :

Condition	Moyenne R	Ie,sub R	Ie,exp R	Ie-833 R	R-833
G.711	82.1404	0	0	0	93.2000
G.726 (32 kbits/s)	56.7768	25.3636	7	23.4075	69.7925
G.729	68.4341	13.7063	10	12.6493	80.5507
IS-54	50.5191	31.6213	20	29.1826	64.0174
2*G.726(32 kbits/s)	32.5162	49.6242	50	45.9771	47.4029
300-3400 Hz	70.5997	11.5407	-	10.6506	82.2131
600-3400 Hz	60.4833	21.6571	-	19.9868	73.2131
300-2000 Hz	39.6207	42.5197	-	39.2405	53.9595
600-2000 Hz	31.7286	50.4118	-	46.5240	46.6760

Tab IV.5. Calcul des valeurs normalisées sur l'échelle R des conditions du test MOS1

Le tableau montre que les références sont évaluées comme ayant une plus grande dégradation que les valeurs théorique.

IV.3.2.Test MOS 2 :

Ce test alliant condition bande-étroite et large-bande va permettre de comparer les références dans ces deux tests et ainsi de trouver une première valeur sur l'échelle R d'une condition large-bande . Comme pour le test MOS , une comparaison entre les locuteurs est d'abord effectuée pour connaître l'influence du sexe du locuteur sur les différentes conditions. Par la suite une valeur globale pour l'ensemble des locuteurs est calculé.

Les notes données par chaque locuteurs sont données dans le tableau suivant :

Codec	Locutrice 1 Femme 	Locutrice 2 Femme 	Locuteur 3 Homme 	Locuteur 4 Homme 
G.711	2.8	3.4	3.3	3.2
G.726 (32kbits/s)	1.9	2.1	2	2.5
G.729A	2.7	2.7	2.6	2.8
IS-54	1.9	1.9	1.8	2.1
G.726 (24kbits/s)*2	1.4	1.7	1.3	1.6
300-3400 Hz	3.2	3.1	2.8	3
600-3600 Hz	2.5	3	2.5	3
300-2000 Hz	1.5	2.2	2	1.9
600-2000 Hz	1.6	1.6	1.9	1.4
G.722	2.6	2.9	2.7	2.7
G.722.2 (23.05 Kbits/s)	4.2	4	4	4.4
G.722.2 (12.65 Kbits/s)	3.7	3.8	4.1	4
G.722.2 (6.6 Kbits/s)	2.8	2.6	3.2	3.2
50-7000 Hz	4.2	4.4	4.3	4.1
200-7000 Hz	3.9	4.5	4.2	4.1
600-7000 Hz	3.2	3	3.4	3.2
100-5000 Hz	3.5	3.9	3.8	3.8
50-3400 Hz	2.8	2.3	2.9	2.9

Tab IV.6. Notes attribuées aux locuteurs du test MOS2

Ici les mêmes stimuli bande-étroite sont utilisés. Or ceux-ci sont bien évalués comme ayant une moins bonne qualité par rapport au test précédent. De plus, on remarque qu'il existe une plus grande différence entre les locuteurs sur les conditions de référence. Un calcul d'ANOVA montre que les conditions large-bande ont peu de différence entre les locuteurs. Par exemple un calcul pour la condition 11 (G.722.2 à 23.05 Kbits/s), montre que la moyenne est identique quelque soit le locuteur ($F=0.37$; $P=0.7768$), et qu'il ya peu de différence entre les réponses des sujets ($F=1.34$; $P=0.1818$). de plus on remarque que pour la condition 13 (G.722.2 à 6.06 Kbits/s), il existe des différences significatives entre les locuteurs ($F=4.49$; $P=0.063$), mais celle-ci est très défavorable aux locutrices.

Un calcul sur l'ensemble des conditions montre qu'il n'existe pas de différence significative entre les locuteurs ($F=1.623$; $P=0.182$).

le tableau montre la moyenne pour chaque condition en valeur MOS ainsi que leurs valeur R

Conditions	Moyenne MOS	Ecart type	Moyenne R	Moyenne MOS1
G.711	3.18	0.876	61.6446	4.10
G.726 (32kbits/s)	2.13	0.744	41.3458	2.93
G.729A	2.75	0.673	53.3235	3.52
IS-54	1.96	0.610	37.7768	2.60
G.726 (24kbits/s)*2	1.50	0.584	27.2688	1.72
300-3400 Hz	3.02	0.825	58.4948	3.63
600-3600 Hz	2.71	0.704	52.4989	3.13
300-2000 Hz	1.89	0.703	36.3646	2.05
600-2000 Hz	1.62	0.608	30.2550	1.68
G.722	2.74	0.797	53.1174	-
G.722.2 (23.05 Kbits/s)	4.16	0.730	83.9168	-
G.722.2 (12.65 Kbits/s)	3.91	0.608	77.1784	-
G.722.2 (6.6 Kbits/s)	2.96	0.847	57.2478	-
50-7000 Hz	.426	0.876	87.0731	-
200-7000 Hz	4.17	0.797	84.2474	-
600-7000 Hz	3.22	0.765	62.2819	-
100-5000 Hz	3.73	0.782	72.8634	-
50-3400 Hz	2.72	0.83	52.7051	-
Total	2.92	0.816	57.2019	2.82

Tab IV.7. Résultats du test MOS2

La figure suivante représente la moyenne pour chaque condition avec écart type :

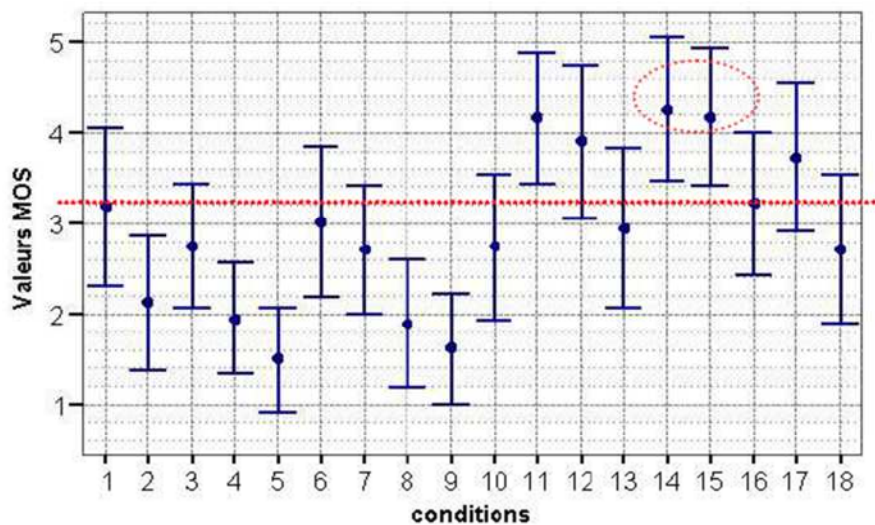


Figure IV.6. Moyenne pour chaque condition avec écart type

Une des premières informations données par la figure est la préférence de presque toutes les conditions large-bande sauf la 10, 13 et 18. Les valeurs de ces dernières montrent qu'une dégradation trop importante du signal liée à un CoDec est plus importante dans l'évaluation globale du stimulus que l'augmentation de la bande passante. Les trois conditions du CoDec AMR-WB montrent effectivement que la qualité de la transmission diminue quand le débit diminue. Enfin, nous observons que l'augmentation de la bande passante dans les fréquences basses sans changer la fréquence de coupure haute, induit une qualité inférieure du stimulus que la référence bande-étroite G.711. Ces fréquences peuvent donc jouer un rôle négatif dans la compréhension du message malgré une sensation de la voix plus proche de la réalité. Il faut aussi remarquer qu'une légère différence, non significative ($F=0.571$; $P=0.451$) de 0.1 point MOS, existe entre les deux conditions 50-7000 Hz et 200-7000 Hz, la première étant préférée sur la deuxième.

Un calcul d'amélioration entre certaines conditions large-bande et la condition de référence G.711, nous donne :

- Une amélioration de 1.08 points MOS pour la condition 50-7000 Hz (soit 34%).
- Une amélioration de 0.99 points MOS pour la condition 200-7000 Hz (soit 31%).

IV.3.3. Comparaison entre MOS1 et MOS2

le tableau suivant montre la répartition des résultats sur l'échelle MOS pour le premier et second test.

Valeur MOS	1	2	3	4	5	Total
% test MOS1	11.9	28..3	31.9	22.1	5.8	100.0
% test MOS2	10.3	27.9	31.5	19.9	10.4	100.0

Tab IV.8. Comparaison des résultats entre MOS1 et MOS2

On remarque que la répartition est différente entre les deux tests, on obtient le double pour la valeur 5. Ce la montre une augmentation globale de la qualité sur l'ensemble des stimuli

IV.3.4. Test avec l'échelle R :

Ce test utilise les 18 conditions bande-étroite et bande élargie. L'interface de jugement correspond à une échelle R du modèle agrandie au dessus de 100 (100 correspond au maximum de qualité pour la téléphonie bande-étroite) et sur laquelle quatre boutons permettant d'écouter le stimulus utilisant les quatre CoDec de référence, en effet la même conversation était jouée quelque soit la condition à évalué par le sujet. Ceci permettait au sujet de mémoriser la qualité des références afin de les comparer plus facilement aux stimuli. Les quatre références sont :

- La référence bande-étroite, le CoDec G.711, positionné à 100 sur l'échelle R.
- Le CoDec G.729, positionné à 90.
- Le CoDec IS-54, positionné à 80.
- Le tandem de CoDecs G.726 à 24 Kbits/s, positionné à 50.

Ce test a été effectué dans les deux pays France et Allemagne avec à chaque fois deux locuteurs, un masculin et un féminin. Il est important de préciser également que tous les stimuli correspondaient à la même conversation processé avec des conditions différentes.

La figure suivante montre les notes attribuées pour chaque locuteur

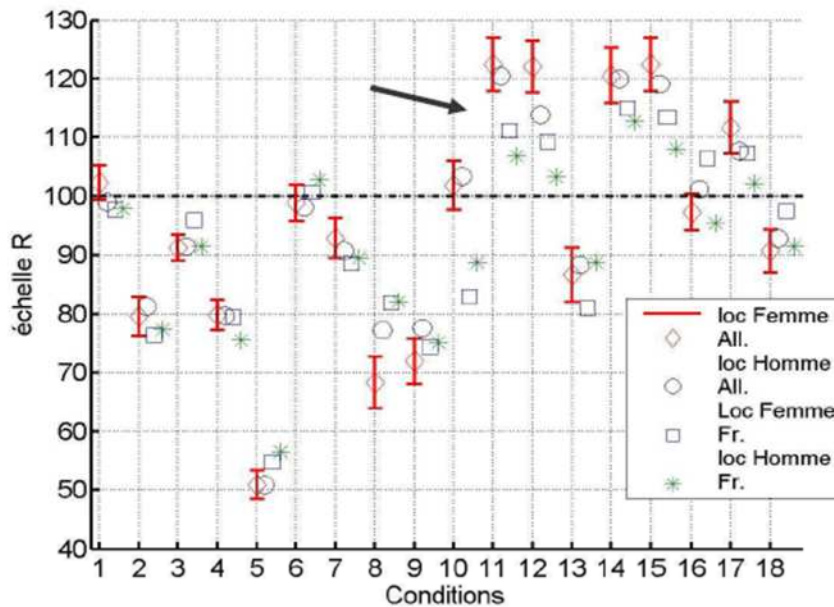


Figure IV.6. Résultats MOS2 pour chaque condition

On remarque des variations plus importantes entre les locuteurs pour chaque condition. Pour certaines conditions, il existe deux groupes distincts entre les deux locuteurs allemands et les deux Français.

Pour les conditions 10 à 18 (celles en bande élargie), on observe qu'une proportion plus importante de points se situe au dessus de la barre 100. Cette ligne représente la qualité théorique de la référence bande-étroite, le G.726.

- Comparaison entre locuteurs

-Test allemand

Les résultats obtenus avec les locuteurs allemands sont donnés dans le tableau qui suit :

Condition	Valeur MOS femme 	Ecart type	Valeur MOS homme 	Ecart type
G.711	102.29	8.615	99.05	5.084
G.726 (32kbts/s)	76.62	10.925	81.33	8.639
G.729	91.24	4.816	91.52	3.544
IS-54	79.86	6.560	79.76	1.513
G.726 (24kbts/s)*2	50.86	5.833	50.90	7.739
300-3400 Hz	98.86	9.345	98.14	11.573
600-3600 Hz	92.90	11.322	90.81	6.161
300-2000 Hz	68.29	19.499	77.29	14.217
600-2000 Hz	71.95	14.746	77.67	14.434
G.722	101.81	17.505	103.38	17.834
G.722.2 (23.05 Kbits/s)	122.43	20.289	120.48	18.597
G.722.2 (12.65 Kbits/s)	122.10	19.131	113.90	17.216
G.722.2 (6.6 Kbits/s)	86.67	21.355	88.48	19.884
50-7000 Hz	120.52	22.2	120.00	21.881
200-7000 Hz	122.48	20.868	119.19	20.532
600-7000 Hz	97.29	9.514	101.19	13.512
100-5000 Hz	111.67	19.417	107.67	12.959
50-3400 Hz	90.76	13.123	92.90	15.016
Total	95.20	24.832	95.09	22.370

Tab IV.9. Résultats du test avec l'échelle R pour les locuteurs allemands

Un calcul d'ANOVA entre les locuteurs, sur l'ensemble des conditions montre qu'il n'existe pas de différence significative entre les locuteurs ($F=0.005$; $P=0.946$). Le tableau montre les résultats pour les deux locuteurs. Un second calcul d'ANOVA pour chaque condition montre qu'il n'y a pas de différence entre les deux locuteurs quelle que soit la condition. Pourtant, la condition G.722.2 à 12.65 Kbits/s ($F=2.921$; $P=0.095$) semble plutôt désavantageux pour le locuteur homme. De même, la condition 300-2000 Hz, montre une différence entre les deux locuteurs ($F=2.127$; $P=0.153$), la fréquence de coupure haute à 2000 Hz semble toucher principalement la locutrice. De plus, la condition G.726 à 32 Kbits/s, semble également être évaluée de moins bonne qualité que sa valeur théorique (qui est normalement 93).

-Test français

Les résultats obtenus avec les locuteurs français sont données dans le tableau qui suit :

Condition	Valeur MOS femme □	Ecart type	Valeur MOS homme *	Ecart type
G.711	97.93	4.713	97.73	7.005
G.726 (32kbits/s)	77.33	11.667	76.4	13.200
G.729	91.47	9.270	95.93	5.489
IS-54	75.60	7.615	97.47	11.969
G.726 (24kbits/s)*2	56.47	10.557	54.80	12.194
300-3400 Hz	102.80	10.913	100.67	11.839
600-3600 Hz	89.53	13.285	88.73	8.271
300-2000 Hz	82.13	15.281	81.93	19.324
600-2000 Hz	75.13	16.561	74.4	13.092
G.722	88.8	17.714	83.07	13.203
G.722.2 (23.05 Kbits/s)	106.8	13.246	111.20	13.311
G.722.2 (12.65 Kbits/s)	103.27	9.225	109.33	10.032
G.722.2 (6.6 Kbits/s)	88.8	14.646	81.07	16.372
50-7000 Hz	112.80	13.177	114.93	16.032
200-7000 Hz	108.13	11.432	113.43	9.234
600-7000 Hz	95.47	21.189	106.53	11.103
100-5000 Hz	102.13	11.012	107.4	15.108
50-3400 Hz	91.47	14.667	97.53	13.851
Total	91.45	20.496	93.03	18.544

Tab IV.10. Résultats du test avec l'échelle R pour les locuteurs français

Pour les résultats de ce test, comme ceux du test précédent, les calculs d'ANOVA ont montré qu'il n'y a pas de différence significative entre les locuteurs sur l'ensemble des conditions ($F=0.888$; $P=0.346$) ni pour une condition particulière. Cependant certaines conditions montrent des différences entre les locuteurs. La condition G.722.2 à 6.6 Kbits/s ($F=1.859$; $P=0.184$), se révèle ici désavantageuse pour la locutrice. La bande passante 600-7000 Hz, montre également une différence entre les deux locuteurs ($F=3.210$; $P=0.084$). En effet, une fréquence de coupure basse semble très désavantageuse pour le locuteur. La condition 50-3400 Hz, montre à l'inverse un désavantage pour le locuteur en raison de l'accentuation trop prononcée des fréquences graves ($F=1.859$; $P=0.184$).

- **Comparaison entre conditions :**

Conditions	Moyenne Allemagne	Ecart type	Moyenne France	Ecart type
G.711	100.67	7.176	97.83	5.867
G.726 (32kbits/s)	80.48	9.766	76.87	12.250
G.729	91.38	4.179	93.7	7.822
IS-54	79.81	4.702	77.53	10.051
G.726 (24kbits/s)*2	50.88	6.769	55.63	11.239
300-3400 Hz	98.50	10.395	101.73	11.240
600-3600 Hz	91.86	9.065	89.13	10.881
300-2000 Hz	72.79	17.459	82.03	17.117
600-2000 Hz	74.81	14.699	74.77	14.673
G.722	102.60	17.473	85.93	15.625
G.722.2 (23.05 Kbits/s)	121.45	19.248	109.00	13.238
G.722.2 (12.65 Kbits/s)	118.00	18.447	106.30	9.959
G.722.2 (6.6 Kbits/s)	87.57	20.400	84.93	15.761
50-7000 Hz	120.26	21.773	113.87	14.460
200-7000 Hz	120.83	20.514	110.80	10.565
600-7000 Hz	99.24	11.710	101.00	17.548
100-5000 Hz	109.67	16.430	104.77	13.263
50-3400 Hz	91.83	13.972	94.50	14.352
Total	95.15	23.618	92.24	19.542

Tab IV.11. Résultats de la comparaison entre conditions

Une comparaison entre les conditions permet d'observer que la qualité des stimuli, pour les conditions G.722.2 baisse avec le débit. Puis un calcul d'ANOVA entre les conditions en comparant chaque condition à la référence G.711 montre que :

- Pour le test effectué en Allemagne, il existe une différence significative entre la référence bande-étroite et toutes les conditions large-bande sauf pour les conditions 300-3400 Hz, G.722 et 600-7000 Hz.
- De même, pour le test en France, il existe une différence significative sauf pour les conditions G.729, 300-3400 Hz, 600-7000 Hz, 100-5000 Hz et 50-3400 Hz.

Nous pouvons en conclure que la fréquence de coupure basse à 50 Hz, est considérée comme une dégradation pour les sujets allemands mais pas pour les Français. Aussi une augmentation de 2 Barks de la bande passante pour cette condition et une augmentation de 1.5

Barks pour la condition 600-7000 Hz ne semblent pas être suffisante pour améliorer la qualité de transmission.

- Les conditions large-bande semblent avoir une meilleure qualité que la référence G.711 sauf la condition G.722.2 à 6.6 Kbits/s. Ceci montre qu'une dégradation trop importante du signal liée à un CoDec ne permet pas une amélioration grâce à la bande élargie.

Discussion :

Ce chapitre a permis de mettre en évidence les différentes techniques utilisées pour évaluer la qualité de la parole. Ces techniques ont été appliquées suivant plusieurs conditions d'acquisition du signal de parole. Selon les résultats d'évaluation, la meilleure qualité est obtenue en utilisant la condition large bande.

CONCLUSION

Conclusion

Ce travail a pour but d'évaluer la qualité du signal de parole et de proposer une solution pour améliorer cette qualité en réduisant les dégradations.

Pour représenter toutes les situations, un ensemble de conditions a été choisi en fonction de type de CoDec et de la bande passante. Pour ce faire, nous avons utilisé une méthode subjective permettant d'évaluer la qualité du signal de parole. Cette méthode est basée sur un ensemble de tests, à savoir MOS et modèle E en utilisant les différentes échelles : ACR, DCR et l'échelle R.

Le premier test appliqué sur notre base de données est MOS1, en utilisant l'échelle ACR il a montré qu'en absence de compression de CoDec la qualité sonore est meilleure et que le codage successif du même CoDec induit une dégradation importante.

Le deuxième test MOS2 appliqué sous la bande-étroite et la bande élargie a montré que la qualité obtenue est meilleure pour la large-bande comparativement à la qualité obtenue sous la bande étroite. En effet, la qualité de transmission diminue quand le débit diminue, et que l'augmentation de la bande passante sans changer la fréquence de coupure induit une dégradation.

Le test du modèle E appliqué sous toutes les conditions, a montré que la fréquence de coupure haute à 200Hz touche principalement les locutrices. Alors que la fréquence de coupure basse à 56 Hz induit une dégradation chez les locuteurs. Une amélioration de 1.08 points sur l'échelle MOS pour la condition 50-7000 (soit 34%) et une amélioration de 0.99 points sur l'échelle MOS pour la condition 200-7000 (soit 31%).

Tous ces résultats montrent qu'on obtient une nette amélioration de la qualité de la parole lors d'une transmission téléphonique grâce à l'utilisation de la bande passante élargie (50-7000 Hz). Par conséquent, il est important d'inclure cette condition dans les systèmes de téléphonie. Comme il est intéressant d'utiliser des CoDecs avec un débit de transmission important.

En perspective, pour bien évaluer la qualité de la parole, les méthodes objectives peuvent compléter l'évaluation globale, en particulier, en utilisant des tests comme le PESQ, PSQM, BSD et MBSD.

Conclusion

Ce travail a pour but d'évaluer la qualité du signal de parole et de proposer une solution pour améliorer cette qualité en réduisant les dégradations.

Pour représenter toutes les situations, un ensemble de conditions a été choisi en fonction de type de CoDec et de la bande passante. Pour ce faire, nous avons utilisé une méthode subjective permettant d'évaluer la qualité du signal de parole. Cette méthode est basée sur un ensemble de tests, à savoir MOS et modèle E en utilisant les différentes échelles : ACR, DCR et l'échelle R.

Le premier test appliqué sur notre base de données est MOS1, en utilisant l'échelle ACR il a montré qu'en absence de compression de CoDec la qualité sonore est meilleure et que le codage successif du même CoDec induit une dégradation importante.

Le deuxième test MOS2 appliqué sous la bande-étroite et la bande élargie a montré que la qualité obtenue est meilleure pour la large-bande comparativement à la qualité obtenue sous la bande étroite. En effet, la qualité de transmission diminue quand le débit diminue, et que l'augmentation de la bande passante sans changer la fréquence de coupure induit une dégradation.

Le test du modèle E appliqué sous toutes les conditions, a montré que la fréquence de coupure haute à 200Hz touche principalement les locutrices. Alors que la fréquence de coupure basse à 56 Hz induit une dégradation chez les locuteurs. Une amélioration de 1.08 points sur l'échelle MOS pour la condition 50-7000 (soit 34%) et une amélioration de 0.99 points sur l'échelle MOS pour la condition 200-7000 (soit 31%).

Tous ces résultats montrent qu'on obtient une nette amélioration de la qualité de la parole lors d'une transmission téléphonique grâce à l'utilisation de la bande passante élargie (50-7000 Hz). Par conséquent, il est important d'inclure cette condition dans les systèmes de téléphonie. Comme il est intéressant d'utiliser des CoDecs avec un débit de transmission important.

En perspective, pour bien évaluer la qualité de la parole, les méthodes objectives peuvent compléter l'évaluation globale, en particulier, en utilisant des tests comme le PESQ, PSQM, BSD et MBSD.

Listes des références bibliographiques

- (Alireza et Siamak, 2004) : Alireza Zolghadr-E-Asli, Siamak Alipour. An effective method for still image compression/decompression for transmission on PSTN lines based on modifications of Huffman coding. 2004.
- .(Beerendes 92) : J.Beerebndes & J.Stemerdink. A perceptual audio quality measurment based on a psychoacoustic sound representation. J.Audio Eng. Soc, vol.40, 1992.
- (Berouti 79) : M. Berouti, R. Schwartz & J. Makhoul. Enhancement of speech corrupted by acoustic noise. In Proc. of ICASSP, volume I, 1979.
- (Buneit 97): L. Bunieti. Traitement automatique de la parole en milieu bruité; études de modeles statistiques et dynamiques . Université Henri-Nancy 1997.
- (G.107 03) : UIT-T G.107. Le mod`ele E : mod`ele de calcul utilis`e pour la planification de la transmission. 2003.
- (G.729 96) : IUT-T Rec. G.729. Coding of speech at 8 kbit/s using conjugate structure algebraic-Code-Excited Linear Prediction (CS-ACELP). 1996.
- (Gopal et al, 1984) : P.M Gopal, J.W Wong, J.C Majithia. Analysis of playout strategies for voice transmission using packet switching techniques. Performance Evaluation, Volume 4, Issue 1, February 1984
- ITU-T Rec.G.107 (2005). The E-model, a Computational Model for Use in Transmission Planning. International Telecommunication Union, CH-Geneva.
- ITU-T Rec. P.833 (2001). Methodology for Derivation of Equipement Impairment Factors from Subjective Listening Tests. International Telecommunication Union, CH-Geneva.
- (Jun-ich et Minoru, 1993): Jun-ichi Mizusawa, Minoru Akiyama. Experiences and expectations of introducing PSTN new services. Computer Networks and ISDN Systems, Volume 25, Issue 10, May 1993.
- (Krebber,1995): Krebber, W. (1995). Sprachübertragungsqualität von Fernsprech-Handapparaten, Vol. 357.Series 10, VDI-Verlag GmbH, DE-Düsseldorf.

- (Manelaos et al, 2001) : Menelaos K. Perdikeas, Iakovos S. Venieris. Parlay-based service engineering in a converged Internet-PSTN environment. Volume 23, Issue 8, 1 April 2000,
- Möller, S. et Raake, A. (2002). Telephone Speech Quality Prediction : Towards Network Planning and Monitoring Models for Modern Network Scenarios. *Speech Communication*, 38(1-2), 47-75.
- (Moore,2003): Moore, B. C. J. and Tan, C-T. (2003). Perceived naturalness of spectrally distorted speech and music. *J. Acoust. Soc. Am.* Vol. 114 (A, 408-19).
- (P.835 03): UIT-T P.835. Subjective test methodology for evaluating speech communication systems that include noise suppression algorithm. 2003.
- (P.861 98) : UIT-T P.861. Objective quality measurement of telephone-band (300-3400 Hz) speech codecs. 1998.
- (Raake,2000): Raake, A. (2000). Perceptual Dimensions of Speech Sound Quality in Modern Transmission Systems. *ICSLP*, 2000.
- (Rajarshi et al, 2009) : Rajarshi Mahapatra, Anindya Sundar Dhar, Debasish Datta. Integrated voice and data transmission employing adaptive modulation in wireless networks. *International Journal of Electronics and Communications*, Volume 63, Issue 12, December 2009.
- (Rogelio, 2008): Rogelio Martínez Perea. Interworking with PSTN/PLMN. *internet Multimedia Communications Using SIP*, 2008.
- (Stuart et Razavi, 2008): Stuart D. Walker, Rouzbeh Razavi. *Voice Transmission. The Cable and Telecommunications Professionals' Reference* (Third Edition), 2008
- (Tohkura 87):Y. Tohkura. A Weighted Cepstral Distance Measure fo Speech Recognition. *IEEE Trans. Acoust., Speech & SignalProcess.*, vol. ASSP-35, pages 1414–1422, 1987.
- (Virole, 2001) :B. Virole. *Psychologie de la surdite*. 2`eme Edition, De Boeck Université, 2001.

- (Wang, 1992): S. Wang, A. Sekey & A. Gersho. An objective measure for predicting subjective quality of speech coders. IEEE J. on Select. Areas in Commun., vol. SAC-10, pages 819–829, Sept 1992.
- (Wang, 1997): S. Wang, A. Sekey & A. Gersho. Modified bark spectral distortion measure which uses noise masking threshold. IEEE Speech Coding Workshop, vol. SAC-10, 1997.
- (Yang, 1997) : W. Yang, M. Dixon & R. Yantorno. Modified bark spectral distortion measure which uses noise masking threshold. In IEEE Speech coding Workshop, 1997.
- (Yang, 1999) : W. Yang, M. Dixon & R. Yantorno. Enhanced modified Bark spectral distortion (EMBSD) : An objective speech quality measure based on audible distortion and cognition model. Phd thesis, May 1999.
- (Zhijun, 2015) : Zhijun Wu. Design of Real-Time Speech Secure Communication over PSTN. Information Hiding in Speech Signal for Secure Communication, 2015.

Listes des références bibliographiques

(Alireza et Siamak, 2004) : Alireza Zolghadr-E-Asli, Siamak Alipour. An effective method for still image compression/decompression for transmission on PSTN lines based on modifications of Huffman coding. 2004.

.(Beerendes 92) : J.Beerebndes & J.Stemerdkink. A perceptual audio quality measurment based on a psychoacoustic sound representation. J.Audio Eng. Soc, vol.40, 1992.

(Berouti 79) : M. Berouti, R. Schwartz & J. Makhoul. Enhancement of speech corrupted by acoustic noise. In Proc. of ICASSP, volume I, 1979.

(Buneit 97): L. Bunieti. Traitement automatique de la parole en milieu bruité; études de modeles statistiques et dynamiques . Université Henri-Nancy 1997.

(G.107 03) : UIT-T G.107. Le mod`ele E : mod`ele de calcul utilis`e pour la planification de la transmission. 2003.

(G.729 96) : IUT-T Rec. G.729. Coding of speech at 8 kbit/s using conjugate structure algebraic-Code-Excited Linear Prediction (CS-ACELP). 1996.

(Gopal et al, 1984) : P.M Gopal, J.W Wong, J.C Majithia. Analysis of playout strategies for voice transmission using packet switching techniques. Performance Evaluation, Volume 4, Issue 1, February 1984

ITU-T Rec.G.107 (2005). The E-model, a Computational Model for Use in Transmission Planning. International Telecommunication Union, CH-Geneva.

ITU-T Rec. P.833 (2001). Methodology for Derivation of Equipment Impairment Factors from Subjective Listening Tests. International Telecommunication Union, CH-Geneva.

(Jun-ich et Minoru, 1993): Jun-ichi Mizusawa, Minoru Akiyama. Experiences and expectations of introducing PSTN new services. Computer Networks and ISDN Systems, Volume 25, Issue 10, May 1993.

(Krebber,1995): Krebber, W. (1995). Sprachübertragungsqualität von Fernsprech-Handapparaten, Vol. 357.Series 10, VDI-Verlag GmbH, DE-Düsseldorf.

- (Manelaos et al, 2001) : Menelaos K. Perdikeas, Iakovos S. Venieris. Parlay-based service engineering in a converged Internet-PSTN environment. Volume 23, Issue 8, 1 April 2000,
- Möller, S. et Raake, A. (2002). Telephone Speech Quality Prediction : Towards Network Planning and Monitoring Models for Modern Network Scenarios. *Speech Communication*, 38(1-2), 47-75.
- (Moore,2003): Moore, B. C. J. and Tan, C-T. (2003). Perceived naturalness of spectrally distorted speech and music. *J. Acoust. Soc. Am.* Vol. 114 (A, 408-19).
- (P.835 03): UIT-T P.835. Subjective test methodology for evaluating speech communication systems that include noise suppression algorithm. 2003.
- (P.861 98) : UIT-T P.861. Objective quality measurement of telephone-band (300-3400 Hz) speech codecs. 1998.
- (Raake,2000): Raake, A. (2000). Perceptual Dimensions of Speech Sound Quality in Modern Transmission Systems. *ICSLP*, 2000.
- (Rajarshi et al, 2009) : Rajarshi Mahapatra, Anindya Sundar Dhar, Debasish Datta. Integrated voice and data transmission employing adaptive modulation in wireless networks. *International Journal of Electronics and Communications*, Volume 63, Issue 12, December 2009.
- (Rogelio, 2008): Rogelio Martínez Perea. Interworking with PSTN/PLMN. *internet Multimedia Communications Using SIP*, 2008.
- (Stuart et Razavi, 2008): Stuart D. Walker, Rouzbeh Razavi. *Voice Transmission. The Cable and Telecommunications Professionals' Reference* (Third Edition), 2008
- (Tohkura 87):Y. Tohkura. A Weighted Cepstral Distance Measure fo Speech Recognition. *IEEE Trans. Acoust., Speech & SignalProcess.*, vol. ASSP-35, pages 1414–1422, 1987.
- (Virole, 2001) :B. Virole. *Psychologie de la surdite*. 2`eme Edition, De Boeck Université, 2001.

- (Wang, 1992): S. Wang, A. Sekey & A. Gersho. An objective measure for predicting subjective quality of speech coders. IEEE J. on Select. Areas in Commun., vol. SAC-10, pages 819–829, Sept 1992.
- (Wang, 1997): S. Wang, A. Sekey & A. Gersho. Modified bark spectral distortion measure which uses noise masking threshold. IEEE Speech Coding Workshop, vol. SAC-10, 1997.
- (Yang, 1997) : W. Yang, M. Dixon & R. Yantorno. Modified bark spectral distortion measure which uses noise masking threshold. In IEEE Speech coding Workshop, 1997.
- (Yang, 1999) : W. Yang, M. Dixon & R. Yantorno. Enhanced modified Bark spectral distortion (EMBSD) : An objective speech quality measure based on audible distortion and cognition model. Phd thesis, May 1999.
- (Zhijun, 2015) : Zhijun Wu. Design of Real-Time Speech Secure Communication over PSTN. Information Hiding in Speech Signal for Secure Communication, 2015.

Annexe A

Échelle de notation de la norme P-835

Note	Evaluation
5	Dépourvu de distorsion
4	Légèrement distordu
3	Quelque peu distordu
2	Assez distordu
1	Très distordu

Tab A.1 Echelle d'évaluation du signal vocal

Note	Evaluation
5	Imperceptible
4	Légèrement imperceptible
3	Perceptible mais non gênant
2	Quelque peu gênant
1	Très gênant

Tab A.2. Echelle d'évaluation du bruit de fond

Note	Evaluation
5	Excellent
4	Bon
3	Passable
2	Médiocre
1	Mauvais

Tab A.3. Échelle d'évaluation de la qualité globale

Annexe A

Échelle de notation de la norme P-835

Note	Evaluation
5	Dépourvu de distorsion
4	Légèrement distordu
3	Quelque peu distordu
2	Assez distordu
1	Très distordu

Tab A.1 Echelle d'évaluation du signal vocal

Note	Evaluation
5	Imperceptible
4	Légèrement imperceptible
3	Perceptible mais non gênant
2	Quelque peu gênant
1	Très gênant

Tab A.2. Echelle d'évaluation du bruit de fond

Note	Evaluation
5	Excellent
4	Bon
3	Passable
2	Médiocre
1	Mauvais

Tab A.3. Échelle d'évaluation de la qualité globale

Annexe B

Table statistiques

B.1. Distribution de la loi de student

dl \ α	0,9	0,5	0,3	0,2	0,1	0,05	0,02	0,01	0,001
1	0,1584	1	1,9626	3,0777	6,3137	12,706	31,821	63,656	636,58
2	0,1421	0,8165	1,3862	1,8856	2,92	4,3027	6,9645	9,925	31,6
3	0,1366	0,7649	1,2498	1,6377	2,3534	3,1824	4,5407	5,8408	12,924
4	0,1338	0,7407	1,1896	1,5332	2,1318	2,7765	3,7469	4,6041	8,6101
5	0,1322	0,7267	1,1558	1,4759	2,015	2,5706	3,3649	4,0321	6,8685
6	0,1311	0,7176	1,1342	1,4398	1,9432	2,4469	3,1427	3,7074	5,9587
7	0,1303	0,7111	1,1192	1,4149	1,8946	2,3646	2,9979	3,4995	5,4081
8	0,1297	0,7064	1,1081	1,3968	1,8595	2,306	2,8965	3,3554	5,0414
9	0,1293	0,7027	1,0997	1,383	1,8331	2,2622	2,8214	3,2498	4,7809
10	0,1289	0,6998	1,0931	1,3722	1,8125	2,2281	2,7638	3,1693	4,5868
11	0,1286	0,6974	1,0877	1,3634	1,7959	2,201	2,7181	3,1058	4,4369
12	0,1283	0,6955	1,0832	1,3562	1,7823	2,1788	2,681	3,0545	4,3178
13	0,1281	0,6938	1,0795	1,3502	1,7709	2,1604	2,6503	3,0123	4,2209
14	0,128	0,6924	1,0763	1,345	1,7613	2,1448	2,6245	2,9768	4,1403
15	0,1278	0,6912	1,0735	1,3406	1,7531	2,1315	2,6025	2,9467	4,0728
16	0,1277	0,6901	1,0711	1,3368	1,7459	2,1199	2,5835	2,9208	4,0149
17	0,1276	0,6892	1,069	1,3334	1,7396	2,1098	2,5669	2,8982	3,9651
18	0,1274	0,6884	1,0672	1,3304	1,7341	2,1009	2,5524	2,8784	3,9217
19	0,1274	0,6876	1,0655	1,3277	1,7291	2,093	2,5395	2,8609	3,8833
20	0,1273	0,687	1,064	1,3253	1,7247	2,086	2,528	2,8453	3,8496
21	0,1272	0,6864	1,0627	1,3232	1,7207	2,0796	2,5176	2,8314	3,8193
22	0,1271	0,6858	1,0614	1,3212	1,7171	2,0739	2,5083	2,8188	3,7922
23	0,1271	0,6853	1,0603	1,3195	1,7139	2,0687	2,4999	2,8073	3,7676
24	0,127	0,6848	1,0593	1,3178	1,7109	2,0639	2,4922	2,797	3,7454
25	0,1269	0,6844	1,0584	1,3163	1,7081	2,0595	2,4851	2,7874	3,7251
26	0,1269	0,684	1,0575	1,315	1,7056	2,0555	2,4786	2,7787	3,7067
27	0,1268	0,6837	1,0567	1,3137	1,7033	2,0518	2,4727	2,7707	3,6895
28	0,1268	0,6834	1,056	1,3125	1,7011	2,0484	2,4671	2,7633	3,6739
29	0,1268	0,683	1,0553	1,3114	1,6991	2,0452	2,462	2,7564	3,6595
30	0,1267	0,6828	1,0547	1,3104	1,6973	2,0423	2,4573	2,75	3,646

Tab B.1. Distribution de la loi de student

B.2. Distribution de la loi de Fisher

	$n = 1$		$n = 2$		$n = 3$		$n = 4$		$n = 5$	
m	$\alpha = 0.05$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.01$
1	161.45	4052.18	199.50	4999.50	215.71	5403.35	224.58	5624.58	230.16	5763.65
2	18.51	98.50	19.00	99.00	19.16	99.17	19.25	99.25	19.30	99.30
3	10.13	34.12	9.55	30.82	9.28	29.46	9.12	28.71	9.01	28.24
4	7.71	21.20	6.94	18.00	6.59	16.69	6.39	15.98	6.26	15.52
5	6.61	16.26	5.79	13.27	5.41	12.06	5.19	11.39	5.05	10.97
6	5.99	13.75	5.14	10.92	4.76	9.78	4.53	9.15	4.39	8.75
7	5.59	12.25	4.74	9.55	4.35	8.45	4.12	7.85	3.97	7.46
8	5.32	11.26	4.46	8.65	4.07	7.59	3.84	7.01	3.69	6.63
9	5.12	10.56	4.26	8.02	3.86	6.99	3.63	6.42	3.48	6.06
10	4.96	10.04	4.10	7.56	3.71	6.55	3.48	5.99	3.33	5.64
11	4.84	9.65	3.98	7.21	3.59	6.22	3.36	5.67	3.20	5.32
12	4.75	9.33	3.89	6.93	3.49	5.95	3.26	5.41	3.11	5.06
13	4.67	9.07	3.81	6.70	3.41	5.74	3.18	5.21	3.03	4.86
14	4.60	8.86	3.74	6.51	3.34	5.56	3.11	5.04	2.96	4.69
15	4.54	8.68	3.68	6.36	3.29	5.42	3.06	4.89	2.90	4.56
16	4.49	8.53	3.63	6.23	3.24	5.29	3.01	4.77	2.85	4.44
17	4.45	8.40	3.59	6.11	3.20	5.18	2.96	4.67	2.81	4.34
18	4.41	8.29	3.55	6.01	3.16	5.09	2.93	4.58	2.77	4.25
19	4.38	8.18	3.52	5.93	3.13	5.01	2.90	4.50	2.74	4.17
20	4.35	8.10	3.49	5.85	3.10	4.94	2.87	4.43	2.71	4.10
21	4.32	8.02	3.47	5.78	3.07	4.87	2.84	4.37	2.68	4.04
22	4.30	7.95	3.44	5.72	3.05	4.82	2.82	4.31	2.66	3.99
23	4.28	7.88	3.42	5.66	3.03	4.76	2.80	4.26	2.64	3.94
24	4.26	7.82	3.40	5.61	3.01	4.72	2.78	4.22	2.62	3.90
25	4.24	7.77	3.39	5.57	2.99	4.68	2.76	4.18	2.60	3.85
26	4.23	7.72	3.37	5.53	2.98	4.64	2.74	4.14	2.59	3.82
27	4.21	7.68	3.35	5.49	2.96	4.60	2.73	4.11	2.57	3.78
28	4.20	7.64	3.34	5.45	2.95	4.57	2.71	4.07	2.56	3.75
29	4.18	7.60	3.33	5.42	2.93	4.54	2.70	4.04	2.55	3.73
30	4.17	7.56	3.32	5.39	2.92	4.51	2.69	4.02	2.53	3.70
40	4.08	7.31	3.23	5.18	2.84	4.31	2.61	3.83	2.45	3.51
80	3.96	6.96	3.11	4.88	2.72	4.04	2.49	3.56	2.33	3.26
120	3.92	6.85	3.07	4.79	2.68	3.95	2.45	3.48	2.29	3.17
∞	3.84	6.63	3.00	4.61	2.60	3.78	2.37	3.32	2.21	3.02

Tab B.2. Distribution de loi de Fisher

Annexe B

Table statistiques

B.1. Distribution de la loi de student

dl \ α d	0,9	0,5	0,3	0,2	0,1	0,05	0,02	0,01	0,001
1	0,1584	1	1,9626	3,0777	6,3137	12,706	31,821	63,656	636,58
2	0,1421	0,8165	1,3862	1,8856	2,92	4,3027	6,9645	9,925	31,6
3	0,1366	0,7649	1,2498	1,6377	2,3534	3,1824	4,5407	5,8408	12,924
4	0,1338	0,7407	1,1896	1,5332	2,1318	2,7765	3,7469	4,6041	8,6101
5	0,1322	0,7267	1,1558	1,4759	2,015	2,5706	3,3649	4,0321	6,8685
6	0,1311	0,7176	1,1342	1,4398	1,9432	2,4469	3,1427	3,7074	5,9587
7	0,1303	0,7111	1,1192	1,4149	1,8946	2,3646	2,9979	3,4995	5,4081
8	0,1297	0,7064	1,1081	1,3968	1,8595	2,306	2,8965	3,3554	5,0414
9	0,1293	0,7027	1,0997	1,383	1,8331	2,2622	2,8214	3,2498	4,7809
10	0,1289	0,6998	1,0931	1,3722	1,8125	2,2281	2,7638	3,1693	4,5868
11	0,1286	0,6974	1,0877	1,3634	1,7959	2,201	2,7181	3,1058	4,4369
12	0,1283	0,6955	1,0832	1,3562	1,7823	2,1788	2,681	3,0545	4,3178
13	0,1281	0,6938	1,0795	1,3502	1,7709	2,1604	2,6503	3,0123	4,2209
14	0,128	0,6924	1,0763	1,345	1,7613	2,1448	2,6245	2,9768	4,1403
15	0,1278	0,6912	1,0735	1,3406	1,7531	2,1315	2,6025	2,9467	4,0728
16	0,1277	0,6901	1,0711	1,3368	1,7459	2,1199	2,5835	2,9208	4,0149
17	0,1276	0,6892	1,069	1,3334	1,7396	2,1098	2,5669	2,8982	3,9651
18	0,1274	0,6884	1,0672	1,3304	1,7341	2,1009	2,5524	2,8784	3,9217
19	0,1274	0,6876	1,0655	1,3277	1,7291	2,093	2,5395	2,8609	3,8833
20	0,1273	0,687	1,064	1,3253	1,7247	2,086	2,528	2,8453	3,8496
21	0,1272	0,6864	1,0627	1,3232	1,7207	2,0796	2,5176	2,8314	3,8193
22	0,1271	0,6858	1,0614	1,3212	1,7171	2,0739	2,5083	2,8188	3,7922
23	0,1271	0,6853	1,0603	1,3195	1,7139	2,0687	2,4999	2,8073	3,7676
24	0,127	0,6848	1,0593	1,3178	1,7109	2,0639	2,4922	2,797	3,7454
25	0,1269	0,6844	1,0584	1,3163	1,7081	2,0595	2,4851	2,7874	3,7251
26	0,1269	0,684	1,0575	1,315	1,7056	2,0555	2,4786	2,7787	3,7067
27	0,1268	0,6837	1,0567	1,3137	1,7033	2,0518	2,4727	2,7707	3,6895
28	0,1268	0,6834	1,056	1,3125	1,7011	2,0484	2,4671	2,7633	3,6739
29	0,1268	0,683	1,0553	1,3114	1,6991	2,0452	2,462	2,7564	3,6595
30	0,1267	0,6828	1,0547	1,3104	1,6973	2,0423	2,4573	2,75	3,646

Tab B.1. Distribution de la loi de student

B.2. Distribution de la loi de Fisher

	$n = 1$		$n = 2$		$n = 3$		$n = 4$		$n = 5$	
m	$\alpha = 0.05$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.01$
1	161.45	4052.18	199.50	4999.50	215.71	5403.35	224.58	5624.58	230.16	5763.65
2	18.51	98.50	19.00	99.00	19.16	99.17	19.25	99.25	19.30	99.30
3	10.13	34.12	9.55	30.82	9.28	29.46	9.12	28.71	9.01	28.24
4	7.71	21.20	6.94	18.00	6.59	16.69	6.39	15.98	6.26	15.52
5	6.61	16.26	5.79	13.27	5.41	12.06	5.19	11.39	5.05	10.97
6	5.99	13.75	5.14	10.92	4.76	9.78	4.53	9.15	4.39	8.75
7	5.59	12.25	4.74	9.55	4.35	8.45	4.12	7.85	3.97	7.46
8	5.32	11.26	4.46	8.65	4.07	7.59	3.84	7.01	3.69	6.63
9	5.12	10.56	4.26	8.02	3.86	6.99	3.63	6.42	3.48	6.06
10	4.96	10.04	4.10	7.56	3.71	6.55	3.48	5.99	3.33	5.64
11	4.84	9.65	3.98	7.21	3.59	6.22	3.36	5.67	3.20	5.32
12	4.75	9.33	3.89	6.93	3.49	5.95	3.26	5.41	3.11	5.06
13	4.67	9.07	3.81	6.70	3.41	5.74	3.18	5.21	3.03	4.86
14	4.60	8.86	3.74	6.51	3.34	5.56	3.11	5.04	2.96	4.69
15	4.54	8.68	3.68	6.36	3.29	5.42	3.06	4.89	2.90	4.56
16	4.49	8.53	3.63	6.23	3.24	5.29	3.01	4.77	2.85	4.44
17	4.45	8.40	3.59	6.11	3.20	5.18	2.96	4.67	2.81	4.34
18	4.41	8.29	3.55	6.01	3.16	5.09	2.93	4.58	2.77	4.25
19	4.38	8.18	3.52	5.93	3.13	5.01	2.90	4.50	2.74	4.17
20	4.35	8.10	3.49	5.85	3.10	4.94	2.87	4.43	2.71	4.10
21	4.32	8.02	3.47	5.78	3.07	4.87	2.84	4.37	2.68	4.04
22	4.30	7.95	3.44	5.72	3.05	4.82	2.82	4.31	2.66	3.99
23	4.28	7.88	3.42	5.66	3.03	4.76	2.80	4.26	2.64	3.94
24	4.26	7.82	3.40	5.61	3.01	4.72	2.78	4.22	2.62	3.90
25	4.24	7.77	3.39	5.57	2.99	4.68	2.76	4.18	2.60	3.85
26	4.23	7.72	3.37	5.53	2.98	4.64	2.74	4.14	2.59	3.82
27	4.21	7.68	3.35	5.49	2.96	4.60	2.73	4.11	2.57	3.78
28	4.20	7.64	3.34	5.45	2.95	4.57	2.71	4.07	2.56	3.75
29	4.18	7.60	3.33	5.42	2.93	4.54	2.70	4.04	2.55	3.73
30	4.17	7.56	3.32	5.39	2.92	4.51	2.69	4.02	2.53	3.70
40	4.08	7.31	3.23	5.18	2.84	4.31	2.61	3.83	2.45	3.51
80	3.96	6.96	3.11	4.88	2.72	4.04	2.49	3.56	2.33	3.26
120	3.92	6.85	3.07	4.79	2.68	3.95	2.45	3.48	2.29	3.17
∞	3.84	6.63	3.00	4.61	2.60	3.78	2.37	3.32	2.21	3.02

Tab B.2. Distribution de loi de Fisher