

**REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE
MINISTRE DE L'ENSEIGNEMENT SUPERIEUR ET DE LA RECHERCHE
SCIENTIFIQUE**

ECOLE NATIONALE POLYTECHNIQUE



DEPARTEMENT D'ELECTRONIQUE

THESE DE DOCTORAT D'ETAT

Présenté par :

M^r. BATEL NOUREDDINE

Magister en électronique à l'ENP
Ingénieur d'état en électronique à l'ENP

**Conception et Réalisation d'un Implant
Cochléaire pour Mal-entendants**

Soutenue publiquement le : 15 / 12 / 2007

Jury :

**Pr. A. Zerguerras, ENP, Président,
Pr. M. Mehenni, ENP, Rapporteur,
Pr. M. Guerti, ENP, Examineur,
Dr. C. Larbes, ENP, Examineur,
Dr. M. Trabelsi, ENP, Examineur,
Dr. M. Talha-Kedir, USTHB, Invitée,
Pr. D. Djennaoui, CHU Mustapha-Bacha, Invité,**

REMERCIEMENTS

Tous mes remerciements au tout puissant d'abord.

Mes vifs remerciements iront ensuite à mes chers parents, surtout pour leurs incessantes prières et inconditionnels soutiens.

Je n'oublierai pas de remercier ma femme, de son aide continue, sa grande patience, et son entière compréhension.

Je remercie chaleureusement le professeur MEHENNI, pour ce travail qu'il m'a proposé, ses articles qui m'ont inspiré, et la patience qu'il a montrée durant tout ce temps.

Mes remerciements vont également aux membres du jury, qui ont bien voulu lire, corriger, et valoriser ma modeste thèse.

Je tiens à remercier en particulier messieurs : KHELIL-CHORFI Smaïl, KERMEZLI Taïb, BENREKIA Fayçal, sans oublier MEMI Fethi, de leurs soutiens jusqu'au dernier moment.

Je remercie humblement le docteur chirurgien OUADIA, qui a été la cause de la réussite des opérations chirurgicales que j'ai subit durant deux années après mon accident de voiture.

Mes remerciements iront également au docteur ORL, ABDELMOUMEN qui m'a fourni ses encyclopédies médicales desquelles j'ai tiré certaines planches anatomiques et principes physiologiques de l'oreille.

Je ne manquerai pas de remercier les responsables de l'institut des sciences de l'ingénieur, ainsi que l'ensemble de mes collègues au département électronique du centre universitaire de Médéa.

A tous ceux ou celles qui m'ont été d'une quelconque aide ou utilité, qu'ils trouvent ici mes vifs remerciements.

A tous ceux ou celles qui m'ont été d'un quelconque soutien morale ou physique, qu'ils trouvent ici mon entière gratitude.

A tous ceux ou celles qui ont contribué de près ou de loin dans ce présent travail, qu'ils trouvent ici ma totale reconnaissance.

Mots-clés

Implant cochléaire, Stimulation électrique, Alimentation par champ externe, Bobines de transfert, Tissu biologique, Facteur de transmission de puissance, Circuits de transmission et de réception, Formants et Pitch, Analyse LPC, Modulation MSK, Sonagramme, Traitement de la parole, Taux d'erreur binaire.

Abstract :

Within the framework of the rehabilitation of hearing at the deaf people by cochlear implant, we propose through this work, the design and realization of the transmitter and receiver of the implant. A new strategy of cochlear stimulation is also proposed. This strategy uses the parametric model of the speech, accomplished in real time by a DSP. The relevant parameters of the speech are thus extracted in short times, and transmitted in a low rate towards the implant. A measurement of the performances of developed circuits and softwares was done.

Résumé :

Dans le cadre de la réhabilitation de l'audition chez les mal-entendants par le biais d'implant cochléaire, nous proposons à travers ce travail, la conception et réalisation du transmetteur et du récepteur de l'implant. Une nouvelle stratégie de stimulation cochléaire est également proposée. Cette stratégie repose sur la modélisation paramétrique de la parole, accomplie en temps réel par un DSP. Les paramètres pertinents de la parole sont ainsi extraits à court terme, et transmis à faible débit vers l'implant. Une mesure des performances des circuits et logiciels développés a été menée.

SOMMAIRE

Introduction	1
CHAPITRE 1. Anatomie et Physiologie de l'oreille humaine	
1.1. Oreille externe et oreille moyenne	2
1.2. Anatomie de l'oreille interne	2
1.2.1. Le labyrinthe	2
1.2.2. La cochlée	2
1.2.3. Le canal cochléaire	3
1.2.4. L'organe de Corti	4
1.2.5. Innervation des cellules ciliées	4
1.2.5.1. Les fibres nerveuses afférentes	5
1.2.5.2. Les fibres nerveuses efférentes	5
1.3. Physiologie de l'oreille interne	5
1.3.1. Comportement mécanique et simulation de fonctionnement	5
1.3.2. L'amplificateur cochléaire	8
1.3.3. Le potentiel de récepteur des CCI	9
1.3.4. Schéma fonctionnel d'une oreille saine	9
1.4. Psychophysique de la stimulation électrique du nerf auditif	10
1.4.1. Ondes rectangulaires ou sinusoïdales	10
1.4.2. Stimulation électrique de la fréquence acoustique	11
1.4.3. Stimulation électrique de l'intensité acoustique	11
1.4.4. Dynamique de la stimulation électrique	11
1.5. La surdité	11
1.5.1. Surdité de transmission	11
1.5.2. Surdité neurosensorielle	11
1.5.2.1. Les surdités totales	12
1.5.2.2. Les surdités profondes	12
1.5.2.3. Les surdités sévères	12
Conclusion	12
CHAPITRE 2. Implants cochléaires commerciaux	
2.1. Candidats à l'implant cochléaire	13
2.2. Caractéristiques d'implant	13
2.2.1. Conception d'électrode	14
2.2.2. Type de stimulation	16
2.2.3. Liaison de Transmission	16
2.2.4. Traitement du Signal	17
2.3. Stratégies de codage	17
2.3.1. Stimulation analogique simultanée	17
2.3.2. Stimulation impulsionnelle séquentielle	20
2.3.2.1. Extraction de formants	20
2.3.2.2. Extraction d'enveloppes	21
2.3.3. Résultats d'implantation avec les systèmes multi-canaux	27
2.4. Les constituants communs d'un implant cochléaire	28
2.5. Choix d'implants	29
2.6. Proposition d'une stratégie améliorée	29
2.6.1. Bilan des données physiologiques et psycho-acoustiques	30
2.6.1.1. Intensité électrique de stimulation	30

2.6.1.2. Aspect de la théorie du temps	30
2.6.1.3. Aspect de la théorie du lieu	30
2.6.1.4. Découpage en tranches de la parole	31
2.6.2. Principe d'analyse spectrale	31
2.6.3. Caractéristiques de la prothèse cochléaire suggérée	31
Conclusion	32

CHAPITRE 3. Alimentation RF transdermale

3.1. Transfert de puissance par couplage magnétique	34
3.1.1. Calcul du potentiel magnétique d'une bobine circulaire	34
3.1.2. Couplage entre bobines désalignées	37
3.1.3. Atténuation en puissance par désalignement des bobines	40
3.2. Interaction des ondes électromagnétiques avec le tissu biologique	40
3.2.1. Propriétés électriques des tissus biologiques	40
3.2.2. Atténuation en puissance due au tissu biologique	42
3.3. Confrontation avec l'expérience	43
3.4. Application au cas d'un implant cochléaire	44
3.4.1. Alimentation de l'implant par champ externe	44
3.4.2. Recherche du coefficient d'induction mutuelle optimal	44
3.4.3. Estimation de la perte due au tissu biologique	45
3.4.4. Calcul optimal des circuits accordés de l'alimentation	47
Conclusion	52

CHAPITRE 4. Analyse de la parole et sonagramme

4.1. Signal de la parole	53
4.1.1. Description phonatoire	53
4.1.2. Caractéristiques de la parole	54
4.1.3. Modélisation de la parole	55
4.1.3.1. Théorie de la production de la parole	56
4.1.3.2. Modèle de production de la parole	57
4.2. Méthodes d'analyse de la parole	59
4.2.1. Analyse par prédiction linéaire	59
4.2.1.1. Estimation du modèle autorégressif	59
4.2.1.2. Méthode de résolution et stabilité du modèle	62
4.2.2. Analyse conjointe temps-fréquence	63
4.2.2.1. Représentation par Ondelettes	63
4.2.2.2. Représentation par Transformée de Fourier glissante	65
4.3. Outil de comparaison par Sonagramme	66
4.3.1. Environnement requis du PC	66
4.3.2. Logiciel de traitement et d'enregistrement / restitution	67
4.3.2.1. L'unité PCoutil.Pas	67
4.3.2.2. L'unité detectSB.Pas	67
4.3.2.3. L'unité Blaster.Pas	67
4.3.2.4. L'unité Analyse.Pas	67
4.3.2.5. Le programme principal Son.Pas	68
4.3.2.6. La procédure d'enregistrement-restitution DSP_RecPlay	70
4.3.2.7. Procédures d'analyse du fichier audio	72
Conclusion	76

CHAPITRE 5. Implémentation de la nouvelle stratégie d'implant sur DSP

5.1. Algorithmes de traitement	77
5.1.1. Prétraitements du signal de la parole	77
5.1.1.1. Recouvrement et pondération	77
5.1.1.2. Préaccentuation	78
5.1.2. Traitements du signal de la parole	78
5.1.2.1. Senseur silence / parole	81
5.1.2.2. Algorithme de Schur	81
5.1.2.3. Mesure de la valeur efficace vraie d'intensité	82
5.1.2.4. Mesure des formants et bandes	82
5.1.2.5. Algorithme de Laguerre	83
5.1.2.6. Senseur de voisement et détection SIFT du pitch	86
5.2. Le processeur DSP de traitement	89
5.2.1. Choix du DSP	89
5.2.2. Architecture du DSP	89
5.2.2.1. Timer	90
5.2.2.2. McBSP	90
5.2.2.3. EDMA	90
5.2.3. Interruptions	92
5.2.4. La carte DSK-C6711	93
5.2.4.1. Caractéristiques principales	93
5.2.4.2. Codec TLC320AD535	93
5.3. Logiciel d'accompagnement DSP/BIOS	94
5.3.1. Catégorie Scheduling	94
5.3.2. Catégorie Instrumentation	96
Conclusion	96

CHAPITRE 6. Transmetteur-Récepteur de l'implant

6.1. Conception et réalisation du transmetteur d'implant	97
6.1.1. Cahier des charges	97
6.1.2. Codage de canal	98
6.1.3. Codage de source	100
6.1.4. Choix de la modulation discrète	100
6.1.4.1. Modulation CPM	101
6.1.4.2. Modulation MSK	102
6.1.5. Implémentation du modulateur MSK sur GAL	104
6.1.5.1. Présentation et description du GAL	104
6.1.5.2. Programmation du GAL	105
6.1.5.3. Application au modulateur MSK	105
6.1.5.4. Application au générateur pseudo-aléatoire	107
6.1.6. Filtre de transmission	108
6.1.7. Modulation HF à PLL	110
6.1.7.1. Filtre de boucle	113
6.1.7.2. Additionneur du signal modulant	114
6.1.8. Amplification HF	115
6.1.8.1. Organe de rayonnement HF	115
6.1.8.2. Amplificateur de puissance	116
6.2. Conception et réalisation du récepteur d'implant	119
6.2.1. Aspect général	119
6.2.2. Circuit limiteur	119

6.2.3. Amplificateur-isolateur HF	120
6.2.4. Discriminateur d'amplitude	122
6.2.5. Démodulateur MSK	123
6.2.5.1. Circuit de mise en forme	123
6.2.5.2. Détecteur de passage par zéro	124
6.2.5.3. Temporisateurs de prétraitement	124
6.2.5.4. Régénérateur de données	125
6.3. Logiciel de la réalisation	126
6.3.1. Développement du projet d'implant sous CCS	126
6.3.1.1. Fichier de configuration	126
a) Application de la catégorie Instrumentation	126
b) Application de la catégorie Scheduling	127
c) Application de la catégorie Input/Output	127
d) Application de la catégorie CSL	127
6.3.1.2. Fichiers-source	128
a) Aquisition.c	128
b) Analyse.c	130
c) Trans_Rcpt.c : Paquetisation des paramètres	130
d) Fichiers en V.Basic	133
6.3.1.3. Fichiers de mesure	135
a) Tracé du spectrogramme à formants	135
b) Routines du BER et diagramme en œil	135
6.3.2. Espaces de travail des projets de mesure	136
6.3.2.1. Espace de travail du projet Parole.pjt	136
6.3.2.2. Espace de travail du projet BER_DiagŒil.pjt	137
6.4. Schéma fonctionnel final de l'implant cochléaire d'essai	138
6.4.1. Synoptique de transmission des paramètres d'implant	138
6.4.2. Synoptique de réception des paramètres d'implant	138
Conclusion	139

CHAPITRE 7. Performance de la réalisation

7.1. Comparaison des traitements LPC-Sonagramme de la parole	140
7.1.1. Présentation des fichiers audio d'essai	140
7.1.2. Description et comparaison des spectrogrammes et signaux obtenus	140
7.1.2.1. Analyse du fichier parenthèse.aud	141
7.1.2.2. Analyse du fichier lapin.aud	143
7.2. Performance des cartes réalisées	147
7.2.1. Mesure des performances d'une transmission de données	147
7.2.1.1. Taux d'erreur binaire	147
7.2.1.2. Diagramme en œil	147
7.2.1.3. Gigue de phase	149
7.2.2. Tests et résultats pratiques des circuits d'implant	149
7.2.2.1. Oscillogrammes	150
7.2.2.2. Mesure du taux d'erreur BER	154
7.2.2.3. L'alimentation par champ HF	154
7.3. Capture des paramètres de stimulation par RTDX	156
Conclusion	158

Conclusion générale	160
----------------------------------	-----

Références	162
-------------------------	-----

Introduction

Une partie appréciable de la population humaine est constituée de gens sourds-muets dits mal-entendants. Les catégories profondément sourdes et totalement sourdes peuvent cependant, bénéficier de la technologie actuelle en matière de réhabilitation de l'audition par le biais d'implants cochléaires. Ces dispositifs microélectroniques ont atteint un stade avancé dans la stimulation électrique des fibres nerveuses de l'oreille interne. Plusieurs stratégies ont été adoptées puis expérimentées sur des milliers de patients à travers le monde.

Bien que les implants cochléaires originaux aient été à canal unique, il y a maintenant plusieurs systèmes d'implant cochléaire multicanal, commercialement disponibles. Ils ont apporté des gains substantiels en reconnaissance de la parole pour la plupart des utilisateurs. Le long des avancés en engineering et conception des processeurs de la parole, des changements ont été apportés dans les critères de candidature pour l'implantation cochléaire. Par exemple, quelques adultes profondément sourds, post-linguaux, étaient initialement considérés comme candidats à l'implantation. Maintenant, l'utilisation d'implant cochléaire multicanal est approuvée pour les enfants sourds pré-linguaux dès l'âge de 12 mois [4, 103].

Nous essayerons dans ce travail d'examiner dans le premier chapitre, les aspects anatomiques et physiologiques de l'oreille humaine, en particulier de l'oreille interne. Nous insisterons sur les exigences physiologiques du nerf auditif dans le cadre de sa stimulation électrique.

Dans le deuxième chapitre, nous tenterons de réunir les principales stratégies utilisées dans la stimulation cochléaire, ainsi que les différents types d'implants commerciaux existants et leurs modèles d'électrodes employés en matière d'implantation cochléaire. Nous aborderons aussi, les critères de candidature et les résultats obtenus en réhabilitation auditive chez les adultes et enfants implantés. L'impact de l'implantation cochléaire sur les enfants sourds est également considéré.

Quant à notre apport, il consiste en une proposition d'une nouvelle stratégie de stimulation cochléaire. Son développement est mis au point à la lumière des constatations physiologiques tirées du premier chapitre et à partir d'une comparaison des différentes solutions adoptées par les implants cochléaires commerciaux, exposées dans ce chapitre.

Dans ce contexte, le troisième chapitre est prévu pour optimiser la conception de l'alimentation radio fréquence transdermale de l'implant. Le chapitre quatre commence par la modélisation de la parole, puis l'utilisation d'un outil d'extraction des paramètres de la parole basé sur le codage autorégressif. Cet outil étant le pilier de notre stratégie proposée, un autre outil de comparaison et validation basé sur le concept des sonagrammes à large bande, a été développé. Le cinquième chapitre, expose les différents algorithmes d'analyse et de traitement en bande de base de la parole, afin de concrétiser notre stratégie de stimulation. Le support matériel étant basé sur un DSP, les possibilités hard et soft de ce dernier sont présentées après avoir argumenté son choix.

Dans un sixième chapitre, nous évoquons la conception et réalisation des circuits de transmission et réception HF, associés à l'implant cochléaire de démonstration. Un dernier chapitre est consacré à l'évaluation des performances obtenues des circuits et logiciels réalisés, ainsi que la capture en temps réel des données mises en jeu, dans le but de les visualiser sur une interface graphique conviviale.

CHAPITRE 1:

Anatomie et Physiologie de l'oreille humaine

Le système auditif humain comprend l'oreille externe, l'oreille moyenne, l'oreille interne et le nerf auditif, ainsi que les centres de traitement des stimuli auditifs situés dans l'encéphale et tronc cérébral. Le système, dans son ensemble, est à la fois d'une extrême complexité et d'une haute précision.

1.1. Oreille externe et oreille moyenne

L'oreille externe [E] et l'oreille moyenne [M], figure 1-1, transmettent le message sonore à l'oreille interne [I]. L'oreille externe allant du pavillon jusqu'au tympan, amplifie les sons aigus entre 1.5 kHz et 7 kHz [20]. Sa fonction est donc une préaccentuation du signal acoustique. L'oreille moyenne assure l'adaptation d'impédance entre le milieu aérien et les liquides de l'oreille interne, ainsi que la protection de cette dernière par un système à 3 osselets.

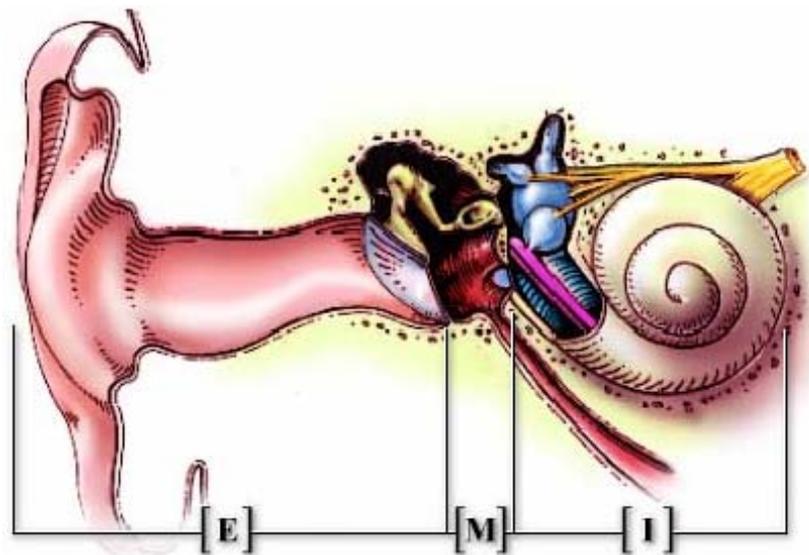


Figure 1-1: Anatomie de l'oreille humaine

1.2. Anatomie de l'oreille interne [15, 37, 58, 86, 100]

1.2.1. Le labyrinthe

L'oreille interne regroupe 2 organes sensoriels distincts, figure 1-1 :

- le vestibule (à canaux semi-circulaires en bleu) : organe de l'équilibration,
- la cochlée (en forme de limaçon) : organe de l'audition.

1.2.2. La cochlée

C'est un tube qui s'enroule sur lui-même en deux tours et demi allant de la base à l'apex, comme une coquille d'escargot, figure 1-2. Le tube a une longueur totale de 30 mm, une hauteur du milieu de la base au sommet de 3 à 4 mm, un diamètre du tube à l'origine de 2 mm et au sommet de 1 mm [58, 100].

Ainsi, les vibrations mobilisent le tympan et la chaîne des osselets, qui transfèrent la vibration au liquide périlymphe de la rampe vestibulaire. L'onde se propage de la base vers l'apex et retourne vers la base par la rampe tympanique. Les deux rampes entourent le canal cochléaire, qui contient l'organe sensoriel auditif appelé organe de Corti, et qui baigne dans le liquide endolymphe. Le nerf auditif se ramifie pour se connecter aux cellules sensorielles de l'organe de Corti.

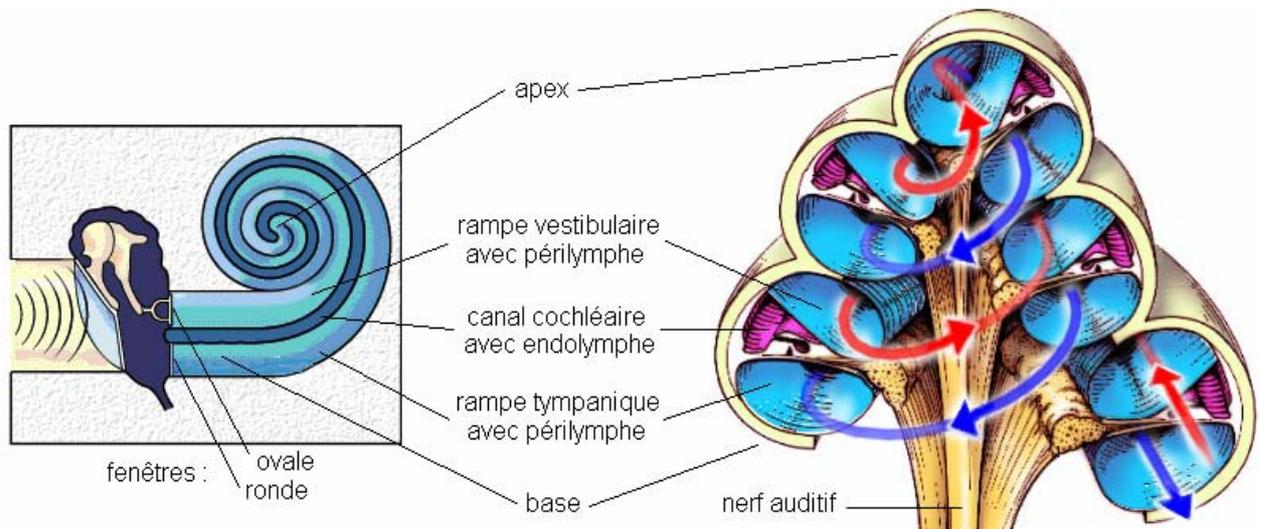


Figure 1-2: Propagation d'onde sonore dans la cochlée

L'organe de Corti repose sur une membrane dite la membrane basilaire, figure 1-3. Elle présente une largeur qui augmente de la *base* à l'*apex*, entraînant des modifications importantes en rigidité et masse réparties [58, 100]. Le lieu de résonance de la membrane varie ainsi en fonction de la fréquence sonore.

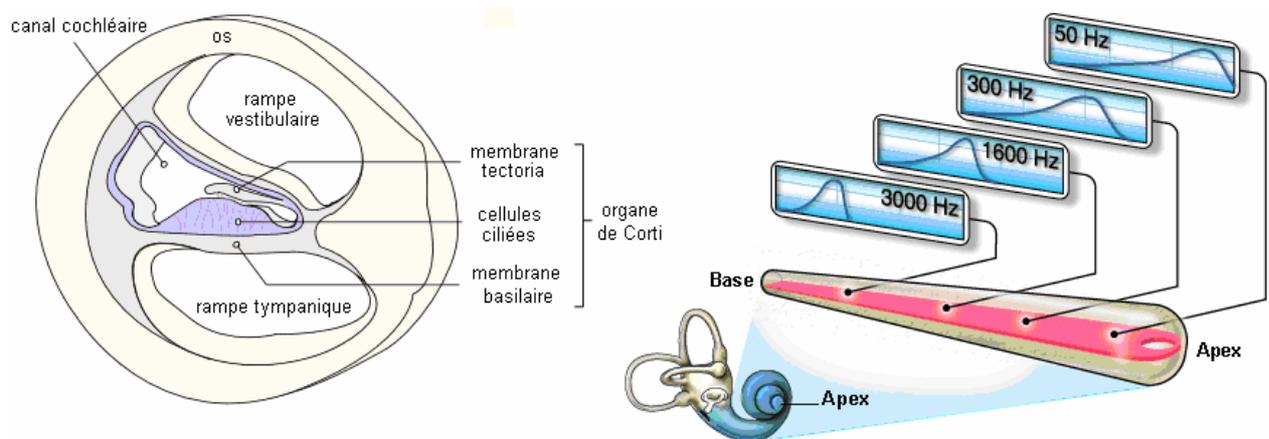


Figure 1-3: Section de la cochlée et membrane basilaire déroulée

1.2.3. Le canal cochléaire

Le canal cochléaire contient l'organe de Corti qui baigne dans l'endolymphe caractérisé par une surcharge en potassium K^+ , figure 1-4. Les rampes vestibulaire et tympanique contiennent la périlymphe, qui se compose de Na^+ et Cl^- . L'électrochimie des liquides de la cochlée, se traduit par un potentiel de +80 mV dans l'endolymphe par rapport à la périlymphe [86].

Les deux liquides ont un rôle physiologique double: le premier est de concourir à l'excitation des cellules ciliées sensorielles par la transmission du signal mécanique, et le second est de participer à la transformation de ce signal en un message nerveux par des échanges chimiques entre les liquides et les cellules ciliées. Les cellules sensorielles, couvertes par la membrane tectoria, sont connectées aux fibres nerveuses.

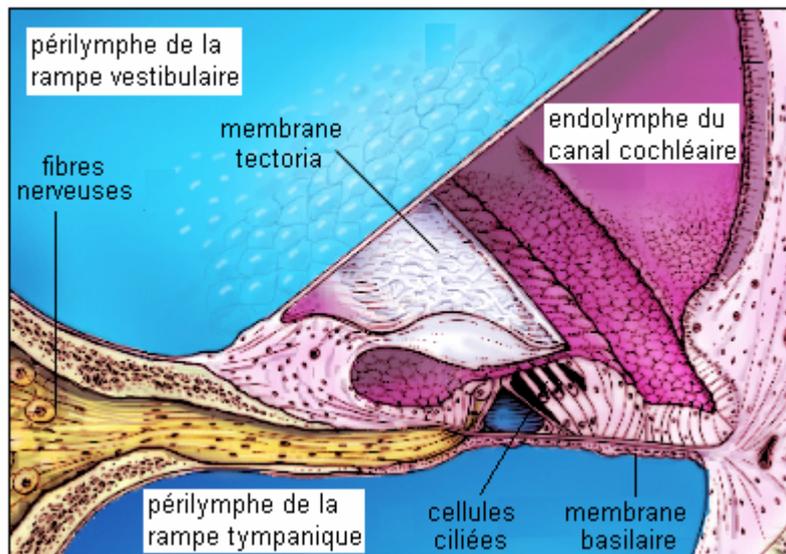


Figure 1-4: Le canal cochléaire

1.2.4. L'organe de Corti

Les cellules sensorielles de l'organe de Corti sont de deux types:

- a- Les cellules ciliées internes (CCI), qui se disposent sur une seule rangée, et présentent des stéréocils qui flote dans l'endolymphe [100].
- b- Les cellules ciliées externes (CCE), qui se disposent en 3 rangées, et présentent des stéréocils qui s'ancrent dans la membrane tectoria [100], figure 1-5.

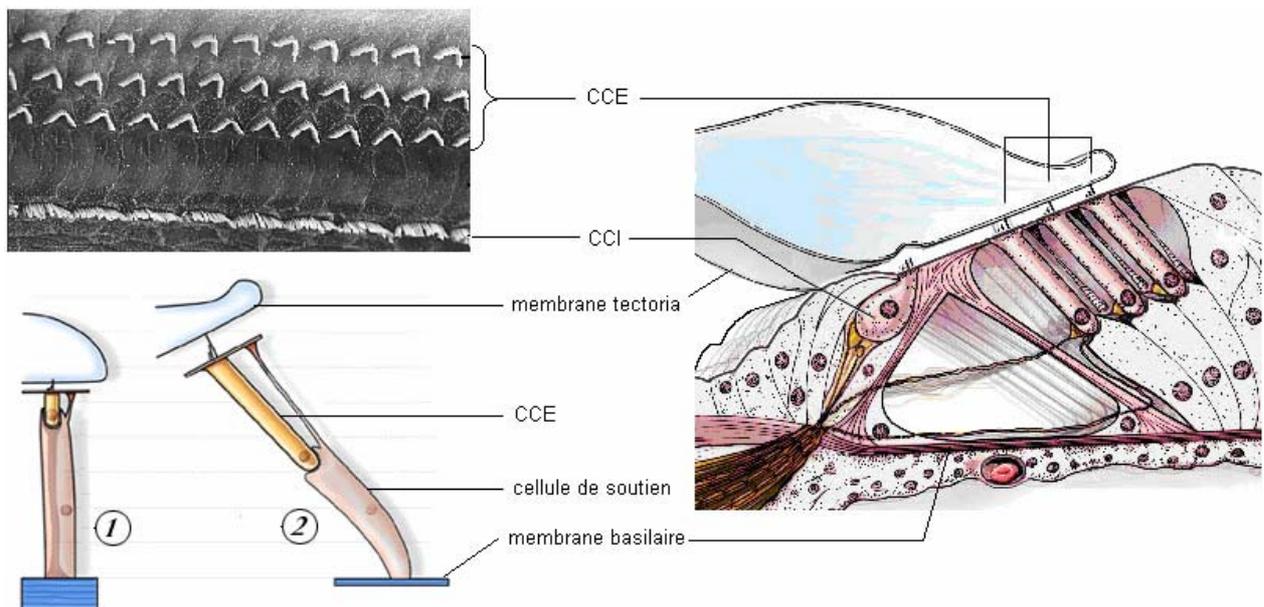


Figure 1-5: Cellules ciliées avec couplage des CCE : (1) ferme à la base, et (2) lâche à l'apex

Le couplage des CCE avec les structures environnantes est plus ferme à la base qu'à l'apex. L'action des CCE est donc plus importante aux fréquences aiguës qu'aux fréquences graves.

1.2.5. Innervation des cellules ciliées

Le nerf auditif débouchant dans la cochlée, contient 30 000 à 35 000 neurones [15]. Il se ramifie en un grand nombre de fibres nerveuses, ordonnées de manière spiralee et se connectant

aux cellules ciliées. Deux types de fibres nerveuses sont recensés.

1.2.5.1. Les fibres nerveuses afférentes

Elles transmettent le message sonore depuis les cellules sensorielles jusqu'aux cellules nerveuses du nerf auditif. En fonction de leurs terminaisons, on distingue :

- Les fibres afférentes des CCI: Au nombre d'environ 45 000 [37, 100], elles représentent 90 à 95% des fibres du nerf cochléaire. Chaque CCI recevant une vingtaine de fibres nerveuses.
- Les fibres afférentes des CCE: Au nombre de 3000 à 5000 [37, 100], elles innervent chacune une dizaine de CCE, figure 1-6.

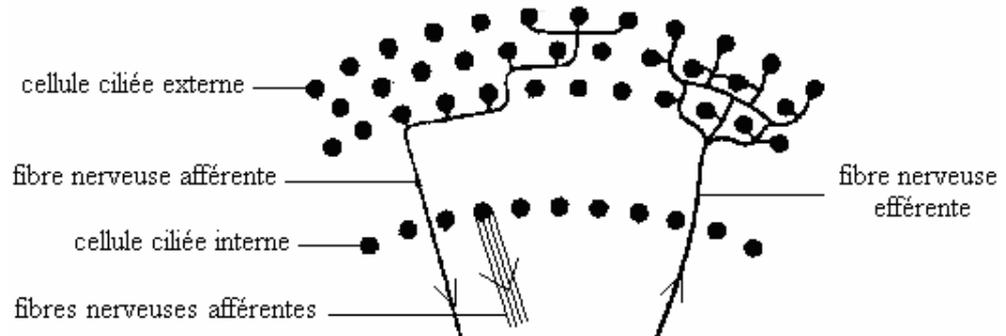


Figure 1-6: Terminaisons des fibres nerveuses

Ainsi, le système d'innervation est divergent pour les CCI et convergent pour les CCE, et ces modalités d'innervation expliquent les hautes possibilités de discrimination tonale des CCI. Les fibres nerveuses possèdent chacune une sensibilité particulière à certaines fréquences, et qu'à cette spécificité correspond une tonotopie fréquentielle dans laquelle, les sons graves se trouvent au niveau de l'apex, alors que les sons aigus sont au contraire situés à la base de la cochlée. Cette sensibilité fréquentielle élective des fibres nerveuses afférentes est une notion anatomophysiologique essentielle sur laquelle repose la technique de réhabilitation des surdités totales par l'implantation d'électrodes multiples.

1.2.5.2. Les fibres nerveuses efférentes

Elles sont fonctionnellement très importantes puisque par leur intermédiaire les centres auditifs régularisent la fonction cochléaire, figure 1.6. On dénombre 40 000 terminaisons efférentes environ [37, 100], distribuées aux CCE.

1.3. Physiologie de l'oreille interne [20, 38, 86]

1.3.1. Comportement mécanique et simulation de fonctionnement

Les vibrations sonores sont transmises à la cochlée par la chaîne ossiculaire. Une onde de pression se propage dans la périlymphe des rampes vestibulaire puis tympanique, faisant mouvoir la membrane basilaire de la base vers l'apex. Les mouvements de la membrane basilaire, support de l'organe de Corti, sont à la base de l'excitation des cellules sensorielles. L'amplitude de l'onde propagée atteint un maximum en un point de la membrane basilaire dont la position dépend de la fréquence du son. Ce point se déplace de l'apex vers la base au fur et à mesure que l'on augmente la fréquence du son [20]. Ce système de filtrage réalise une répartition topographique des fréquences le long de la membrane basilaire avec les fréquences aiguës à la base et les fréquences graves à l'apex. Dans le cas d'un son complexe, le filtre basilaire en réalise une analyse spectrale.

La membrane basilaire, siège du traitement des vibrations sonores et leur conversion en décharges électriques, se trouve le long de la cochlée spiralée (environ 34 mm). La loi de Hartmann (1996) [38], permet de tirer la distance cochléaire D à partir de l'apex, correspondante à la fréquence de résonance f :

$$D \text{ (mm)} = 7.24 \ln [1 + f \text{ (Hz)} / 165] \quad 1.1$$

La même distance mesurée à partir de la base devient : $L \text{ (mm)} = 34 - D \quad 1.2$

D'autre part, l'impédance acoustique d'un système vibratoire est donnée par [19] :

$$Z(f, x) = \sqrt{R^2 + (M(x) \cdot f - \frac{S(x)}{f})^2} \quad 1.3$$

où $x = \frac{L \text{ (mm)}}{34} 99 \quad 1.4$

Avec : M(x) et S(x), masse et rigidité réparties dépendant de l'endroit sur la membrane, f fréquence, x distance d'éloignement normalisée par rapport à la base de la cochlée, (x = 0 à 99 pour L = 0 à 34 mm), et les frottements R=40 g/s.mm² [20].

On suppose que les lois d'augmentation de la masse et de diminution de rigidité, sont linéaires en fonction de la distance x, d'où :

$$M(x) = M_b (a x + 1) \quad 1.5$$

et $S(x) = S_a (100 - x) \quad 1.6$

Tel que M_b est la masse à l'origine (ou base), S_a est la rigidité à l'extrémité (ou apex), a étant un coefficient (pouvant dépendre de x pour une variation non linéaire de la masse).

Des facteurs de résonance peuvent être définis à la fréquence de résonance f₀ [20]:

$$Q = M f_0 / R = S / R f_0 \quad 1.7$$

La bande passante étant donnée par : $B = f_0 / Q \quad 1.8$

Dans [20], on peut relever les bandes passantes suivantes :

$$B_1 = 2 \text{ kHz} \text{ autour de la fréquence } f_1 = 10.5 \text{ kHz, d'où } Q_1 = 5.25.$$

$$B_2 = 0.4 \text{ kHz} \text{ autour de la fréquence } f_2 = 1.4 \text{ kHz, d'où } Q_2 = 3.25.$$

Ainsi selon (1.7): $Q_1/Q_2 = M_1 f_1 / M_2 f_2 = S_1 f_2 / S_2 f_1 = 1.6$

Ce qui donne: $M_1/M_2 = 0.21$ et $S_1/S_2 = 12$.

Les distances à partir de la base sont tirées de (1.1) et (1.2) :

$$L_1 = 3.8 \text{ mm (} x_1 = 11) \text{ pour } f_1 = 10.5 \text{ kHz}$$

$$L_2 = 18.2 \text{ mm (} x_2 = 53) \text{ pour } f_2 = 1.4 \text{ kHz.}$$

Le rapport $M_1/M_2 = (a x_1 + 1) / (a x_2 + 1) = 0.21$ donne $a = 6.1$.

D'où: $M_b = R Q_1 / f_1 (6.1 x_1 + 1) = R Q_2 / f_2 (6.1 x_2 + 1) = 0.3 \text{ mg/mm}^2$

A partir de l'expression (1.3), on tire à la résonance : $f^2 = S / M \quad 1.9$

Ainsi pour x=0, on tire d'après (1.1), (1.5) et (1.6) : $f = 17.9 \text{ kHz, } M = M_b, \text{ et } S = 100 S_a.$

L'expression (1.9) permet de calculer le rapport : $S_a/M_b = 3204100 \text{ s}^{-2}$, D'où $S_a = 961.23 \text{ g/s}^2 \text{ mm}^2$.

Dans un milieu liquide ou solide, l'onde acoustique est longitudinale [27]. L'équation d'onde acoustique progressive, se propageant le long de la membrane basilaire, est donnée par [27] :

$$\frac{d^2 e}{dx^2} = \frac{1}{v^2} \frac{d^2 e}{dt^2} = \frac{\rho^2}{Z^2(x)} \frac{d^2 e}{dt^2} \quad 1.10$$

Avec: e amplitude de l'onde, v vitesse de propagation de l'onde, ρ masse volumique de la membrane, Z(x) impédance acoustique, et x longueur de la membrane.

Une solution de cette équation est donnée par:

$$e = e_0 \sin(kx) \sin(\omega t) \quad 1.11$$

Où : $k = 2\pi / \lambda = \omega / v$, tel que ω est la fréquence angulaire, et λ, la longueur d'onde.

La partie spatiale de la solution, représente la sélectivité de la membrane basilaire :

$$\sin(kx) = \sin(\omega x/z(x)) \quad 1.12$$

L'amplitude relative de la pression sonore, présente le long de la cochlée, est alors :

$$E(\omega, x) = \frac{e}{e_0} = \sin(\omega t) \sin \left(\frac{\omega x \rho/R}{\sqrt{1 + \left[\frac{M_b}{R} (ax + 1) \frac{\omega}{2\pi} - \frac{S_a}{R} (100 - x) \frac{2\pi}{\omega} \right]^2}} \right) \quad 1.13$$

En prenant la densité de la membrane $\rho = 10 \text{ g/m}^3$ (dix fois celle de l'air 1.2 g/m^3), nous avons tracé l'amplitude relative de la pression sonore en fonction du lieu sur la membrane basilaire, pour une fréquence acoustique donnée.

A 100 Hz, figure 1-7, le maximum d'amplitude de l'onde propagée est atteint dans la région apicale de la membrane. Comme en apex, la masse de la membrane est maximale et la rigidité minimale, la transmission des aigus est limitée et le point de résonance est abaissé.

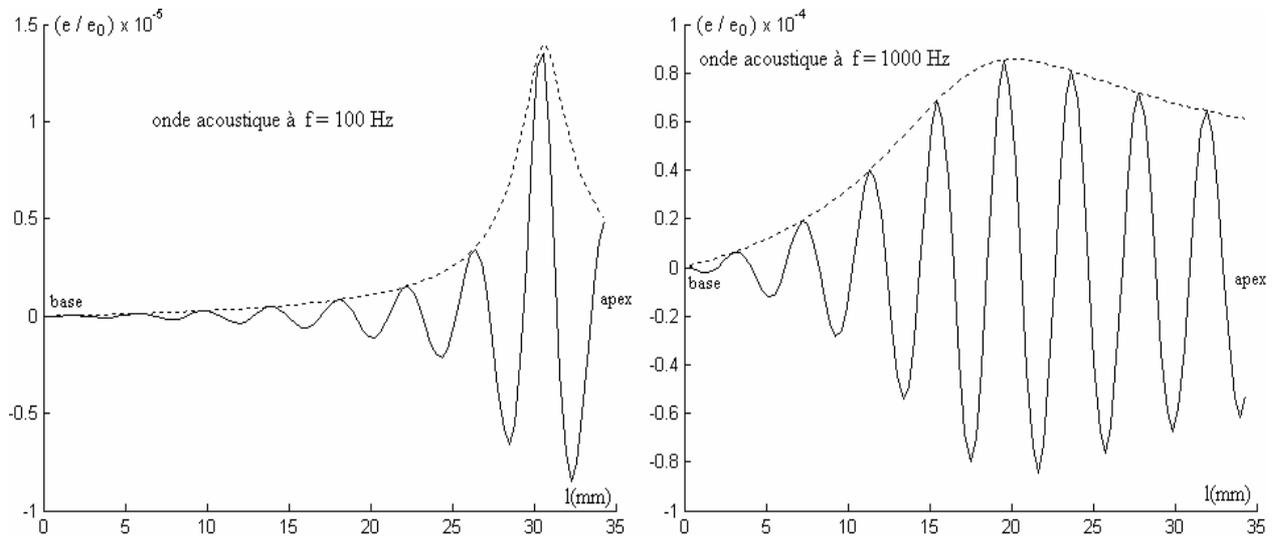


Figure 1-7: Onde acoustique se propageant sur la membrane basilaire avec sélectivité centrée sur 100 puis 1000 Hz

A 1000 Hz, la résonance s'effectue en un lieu de la membrane où la masse diminue, tandis que la rigidité augmente. La transmission des fréquences graves est diminuée et le point de résonance glisse vers les aigus. Cet aspect montre la tonographie de la cochlée.

Dans cette simulation, un fait contradictoire avec la réalité est constaté. La sélectivité fréquentielle de la membrane est normalement plus prononcée pour les fréquences aigues que les fréquences graves. Une solution éventuelle serait de prendre pour la masse répartie, un coefficient a (1.5) non constant mais dépendant du lieu x sur la membrane basilaire.

Des résultats récents [20] ont montré que la sélectivité mécanique de la membrane basilaire est identique à celle des CCI et leurs fibres nerveuses, figure 1-8, où l'intensité mesurée en dB SPL (Sound Pressure Level) est tracée en fonction de la fréquence acoustique.

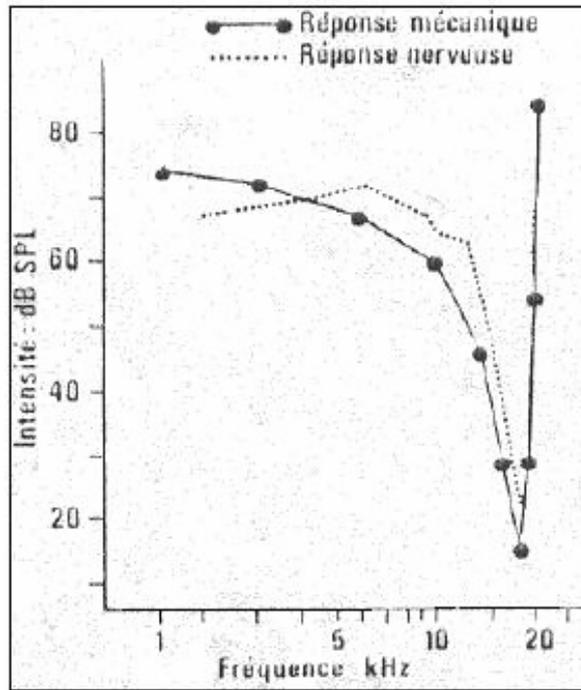


Figure 1-8: Sélectivité mécanique et électrique en réponse à une fréquence de 18 kHz

1.3.2. L'amplificateur cochléaire

Les 95% des fibres nerveuses afférentes sont connectées aux CCI, figure 1.6, qui constituent ainsi le seul véritable récepteur sensoriel de la cochlée.

Connectées aux fibres nerveuses efférentes, les CCE constituent par contre, une chaîne de retour par laquelle une rétroaction est assurée. Le filtre actif constitué par les CCE, amplifie de 50 dB la vibration de la membrane basilaire sur une portion très étroite, assurant ainsi une meilleure sélectivité en fréquence, figure 1.9. Cet accord en fréquence dépend de la contraction des CCE, qui est plus importante aux fréquences aiguës qu'aux fréquences graves, et qui se retrouve transmise aux CCI par l'intermédiaire de la membrane tectoria [20], figure 1-5.

D'où le principe de la théorie spatiale ou du lieu (place theory), postulant que la fréquence acoustique est perçue suivant l'endroit ou site d'excitation des cellules ciliées. Une bonne sélectivité étant attribuée aux fréquences supérieures à 1 kHz, le codage est essentiellement tonotopique pour les fréquences aiguës.

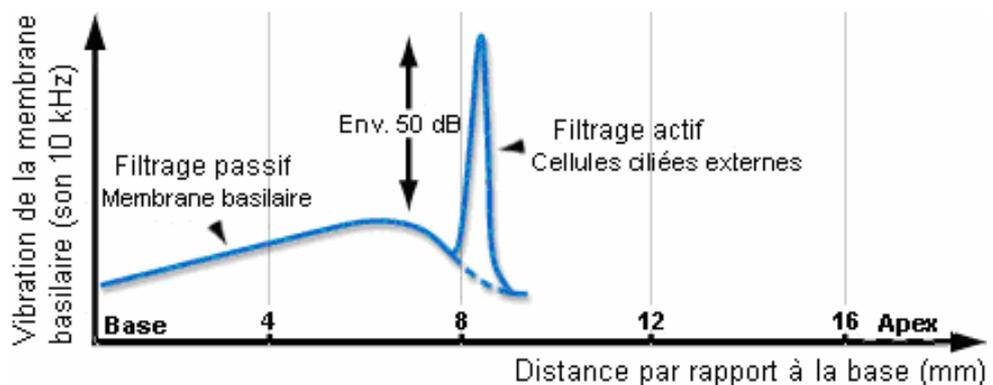


Figure 1-9: L'amplificateur cochléaire

1.3.3. Le potentiel de récepteur des CCI

L'activité électrique des CCI pendant la stimulation sonore a pu être enregistrée au moyen de microélectrodes intracellulaires (Russel et Sellick, 1983) [20]. Les potentiels unitaires enregistrés, appelés potentiels de récepteur, sont constitués d'une composante continue reproduisant l'enveloppe de la stimulation acoustique, et d'une composante alternative superposée qui suit la fréquence du son jusqu'à 1 kHz au maximum, figure 1-10.

Au-delà de 1 kHz, les caractéristiques de filtrage de la membrane cellulaire déterminent une atténuation progressive de la composante alternative, ne laissant subsister à partir de 5 kHz que l'enveloppe de la stimulation. De même, la période réfractaire des fibres nerveuses associées ne leur permet pas de suivre une fréquence supérieure à 1 kHz [20].

D'où le principe de la théorie temporelle (volley theory), postulant que la fréquence acoustique est perçue par des potentiels de récepteur synchrones des cellules ciliées. La fréquence de 1 kHz étant une limite supérieure, le codage est essentiellement temporel pour les fréquences graves.

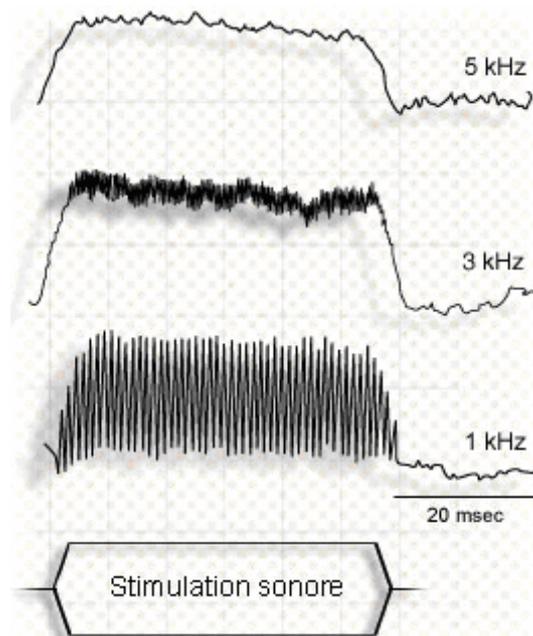


Figure 1-10: Potentiel de récepteur dans les CCI

1.3.4. Schéma fonctionnel d'une oreille saine

Une récapitulation des principales données anatomiques et physiologiques précédentes, nous permet d'en déduire le schéma fonctionnel d'une oreille saine, figure 1-11.

L'oreille externe assure une préaccentuation du son, tandis que l'oreille moyenne assure l'adaptation d'impédance entre le milieu aérien et la cochlée.

Les vibrations sonores sont transmises aux liquides de l'oreille interne, où la membrane basilaire assure une analyse spectrale du son. L'action des CCE assure une amplification sélective des constituants fréquentiels du son, dont les CCI accomplissent une transduction mécano-électrique.

La forte différence de potentiel existant entre le corps cellulaire des CCI (-60 mV) et l'espace endolymphatique (+80 mV) est à l'origine d'un courant transmembranaire permanent. La déformation des stéréocils entraînée par l'onde propagée dans l'endolymphe entraîne une modification de perméabilité membranaire des CCI. Il en résulte une modification du courant transmembranaire engendrant la génération de potentiels de récepteur. L'activation de la CCI s'accompagne de la libération des neurotransmetteurs (médiateurs chimiques) [20], qui

permettant la dépolarisation des fibres nerveuses afférentes. Les fibres ainsi activées vont donner naissance à des potentiels générateurs destinés aux neurones du nerf auditif. Ces dernières réagissent par des décharges électriques, dont la distribution temporelle et spatiale constitue un codage du son, en destination du centre auditif [20].

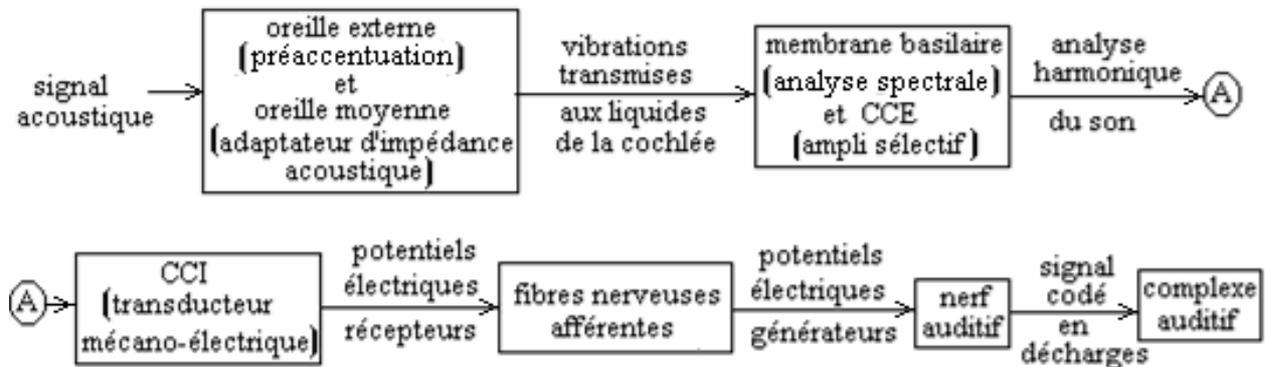


Figure 1-11: Schéma fonctionnel d'une oreille

1.4. Psychophysique de la stimulation électrique du nerf auditif [22, 23]

La sécurité des paramètres du stimuli électrique est une exigence physiologique du nerf, et elle est assurée sous les aspects suivants.

1.4.1. Ondes rectangulaires ou sinusoïdales

La stimulation électrique d'une fibre nerveuse est d'autant plus efficace, c'est-à-dire économe en énergie, qu'elle est brutale avec aspect à fronts raides.

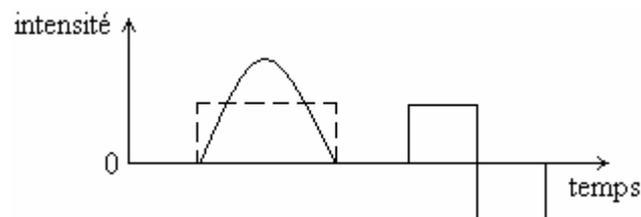


Figure 1-12: Impulsions de stimulation : rectangulaire, sinusoïdale et biphasique

Pour une même quantité d'électricité dans l'onde rectangulaire et l'onde sinusoïdale (mêmes surfaces et durées), figure 1-12, l'intensité de l'onde rectangulaire est plus faible. Ainsi, une stimulation par onde rectangulaire (ou stimulation digitale) demandera pour être efficace une intensité moindre qu'une stimulation par onde sinusoïdale (ou stimulation analogique). C'est pourquoi la stimulation nerveuse est très généralement réalisée par des ondes rectangulaires.

Des études neurophysiologiques ont montré que les impulsions doivent être balancés en charge pour ne pas détériorer les fibres nerveuses. Ainsi la polarité, généralement positive, de la stimulation rectangulaire doit être suivie d'un apport de courant de polarité inverse mais quantitativement équivalent pour éviter les phénomènes de polarisation électrolytique. Certains implants cochléaires actuels stimulent le nerf auditif par l'intermédiaire d'impulsions biphasiques modulés en amplitude, figure 1-12.

Pour exciter le nerf auditif, la largeur d'impulsion ne peut être plus courte que 0.1 ms [22, 23]. D'autre part, la corrosion de l'électrode (en platine) croît avec la densité de courant et la durée d'impulsion. C'est pour ces raisons, qu'un compromis utilise des durées d'impulsion de 180 μ s/phase à forme biphasique [22, 23].

1.4.2. Stimulation électrique de la fréquence acoustique

Dans la gamme de fréquences 60-300 Hz, théoriquement seront seuls perçus le rythme de la parole (pitch) et ses variations de hauteur. En pratique, cette discrimination est souvent possible jusqu'à 1000 Hz, notamment lorsque le pourcentage des fibres nerveuses auditives valides est élevé. Au-delà, le codage de la fréquence dépend plutôt de la position de l'électrode stimulante sur le clavier cochléaire : placée vers l'apex, cette électrode donnera un son grave même si la fréquence de stimulation est très élevée; placée vers la base, elle donnera au contraire un son aigu.

1.4.3. Stimulation électrique de l'intensité acoustique

La fibre nerveuse est sensible à la quantité d'énergie électrique délivrée par seconde. La plupart des implants travaillent avec des sources de courant pour pallier les variations dans le temps de l'impédance des électrodes. Ainsi, à impédance donnée des tissus biologiques, la quantité d'énergie dépend de l'intensité du courant et de la durée de son application. C'est pourquoi, la plupart des appareillages peuvent faire varier la durée en plus de l'intensité, pour augmenter les possibilités de discrimination de l'intensité sonore [22, 23]. Une rangée de courant d'environ 1 mA, est sélectionnée pour s'accommoder avec les différences présentées par les patients dues aux variations dans la position de l'électrode et la population des fibres nerveuses résiduelles.

1.4.4. Dynamique de la stimulation électrique

Dans une stimulation électrique du nerf cochléaire, il y a une très faible différence entre la quantité d'énergie nécessaire pour avoir une sensation sonore minimale et celle qui donne une sensation à peine douloureuse. Exprimé en décibels, ce rapport d'énergie du seuil minimum au seuil maximum de sensation supportable est habituellement proche de 6 dB, c'est-à-dire très faible par rapport aux 100 dB de la stimulation sonore d'une oreille saine [23, 61]. On mesure la grande efficacité des structures de l'oreille et la stimulation mécanique des liquides cochléaires, que l'implant doit court-circuiter parce qu'elles sont devenues inefficaces. Dans un implant cochléaire, il devient donc indispensable de comprimer l'énergie électrique reflétant le message sonore à la sortie du microphone, de façon à ce que les fluctuations d'intensité soient contenues dans une plage étroite proche de 6 dB. On appelle cette plage la "dynamique de la stimulation".

1.5. La surdité [4]

Il existe deux types de surdité.

1.5.1. Surdité de transmission

La surdité de transmission intervient lorsque le tambour et les os de l'oreille moyenne ne transmettent plus de vibrations sonores à l'oreille interne. Cependant, la cochlée et le nerf auditif peuvent fonctionner normalement.

1.5.2. Surdité neurosensorielle

La surdité de perception est provoquée par un dysfonctionnement de la cochlée (oreille interne), pouvant être le signe de dommages au niveau des fines cellules auditives le long de la cochlée. Ces cellules auditives endommagées ne peuvent pas envoyer de signaux électriques au nerf auditif, par conséquent le cerveau ne reçoit pas les informations sonores complètes. Plus les cellules auditives endommagées sont nombreuses, plus la perte d'audition est susceptible d'être sévère. Les surdités neurosensorielles peuvent entraîner une altération faible à profonde de l'audition. La perte neurosensorielle peut être provoquée par de nombreux facteurs notamment la génétique, une blessure, une maladie ou encore le processus de vieillissement naturel.

Le mécanisme de la surdité s'exprime clairement par l'audiogramme. Cet examen consiste à faire entendre un son de fréquence pure avec une intensité de plus en plus forte, pour apprécier à partir de quelle intensité le patient sourd commence à percevoir cette fréquence. Après la mesure du seuil auditif "liminaire", on apprécie pour cette même fréquence le seuil douloureux ou inconfortable, qui chez le sujet normal avoisine les 100 dB. La même opération est effectuée pour l'éventail des sept principales fréquences (de 125 Hz à 8000 Hz) du spectre auditif de l'homme, et réalisée successivement pour les deux oreilles. La distinction des surdités sévères, profondes ou totales, est clinique.

1.5.2.1. Les surdités totales

Elles sont caractérisées par une inefficacité totale des prothèses amplificatrices conventionnelles, qui n'apportent aucune information sonore globale, et ne fournissent aucune aide à la lecture labiale. Un audiogramme "blanc" est donc obtenu, ne comportant aucun "restes" auditifs, quelles que soit la fréquence et l'intensité étudiée. Elles sont dues à une destruction totale de l'organe de Corti, associée souvent, à une atteinte plus ou moins importante du nerf auditif. Ces surdités totales sont accessibles à l'implantation cochléaire, pourvu que les fibres du nerf cochléaire soient encore suffisamment nombreuses pour pouvoir transmettre l'information sonore, transformée en signaux électriques par l'implant, jusqu'aux centres cérébraux. Ceci est possible même si l'organe de Corti est complètement détruit.

1.5.2.2. Les surdités profondes

Elles sont caractérisées par une inefficacité pratique des prothèses amplificatrices conventionnelles, qui n'apportent qu'une information sonore globale, représentant toutefois une aide à la lecture labiale. Elles sont dues à une atteinte grave, mais incomplète de l'organe de Corti ou du nerf auditif. Le rôle de l'implant sera de remplacer, au moins partiellement, l'organe de Corti, dont la fonction n'est pas seulement d'effectuer une transduction d'énergie, mais aussi de réaliser un codage du message acoustique.

1.5.2.3. Les surdités sévères

Quel que soit leur mécanisme pathologique, les surdités sévères se définissent actuellement par l'existence de restes auditifs sur toutes les fréquences, et par une efficacité certaine mais limitée des prothèses conventionnelles. Mais l'intelligibilité de la parole sans lecture labiale de ces patients est tellement mauvaise, que l'on s'est aperçu vers l'année quatre-vingt quinze, que l'implant cochléaire pouvait faire beaucoup mieux que leur organe de Corti, trop endommagé pour être assez efficace.

Conclusion

L'oreille interne comporte la cochlée : un tube creux enroulé sur lui-même. Le nerf cochléaire s'insère en ramification dans ce tube. A l'origine de cette insertion se trouve un élément fondamental: l'organe de Corti. C'est lui qui permet la transduction du message sonore, et aussi le codage de ce message. En effet, l'organe de Corti, situé tout le long de la cochlée, a une activité spécifique en fonction de la fréquence qui est détectée. Les fréquences aiguës étant prises en charge par les organes à la base de la cochlée, et les graves au sommet de la cochlée. Une implantation cochléaire est prévue pour des surdités totales, profondes, voire sévères, à la condition qu'une certaine population de fibres nerveuses subsiste encore. Une stimulation électrique doit respecter les propriétés des neurones cochléaires. Pour cela, elle doit utiliser des signaux électriques symétriques par rapport à la masse, avec une dynamique et un rythme n'excédant pas respectivement, 7dB et 1 kHz.

CHAPITRE 2:

implants cochléaires commerciaux

2.1. Candidats à l'implant cochléaire :

Un audiogramme est utilisé pour différencier les divers degrés de surdité. L'implant cochléaire est indiqué dans les surdités post-linguales profondes (perte d'audition de 90 dB ou plus) [23, 49] voire totales avec un résidu encore intact de fibres nerveuses, et les surdités sévères lorsqu'il existe une intelligibilité sans lecture labiale, inférieure à 30 % avec prothèses conventionnelles. La surdité doit être bilatérale (dans les deux oreilles). Le bilan préopératoire s'attache aussi à rechercher les causes de la surdité (traumatique, infectieuse, toxique, ou héréditaire) et s'assure de la faisabilité de l'implantation en vérifiant que la cochlée n'est pas ossifiée, etc. L'intervention en elle-même est relativement simple. Elle se fait sous microscope opératoire. Pour mettre en place la gaine d'électrode à l'intérieur de la cochlée, on pratique au niveau de l'oreille interne une petite ouverture d'insertion. La gaine d'électrodes s'enroule dans le tube cochléaire. La prise en charge opératoire est triple : médicale, ajustage, et orthophonique. Dès la mise en place de l'implant après le premier réglage au 1^{er} mois, le patient perçoit son environnement sonore, et progressivement, il va devoir le décoder. C'est le rôle essentiel de l'orthophoniste qui accompagne alors le malade.

Dans le réglage des seuils, on demande au patient de déterminer à quel moment il commence à avoir une sensation auditive, et à quel moment la sensation est confortable à légèrement forte. On fait ce réglage pour toutes les électrodes [38, 56, 61].

Les tests couvrent aussi l'aspect du décalage en fréquence dû à l'insertion d'électrode. Ceci revient à ce que les électrodes soient insérées partiellement dans la cochlée, typiquement 22 à 30 mm, dépendant de l'état de la cochlée [111, 113]. Le fait que les électrodes ne soient pas pleinement insérées dans la cochlée, une erreur se produit en correspondance entre la fréquence d'analyse et la fréquence de stimulation. Comme résultat, le signal codé de la parole est décalé en fréquence et devient moins intelligible. Ainsi, le nombre de sites disponibles à la stimulation dépend de la profondeur d'insertion de la multi-électrode à l'instant de l'implantation. En plus et dans certains cas, la location physique des électrodes à l'intérieur de la cochlée peut ne pas être à réponse maximale à une stimulation électrique, du moment que la densité et l'uniformité de la population neuronale peuvent varier avec le type de pathologie ayant causé la surdité. Ces deux faits réduisent la flexibilité de la programmation et peuvent limiter les performances chez certains individus.

L'identification de la position d'électrode dans la cochlée est faite de plusieurs manières :

Dans une étude [22, 23], les variations du pitch avec la position d'électrode sont évaluées en utilisant une procédure de comparaison par paire d'électrodes. Les patients sont interrogés si la sensation à une électrode est plus aiguë ou plus grave que la sensation de l'autre électrode.

Après ajustement du processeur de traitement de la parole avec la collection recueillie des caractéristiques psychophysiques comparatives, un programme de rééducation auditive et réhabilitation du patient, est entrepris.

Les capacités de perception du langage chez le patient, sont typiquement évaluées par des tests incluant des phrases, des mots monosyllabiques, des voyelles et des consonnes.

Les résultats sont très satisfaisants lorsque le patient peut parler au téléphone. Ils sont considérés satisfaisants quand le patient peut soutenir une conversation sans l'aide de la lecture labiale. Alors qu'ils sont jugés médiocres dans le cas où le patient doit utiliser la lecture labiale pour une bonne compréhension.

2.2. Caractéristiques d'implant :

Plusieurs types d'implant cochléaire ont été développés. Ces dispositifs diffèrent par les caractéristiques suivantes.

2.2.1. Conception d'électrode

Certains aspects associés à la conception d'électrode, sont :

- l'emplacement des électrodes,
- l'orientation des électrodes par rapport au tissu excitable,
- le nombre d'électrodes et espacement des contacts,
- la configuration d'électrode.

Les électrodes peuvent être placées sur la surface de la cochlée (extra cochléaire) mais demandent alors des niveaux élevés de courant de stimulation. En pratique, les électrodes sont souvent placées dans la rampe tympanique (intra cochléaire) à travers une incision dans la membrane de la fenêtre ronde au niveau de la base cochléaire [22, 23]. Cela permet un contact direct et une localisation plus facile des groupes de fibres nerveuses à caractéristiques fréquentielles distinctes. L'insertion de l'électrode cause un minimum de traumatisme si celle-ci est flexible. Le nombre d'électrodes et leur espacement, affectent la résolution en site dans la cochlée du codage fréquentiel.

Cependant, deux facteurs inhérents aux prothèses cochléaires contraignent le codage fréquentiel :

- 1- Le nombre de neurones auditives résiduelles qui peuvent être stimulées sur un site particulier dans la cochlée.
- 2- La diffusion de l'excitation associée à la stimulation électrique.

Le premier problème dépend du type (étiologie) de la surdité. Ainsi, si le nombre de neurones résiduelles est réduit à une petite zone dans la cochlée, quelques électrodes implantées près de cette zone seront meilleures qu'une centaine d'électrodes distribuées le long de la cochlée.

Pour cela, l'utilisation d'un grand nombre d'électrodes ne conduit pas nécessairement à une meilleure performance, du fait que le codage fréquentiel est influencé par le nombre résiduel des neurones auditives stimulables. En outre, le codage fréquentiel est aussi influencé par la diffusion de l'excitation causée par la stimulation électrique. Le courant électrique injecté dans la cochlée, ne stimule pas un unique site de neurones auditives mais plusieurs. Une telle diffusion dans l'excitation est plus prononcée dans la configuration monopolaire des électrodes. Dans cette configuration, l'électrode active est située au voisinage des fibres nerveuses alors que l'électrode de masse commune à toutes les électrodes, est placée à distance, généralement à l'une des extrémités de la gaine d'électrodes, figure 2-1.

La diffusion de l'excitation due à la stimulation électrique, peut être limitée par l'utilisation de la configuration bipolaire des électrodes. Dans cette configuration, l'électrode active et sa référence placées au niveau des fibres nerveuses, sont proches l'une de l'autre, figure 2-1. La paire d'électrodes stimulée peut être séparée par n électrodes non actives (bipolaire BP+n).

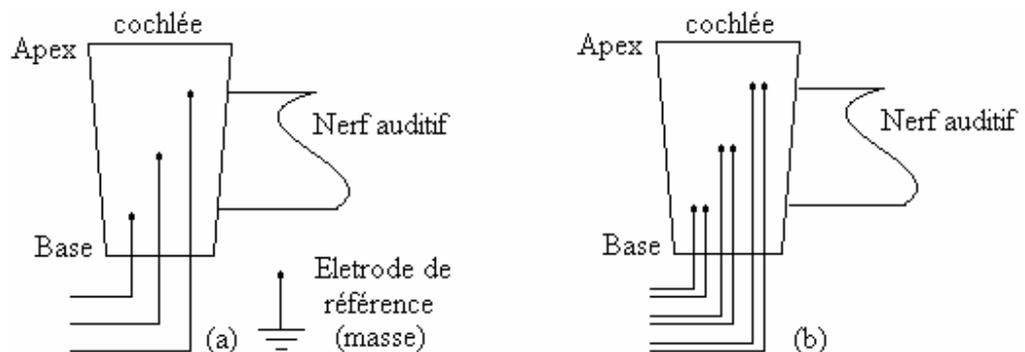


Figure 2-1 : Configurations d'électrodes : (a) monopolaire et (b) bipolaire

La propagation du courant dans la rampe tympanique a été étudiée en utilisant des électrodes bipolaires et monopolaires, figure 2-2 [22, 23, 61]. La stimulation bipolaire présente une atténuation de courant de 3 à 4 dB/mm autour de l'emplacement de l'électrode; tandis que celle

de la stimulation monopolaire est moins de 1dB/mm, présentant ainsi une courbe moins sélective. Les électrodes bipolaires produisent une stimulation plus localisée que les électrodes monopolaires [23, 104].

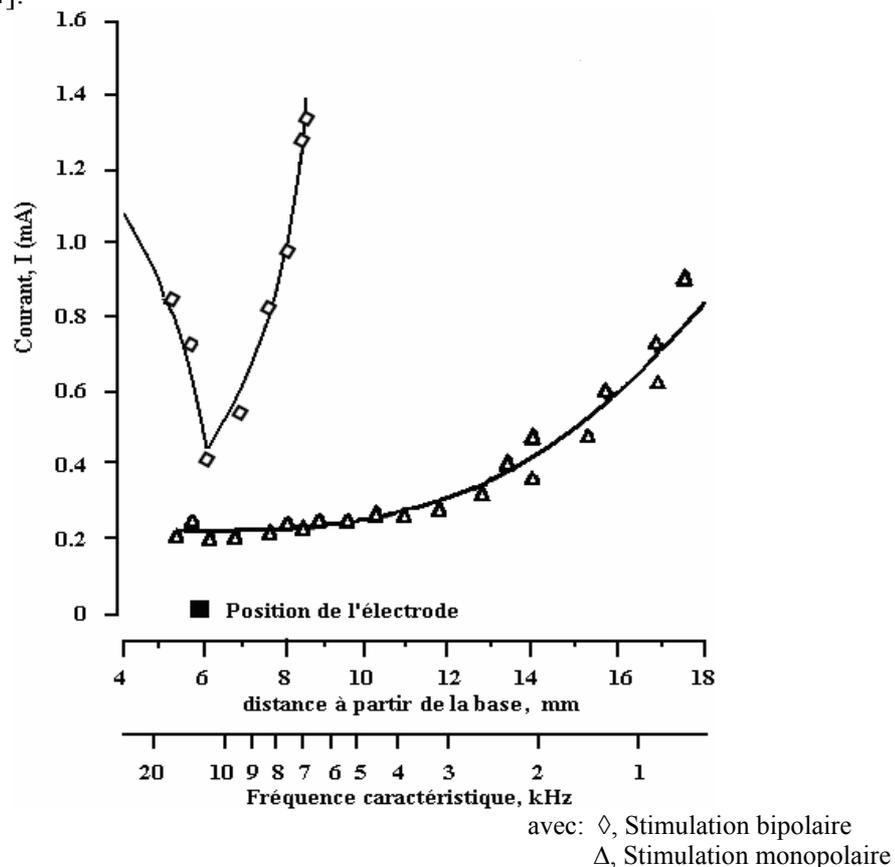


Figure 2-2: Courant seuil de stimulation en fonction de la tonotopie cochléaire

En revanche, la stimulation monopolaire nécessite moins de courant pour obtenir une même sensation sonore que la stimulation bipolaire. Cependant pour un patient particulier, rien ne peut être affirmé en terme de performance obtenue vis-à-vis des configurations d'électrodes employées.

En pratique, certains implants utilisent des électrodes monopolaires, d'autres emploient des électrodes bipolaires, alors que d'autres fournissent les deux, laissant au clinicien le choix entre ces deux types de stimulation.

La table 2.1, montre une liste d'implants cochléaires courants et leurs caractéristiques.

Implant	Electrodes			Type de stimulation	Lien de Transmission
	Nombre	Espacement	Configuration		
Ineraid	6	4 mm	Monopolaire	Analogique	Percutané
Nucleus	22	0.75 mm	Bipolaire	Impulsionnelle	Transcutané
Clarion	16	2 mm	Monop./Bipol.	Analog./Impul.	Transcutané
Med-El	12	2.8 mm	Monopolaire	Impulsionnelle	Transcutané

Table 2.1: Caractéristiques des implants cochléaires du commerce

Ineraid (appelé aussi Symbion) utilise six électrodes espacées de 4 mm. Seulement quatre électrodes, les plus apicales, sont utilisées en configuration monopolaire. Nucleus utilise 22 électrodes espacées de 0.75 mm. Les électrodes espacées de 1.5 mm, sont utilisées comme des paires bipolaires (dit mode BP+1). L'électrode multiple est faite de fils fins en platine de

diamètre 0.025 mm et connectés individuellement aux 22 contacts métalliques, l'isolation étant faite par du silastène biocompatible, [22, 23, 111], figure 2-3.

Clarion fournit les deux configurations, monopolaire et bipolaire. Huit paires d'électrodes espacées de 2 mm sont utilisées [112]. Med-El utilise huit électrodes parmi douze disponibles et sont espacées de 2.8 mm en configuration monopolaire [113].

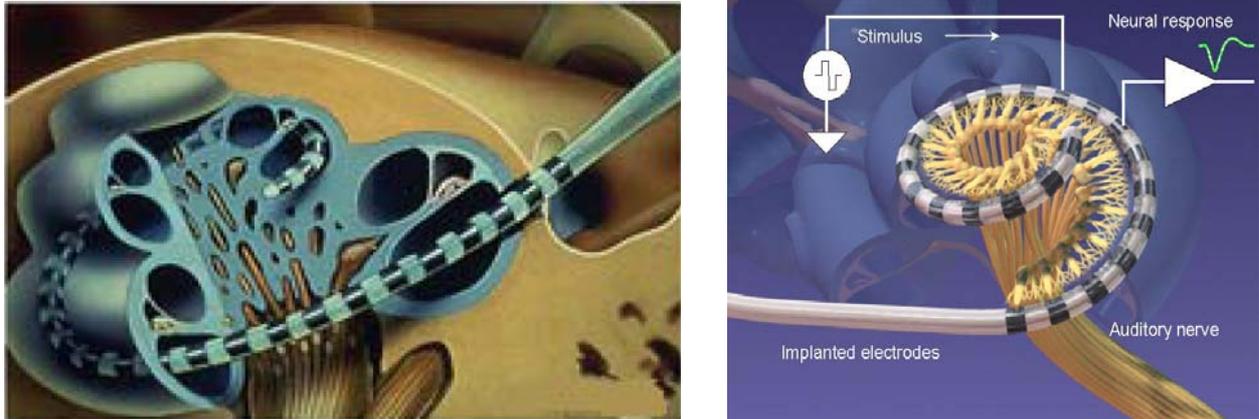


Figure 2-3: Multi-électrodes insérée dans une cochlée pour la stimulation des fibres nerveuses

2.2.2. Type de stimulation

Il y a généralement deux types de stimulation relatives à la présentation de l'information aux électrodes : la stimulation analogique et la stimulation impulsionnelle.

En stimulation analogique, la forme d'onde acoustique est filtrée en bandes fréquentielles, et les signaux analogiques filtrés sont présentés simultanément à toutes les électrodes. Un inconvénient de la stimulation analogique est que la stimulation simultanée peut causer une interaction entre canaux et occasionner une diminution dans la compréhension de la parole (Wilson et al, 1991), en plus d'une consommation prohibitive en puissance suite à la stimulation simultanée de toutes les électrodes.

En stimulation impulsionnelle, l'information est délivrée aux électrodes en utilisant un ensemble d'impulsions brèves. Dans certains composants, les amplitudes de ces impulsions sont extraites à partir d'enveloppes des formes d'ondes filtrées. L'avantage de ce type de stimulation est que les impulsions peuvent être délivrées non simultanément (mode de non recouvrement), ce qui minimise les interactions entre canaux.

Le rythme auquel les impulsions sont délivrées aux électrodes, affecte les performances en reconnaissance de la parole. Des rythmes élevés d'impulsions tendent à donner de meilleures performances que les bas rythmes [38, 56, 109], en fournissant de plus amples détails temporels.

2.2.3. Liaison de Transmission

Les modalités du passage de l'information sonore à travers la barrière cutanée sont un élément de confort très important pour le patient. Deux voies de transmission des signaux du processeur externe vers les électrodes implantées, sont utilisées : à travers une connexion transcutanée ou à travers une connexion percutanée.

Le système percutané transmet directement les stimuli aux électrodes à travers une fiche de connexion. Dans ce système très simple, il n'y a pas d'électronique implantée autre que les électrodes, et les impulsions qui transitent à travers la fiche de connexion sont directement délivrées aux électrodes. Ineraid, fabriqué par Symbion, est l'unique dispositif utilisant un connecteur percutané. Mais même avec les techniques employant les matériaux les plus adéquats, 25% des patients d'Ineraid ont présenté des surinfections cutanées.

La connexion transcutanée, quant à elle, transmet les stimuli à travers une liaison radio fréquence. Le passage par les variations d'un champ électromagnétique font que cette modalité est la plus employée maintenant par les implants.

Un exemple de transfert par bobines couplées, extérieure et implantée, est celui effectué par un implant [23] où l'énergie d'alimentation est transmise à la fréquence porteuse de 112 kHz et les données de stimulation à 2 MHz.

Quand le boîtier du récepteur est hermétiquement scellé en céramique, substance transparente aux ondes magnétiques (cas de Clarion et Med-El), la bobine réceptrice est incluse dans le boîtier, ce qui diminue l'encombrement de celui-ci. Lorsque le boîtier du récepteur est métallique (en titanium) et donc opaque aux ondes magnétiques (cas de Cochlear), la bobine sous-cutanée lui est extérieure. Le meilleur emplacement de la bobine réceptrice est dans la mastoïde avec des creusets de 6 mm. Le maximum de diamètre du boîtier, pouvant être placé confortablement dans des adultes, étant de 35 à 40 mm [22, 23].

L'avantage de la transmission transcutanée, est que la peau est laissée se refermer après opération chirurgicale éliminant ainsi de possibles infections. Le désavantage de ce système est que l'électronique implantée (la circuiterie du récepteur) peut tomber en panne, et cela requiert une chirurgie pour la remplacer.

2.2.4. Traitement du Signal

La plus importante différence entre les dispositifs d'implant, est dans la stratégie de traitement du signal utilisée pour transformer le signal de la parole en stimuli électrique. Plusieurs techniques de traitement du signal ont été développées durant les 25 ans passés. Certaines de ces techniques visent à préserver l'information de la forme d'onde, d'autres visent à en préserver les composantes spectrales. Une description détaillée de ces techniques est présentée dans les sections suivantes. Les résultats succincts de chaque stratégie de traitement du signal sont également présentés.

2.3. Stratégies de codage

Une stratégie de codage définit la manière dont le système d'implant transforme le son en stimulation électrique envoyé au nerf auditif. Afin de représenter la parole de façon précise, la stratégie de codage doit refléter trois paramètres dans son code de stimulation électrique : fréquence, amplitude, et temps. L'information sur la fréquence acoustique est donnée par le site de la stimulation, l'amplitude acoustique est codée par l'amplitude du courant de stimulation, tandis que le caractère temporel est converti en rythme et configuration de la stimulation.

La majorité des adultes implantés sont des sourds post linguaux. En d'autres termes, leur perte d'audition s'est produite après qu'ils aient acquis parole et langage. Les tests psychophysiques après implantation, ont double rôle: ajuster les processeurs-émetteurs des patients et développer de nouvelles stratégies de traitement de la parole destinées aux malentendants [22, 49, 56].

Les stratégies de traitement de la parole disponibles dans les systèmes d'implant, peuvent être divisées en deux principales catégories: stratégies présentant une stimulation analogique simultanée, et celles présentant une stimulation impulsionnelle séquentielle basée sur l'extraction des paramètres de la parole (formants, enveloppes).

2.3.1. Stimulation analogique simultanée

La figure 2-4 montre le diagramme-bloc des implants utilisant une stimulation analogique simultanée [34, 56, 59, 112].

Le signal acoustique est préamplifié sous une CAG éventuelle, puis filtré par un banc de filtres passe-bandes. Les formes d'onde filtrées sont soit pondérées par des gains ajustables propres à chaque voie (ou canal), soit comprimées selon la rangée dynamique du patient. Une liaison transcutanée ou percutanée est ensuite utilisée, afin de stimuler simultanément les électrodes

implantées de manière analogique.

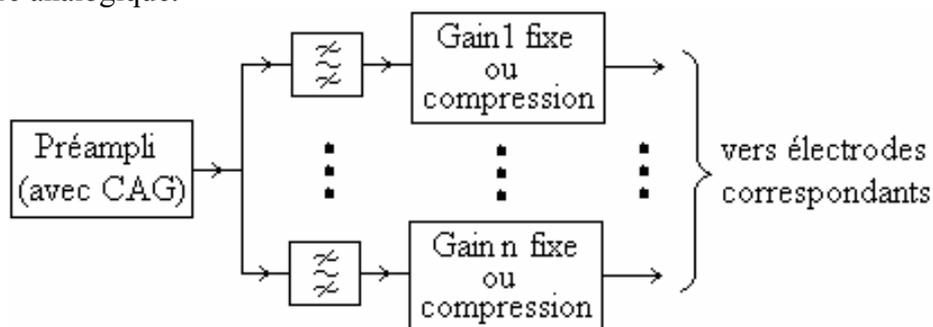


Figure 2-4: Implants à stratégie analogique simultanée

Les implants à canal unique ont été les premiers à être utilisés sur des sujets humains. Ces dispositifs sont, l'implant House/3M fabriqué en 1970 [52] et l'implant Vienna/3M fabriqué en 1980 [50], qui fournissent la stimulation électrique à un seul site dans la cochlée, en utilisant une électrode active unique (une seule voie dans la figure 2-4).

L'implant House/3M filtre le son dans la gamme 340-2700 Hz et n'utilise pas de CAG, mais un amplificateur de sortie permettant au patient de contrôler l'intensité de la stimulation. Le signal reçu modulé en AM, est envoyé sans démodulation à l'électrode active implantée.

L'information temporelle de la parole est contenue dans l'enveloppe du signal modulé. Elle subit une distorsion pour un fort niveau sonore non limité par CAG [35]. Vu cette contrainte, la majorité des patients n'ont pas obtenu une reconnaissance du langage ouvert sans lecture labiale [35, 61]. Un score moyen de 37% en identification correcte des consonnes chez quatre patients, a été mesuré par Rosen et al [76]. Dans une étude faite par Danhauer et al. [26], seulement quatre patients (sur 18) ont achevé un score de 2% en identification correcte des mots monosyllabiques.

L'implant Vienna/3M quant à lui, est muni d'une CAG et utilise un filtre d'égalisation fréquentielle dans la gamme 100-4000 Hz. Le signal reçu modulé en AM, est envoyé après démodulation à l'électrode active implantée.

L'écrêtage des niveaux forts du signal d'entrée est évité par CAG. Le filtre d'égalisation est ajusté pour chaque patient de façon à ce que les fréquences filtrées soient d'égal volume sonore (préaccentuation des aigus) [83, 61]. Dans l'étude faite par Hochmair-Desoyer et al. [51] sur un groupe de 22 patients porteurs d'implant Vienna/3M, un score moyen de 30% dans l'identification correcte des mots monosyllabiques et un score moyen de 45% dans l'identification correcte de mots dans des phrases, ont été obtenus. Cependant, les chercheurs Gantz et al. [61], ont constaté que les patients utilisant l'implant Vienna/3M, avaient des scores nulles en reconnaissance du langage ouvert sans lecture labiale.

La stimulation à canal unique impose un codage fréquentiel selon la théorie temporelle [61, 83]. Or le codage temporel de la fréquence par les fibres nerveuses est restreint à 1 kHz, dû à la période réfractaire neurale [1, 23]. Ceci n'est cependant pas suffisant pour une perception de la parole, à cause de l'importante information du signal de la parole jusqu'à 4000 Hz, et au delà. La stimulation à canal unique n'exploite pas le mécanisme du codage fréquentiel en site sur la cochlée, puisque un seul site est stimulé. De ce fait, les implants à canal unique sont capables de transmettre les informations temps/enveloppe (fréquence fondamentale et premier formant relatif aux voyelles), en plus d'une discrimination de certaines consonnes (cas de l'implant Vienna/3M) [31]. L'information en fréquence transmise, est cependant limitée et insuffisante pour une reconnaissance vocale.

Les systèmes multi-canaux exploitent l'arrangement tonotopique de la cochlée en utilisant un nombre suffisant d'électrodes, tel que des groupes discrets de fibres nerveuses peuvent être stimulés. Ainsi, l'information spectrale du signal de la parole est transférée, en plus du codage de la durée et de l'intensité fournies par un système d'implant cochléaire à canal unique.

Il est impossible de découper l'information sonore en autant de bandes de fréquences qu'il y a de fibres nerveuses (le nerf normal contient près de 40.000 fibres). Le nombre le plus élevé d'électrodes utilisées n'est que de 24 (Nucleus).

L'implant Ineraid (ou Symbion) utilise la stratégie CA (Compressed Analog) [34]. Le signal est filtré dans quatre bandes fréquentielles contiguës, centrées sur les fréquences 0.5, 1, 2, et 3.4 kHz. Les formes d'onde filtrées sont pondérées par des gains ajustables, puis sont directement envoyées à travers une connexion percutanée aux quatre électrodes intracochléaires, figure 2-4. Les électrodes espacées de 4 mm, opèrent dans une configuration monopolaire. Le grand espacement entre électrodes, conserve une action précise sur le clavier fréquentiel du tube cochléaire même si les électrodes sont mal isolées. Dorman *et al.* [28] ont établi, pour un échantillon de 50 patients porteurs d'implants Ineraid, un score moyen de 45% pour l'identification correcte de mots dans des phrases, et un score moyen de 14% pour l'identification correcte de mots monosyllabiques. L'approche CA multi canaux a présenté tôt, une performance supérieure dans la reconnaissance du langage par rapport à l'approche à canal unique [34, 61].

L'implant Clarion utilise la stratégie SAS (Simultaneous Analog Stimulation) [56, 112]. Le signal acoustique est traité par huit filtres, figure 2-4. Les formes d'onde analogiques sont délivrées à chaque électrode au rythme de 13000 échantillon/sec par canal, à travers une liaison transcutanée.

Pour certains patients, l'utilisation d'une stimulation simultanée conduit à une perte de discrimination de la parole due à l'interaction entre canal. Lorsque plusieurs électrodes sont stimulées simultanément, les champs électriques peuvent se chevaucher causant des interférences entre canaux. Cette interaction réduit la qualité du son, particulièrement dans un environnement bruité ou musical.

Ce problème est réduit dans l'implant Clarion, par l'utilisation d'un mode bipolaire amélioré (Hi-Focus) qui minimise la probabilité d'interaction entre canaux. D'autre part, le récent implant cochléaire pulsar utilise une stimulation parallèle intelligente qui est une stimulation simultanée sans interaction entre canaux. Pour cela, un procédé avancé de compensation d'interaction de canal est utilisé : le CIC (Channel Interaction Compensation). CIC calcule l'étendue du courant électrique en temps réel et réduit automatiquement le potentiel d'interaction du canal [112], figure 2-5.

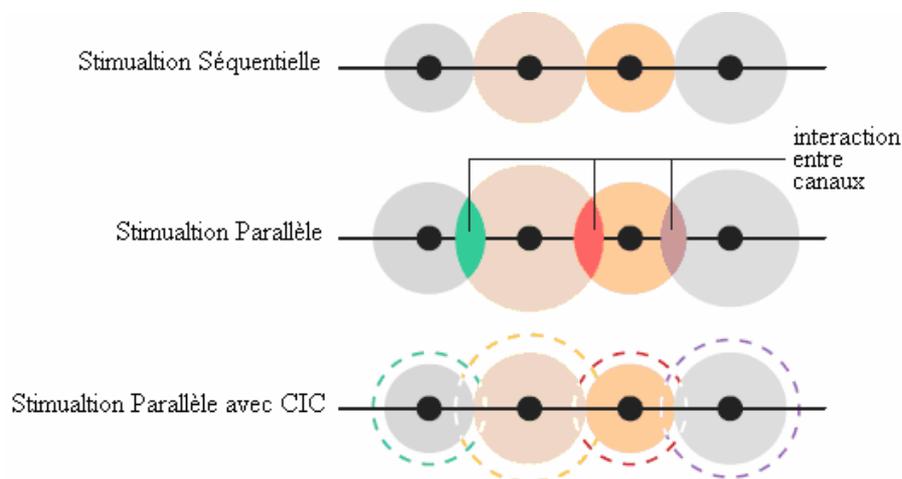


Figure 2-5: Compensation CIC effectuée par l'implant Pulsar

2.3.2. Stimulation impulsionnelle séquentielle :

Dans une stimulation simultanée, la sommation des champs électriques au niveau des électrodes individuelles, cause une interaction entre les canaux et rend les réponses neurales aux stimuli d'une électrode, distordues par les stimuli d'autres électrodes. Ces interactions dégradent la compréhension du langage et restent un problème majeur même s'il est partiellement levé par une technique élaborée (CIC) [112].

Comme une stimulation simultanée consomme plus d'énergie qu'une stimulation non simultanée, cette dernière est adoptée avec des impulsions intercalées non recouvertes. Des trains d'impulsions biphasiques, voir paragraphe 1.4.1, sont délivrés aux électrodes, telle qu'une seule électrode est stimulée à la fois, figure 2-6. La stimulation séquentielle est issue d'une approche basée sur l'extraction des paramètres de la parole. Les amplitudes des impulsions suivent l'enveloppe des formes d'onde filtrées, tandis que la sélection des électrodes à stimuler est contrôlée par les paramètres spectraux extraits de la parole.

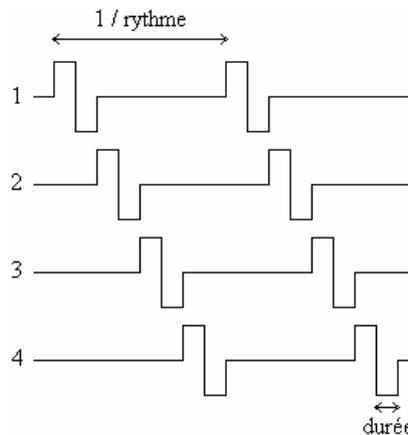


Figure 2-6 : Impulsions intercalées sans recouvrement

L'inconvénient d'une stimulation impulsionnelle est que la quantité d'information par unité de temps, dépend directement de la durée des impulsions individuelles et de leurs rythmes.

Les stratégies d'implants séquentielles se contentent de l'extraction d'enveloppe spectrale de la parole, ou la combinent avec l'extraction de formants.

2.3.2.1. Extraction de formants :

Une redondance d'information se trouve dans l'audition saine. On peut alors se contenter de n'envoyer à l'oreille que les informations les plus pertinentes de la parole comme les formants, bien que cela prive de cette redondance une oreille altérée qui aurait eu sans doute fort besoin du maximum d'informations possible. Ce choix a permis de réaliser un système peu volumineux et beaucoup plus efficace qu'un système monocanal.

La figure 2-7 montre le diagramme-bloc des implants utilisant une stimulation impulsionnelle non simultanée, basée sur l'extraction des formants de la parole [22, 23, 61].

Le signal acoustique est préamplifié avec une CAG. Des détecteurs de passage par zéro sont utilisés pour estimer la fréquence fondamentale F_0 à la sortie d'un filtre passe bas, et les formants F_i à la sortie de filtres passe bandes. Les amplitudes des formants sont estimées avec des détecteurs d'enveloppe, par redressement et filtrage passe bas du signal de bande correspondant. L'amplitude des impulsions de stimulation étant proportionnelle à celles des formants, elles sont appliquées aux électrodes appropriées sur le clavier cochléaire, selon les fréquences des formants obtenues.

L'information du voisement est véhiculée par F_0 en stimulant les électrodes sélectionnées au rythme de F_0 . Durant les segments non voisés, les électrodes sélectionnées sont stimulées à un rythme quasi-aléatoire.

Seul l'implant Nucleus à 22 électrodes, utilise la stratégie d'extraction de formants.

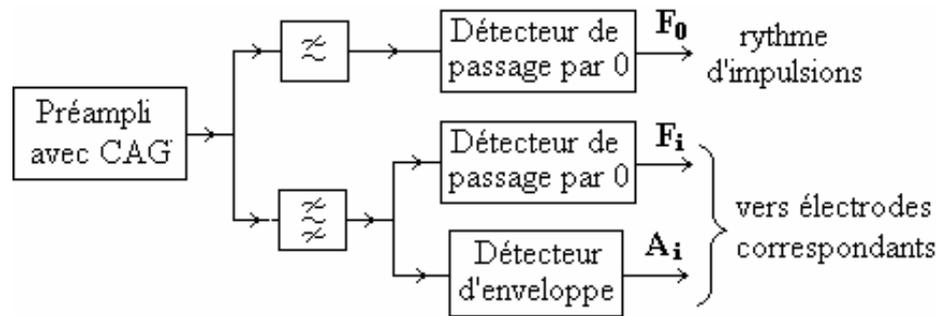


Figure 2-7 : Extraction du pitch et formants de la parole

La stratégie F0/F2 est la première stratégie développée en 1980 [23, 81]. Elle extrait le pitch F_0 à la sortie d'un filtre passe bas à 270 Hz, et le second formant F_2 avec son amplitude A_2 à la sortie d'un filtre passe bande à 1000-4000 Hz [61]. Les résultats avec cette stratégie étaient encourageants car quelques patients ont pu obtenir une compréhension du langage ouvert [32]. Des enfants implantés, ont présenté une amélioration significative dans l'identification de mots fermés mais une reconnaissance très limitée de mots ouverts (Miyamoto, 1989; Staller, 1991). Staller et al. (1991) ont rapporté des scores moyens de 10% en reconnaissance correcte de mots monosyllabiques pour un groupe de 80 enfants. Une performance similaire en reconnaissance auditive-seule de mots ouverts, a été reportée par Osberger (1991) pour 28 enfants. Bien que la performance en audition seule fût limitée, les auteurs ont trouvé que la majorité des enfants qui ont porté tôt le dispositif Nucleus dans leurs vies, ont montré une amélioration significative en reconnaissance verbale avec l'aide de la lecture labiale.

La stratégie F0//F1/F2 devint disponible en 1985 [12]. Des détecteurs additionnels sont inclus pour estimer le premier formant F_1 avec son amplitude A_1 à la sortie d'un filtre passe bande à 280-1000 Hz. Pour les segments voisés, une impulsion électrique est appliquée à une électrode choisie selon F_1 parmi les cinq électrodes les plus apicales, et une autre est appliquée à une électrode choisie selon F_2 parmi les 15 électrodes restantes. Les impulsions biphasiques à durée de phase de 200 μ s, sont espacées de 800 μ s pour éviter toute interaction des canaux pouvant se produire. Comme dans la stratégie F0/F2, les électrodes sont stimulées au rythme de F_0 pour les segments voisés et au rythme moyen de 100 imp/s pour les segments non voisés.

L'apport de l'information F_1 a amélioré la performance de reconnaissance vocale des patients porteurs d'implant cochléaire Nucleus. Dowell et al. [33] ont trouvé une augmentation des scores moyens en reconnaissance correcte de mots, allant de 30% avec la stratégie F0/F2 à 63% avec la stratégie F0/F1/F2. Tye-Murray et al. [102] ont aussi rapporté que les scores moyens en identification correcte de mots monosyllabiques ont passé de 8% avec F0/F2 à 28% avec F0/F1/F2. Cependant, il n'y a pas de différence significative trouvée entre F0/F2 et F0/F1/F2 durant les tests de reconnaissance des consonnes en audition seulement. Ceci est dû à ce que la majorité des consonnes contiennent une information haute fréquence, et que la stratégie F0/F1/F2 apporte plus d'information sur les basses fréquences, nécessaire à la reconnaissance des voyelles.

2.3.2.2. Extraction d'enveloppes

C'est la stratégie la plus utilisée actuellement par les implants cochléaires. Au lieu d'extraire, formants et pitch de la parole, on se contente d'extraire les enveloppes d'ondes filtrées. La stratégie favorise la représentation fréquentielle du message vocal.

La figure 2-8 montre le diagramme-bloc de la stimulation impulsionnelle non simultanée, basée sur l'extraction d'enveloppes de la parole filtrée [46, 61, 63, 71, 80, 108, 118].

Le signal acoustique est préamplifié sous une CAG, puis filtré par un banc de filtres passe

bandes. Les enveloppes des formes d'onde sont alors extraites par redressement et filtrage passe bas. Les sorties d'enveloppes sont généralement comprimées et utilisées pour moduler des impulsions biphasiques de stimulation.

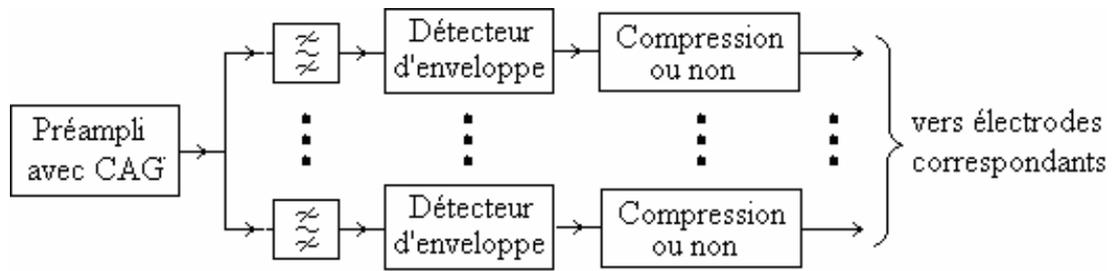


Figure 2-8 : Extraction d'enveloppes d'ondes filtrées

La stratégie MPEAK (MultiPeak) a été utilisée dans l'implant Nucleus pour extraire de la parole, l'information haute fréquence en addition avec l'information issue de la stratégie F0/F1/F2. Cette nouvelle information est extraite en utilisant des détecteurs d'enveloppe, à partir des bandes fréquentielles 2-2.8 kHz, 2.8-4 kHz, et 4-6 kHz, puis envoyée aux électrodes 7, 4, et 1. Les électrodes sont stimulées au rythme de F_0 pour les sons voisés, et à un rythme quasi-aléatoire moyen de 250 imp/s pour les sons non voisés.

Wallenberger et Battmer [106] ont trouvé que la stratégie MPEAK a apporté une amélioration de 17% en identification des consonnes pour un groupe de cinq patients. Une amélioration de 28% en reconnaissance de phrases ouvertes a été également trouvée avec six patients (Sutter et Schreiner, 1991). Cependant, la stratégie MPEAK ainsi que les stratégies F0/F2 et F0/F1/F2, présentent une limitation majeure due aux algorithmes d'extraction des formants. Elles tendent à commettre des erreurs, spécialement dans les situations où le signal de la parole est bruité.

La stratégie de l'implant SMSP (Spectral Maxima Sound Processor) a été développée en 1990 [63]. Elle utilise un banc de 16 filtres passe bandes répartis sur la gamme 250 à 5400 Hz, pour extraire les enveloppes par redressement et filtrage passe-bas à 200 Hz, suivant la figure 2-8. Un détecteur du maxima spectral procède à sélectionner les six plus grandes sorties de filtre, à des intervalles de 4 ms. Ces amplitudes maximales sont comprimées logarithmiquement, pour adaptation à la rangée électrique dynamique du patient, puis transmises aux six électrodes sélectionnées à travers une liaison radio fréquence. Cette technique est souvent désignée par stratégie « n de m ». Il faut noter cependant, que le terme "maxima spectral" réfère aux plus grandes amplitudes des filtres qui ne sont pas nécessairement des pics spectraux. Comme illustré dans la figure 2-9, plusieurs maximums peuvent venir d'un seul pic spectral.

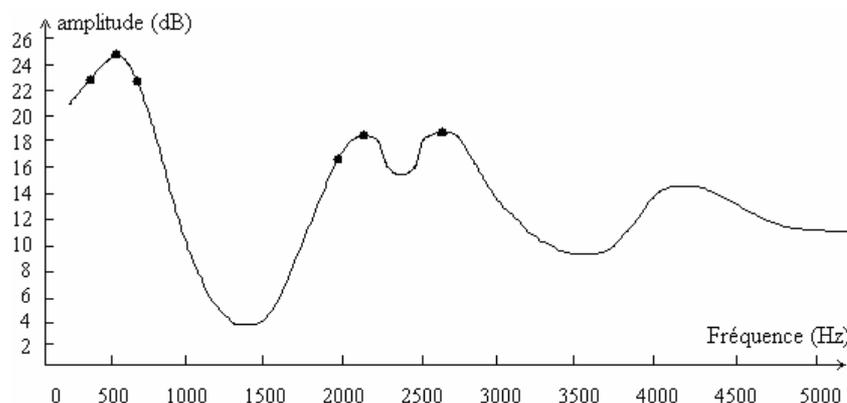


Figure 2-9: Exemple de la sélection des maxima spectraux dans la stratégie SMSP

Les impulsions biphasiques sont délivrées aux six électrodes sélectionnées en mode intercalé au rythme de 250 imp/s. Ce rythme est gardé constant pour les sons voisés et non voisés. Des comparaisons entre les stratégies SMSP et MPEAK menées sur un seul patient, ont montré des améliorations avec la stratégie SMSP pour la reconnaissance de mots, consonnes et voyelles [63].

La stratégie SPEAK (Spectral Peak) a été implémentée dans le processeur Spectra 22 du système d'implant Nucleus à 22 canaux, fabriqué en 1995 [80]. Ce processeur peut être programmé avec une stratégie d'extraction de paramètres (F0/F1/F2, MPEAK), ou la stratégie du pic spectral SPEAK, qui est une stratégie n de m stimulant un certain nombre d'électrodes où les énergies sont les plus élevées. La stratégie SPEAK utilise un banc de 20 filtres (au lieu de 16 dans SMSP) avec des fréquences centrales allant de 250 Hz à 10 kHz. Elle estime de façon continue les sorties des 20 filtres et sélectionne les plus grandes amplitudes (de 5 à 10) dépendant de la composition spectrale du signal d'entrée. Pour un spectre large, plus de maxima sont sélectionnés et le rythme de stimulation est abaissé. Pour un spectre limité, peu de maxima sont sélectionnés et le rythme de stimulation est augmenté afin de fournir plus d'information temporelle. Le rythme de stimulation varie ainsi entre 180 à 300 imp/s dépendant du nombre de pics spectraux identifiés. La stimulation se fait en mode bipolaire BP+n où le nombre n d'électrodes d'espacement situées entre les deux électrodes actives, est fixé selon la sensibilité individuelle.

Une comparaison des stratégies SPEAK et MPEAK effectuée par Skinner et al. [84] sur soixante patients, a montré que la stratégie SPEAK avait une meilleure performance que la stratégie MPEAK, concernant la reconnaissance des voyelles, consonnes, mots monosyllabiques et phrases fermées. De larges améliorations en performance ont été obtenues avec des tests dans le bruit. Ceci est dû à ce que la stratégie MPEAK est basée sur des algorithmes d'extraction de paramètres qui sont connus être sources d'erreurs, spécialement en environnement bruité.

La stratégie SPEAK est maintenant incorporée dans le nouveau système d'implant cochléaire (1998), Nucleus 24 (CI24M) et sa variante Nucleus 24 Contour (CI24RCS). Ce système à aimant interne amovible permet l'imagerie par réponse magnétique (IRM). Selon Abbas [1], les électrodes servent aussi à enregistrer les potentiels auditifs évoqués électriquement dans la cochlée, qui aident dans la programmation du processeur de la parole chez les très jeunes enfants (<2ans). La variante Nucleus 24 Double-Array est un implant spécialement conçu pour les personnes qui présentent une ossification ou une croissance osseuse dans la cochlée, empêchant l'utilisation d'électrodes simples. Cet implant comprend deux rangées d'électrodes plus courtes, avec chacune 11 canaux de stimulation. Le chirurgien perce deux trous dans la cochlée pour y insérer les deux rangées d'électrodes.

Les implants Nucleus 24 à côté des nouveaux processeurs vocaux numériques, Sprint et Esprit, incluent les stratégies CIS et ACE à stimulation rapide, développés par d'autres firmes.

La stratégie CIS (Continuous Interleaved Sampling) est utilisée dans l'implant Clarion en 1991 [108], avec stimulation monopolaire de durée d'impulsion de 75 ms/phase au rythme de 833 imp/s. Elle est la plus rapide actuellement (jusqu'à 18000 échantillons par seconde), favorisant la représentation temporelle du message vocal par une mise à jour relativement rapide. Une performance améliorée en reconnaissance des consonnes est donc constatée (Lawson et al., 1996). Les enveloppes des formes d'onde sont extraites par redressement et filtrage passe bas à 200 ou 400 Hz. Une fonction de compression logarithmique est utilisée pour assurer que les sorties d'enveloppe s'adaptent à la gamme dynamique d'audition du patient. Les sorties d'enveloppes comprimées modulent des impulsions biphasiques avec des amplitudes proportionnelles aux enveloppes, et sont délivrées à six électrodes à un rythme constant sans recouvrement, figure 2-8.

Le rythme auquel les impulsions sont délivrées aux électrodes a un impact majeur sur la reconnaissance du langage. Pratiquement, un rythme élevé des impulsions de stimulation fournit

de meilleures performances qu'un bas rythme. Plusieurs études [13, 29, 30, 108, 109] ont comparé les différences en performance entre les stratégies CA et CIS. Sept patients ont été testés sur la reconnaissance ouverte de 50 mots monosyllabiques et 100 mots clé [108]. Les patients portaient le dispositif Ineraid, qui emploie l'approche CA. Les mêmes patients ont été aussi testés en utilisant une implémentation « laboratoire » de l'approche CIS. Les scores moyens obtenus avec le processeur CIS ont été significativement plus élevés que ceux obtenus avec l'approche CA. Plusieurs facteurs peuvent être responsables d'un tel succès de l'approche CIS sur l'approche CA: L'utilisation d'une stimulation non simultanée, l'utilisation de six canaux plutôt que quatre, et la rapide représentation des variations de l'enveloppe avec l'utilisation d'une stimulation à rythme élevé.

L'implant Clarion 1.2, apparu en 1996, possède 16 électrodes rangées en huit paires (8 canaux) programmées en mode monopolaire ou bipolaire avec stimulation séquentielle ou simultanée. Cette dernière possibilité revient au large espacement entre électrodes qui procure aussi de faibles seuils. Les nouveaux implants Clarion CII, Series-S et Pulsar sont devenus en 2002, les uniques implants pouvant stimuler avec des formes d'ondes analogiques grâce à une multi-électrode Hi-Focus (voir stratégie SAS dans 2.3.1). Cette électrode a 16 contacts et 16 canaux assurant un rythme de 2840 imp/s par canal pour CIS lorsque tous les 16 canaux sont activés, et 5980 imp/s par canal lorsque 8 canaux sont programmés.

Loizou et autres (1997) ont comparés SPEAK et CIS. CIS s'est avérée meilleure pour la perception de voyelle dans le bruit quand le rapport du signal sur bruit (SNR) était à +5dB. À tous les niveaux de SNR examiné, la stratégie CIS a fourni une meilleure perception des consonnes que SPEAK. Aucune différence n'a été trouvée, cependant, dans la perception des phrases. Dans une étude faite par Loeb et Kessler [60], trente deux patients porteurs d'implant Clarion ont obtenu des scores de 30% à 100% en reconnaissance du langage ouvert. Les résultats montrent aussi que près de 50% des utilisateurs de Clarion Series-S, préfèrent la stratégie analogique SAS via le mode de couplage bipolaire amélioré, que la stratégie CIS. De plus, le processeur Clarion Series-S fournit l'option d'une nouvelle stratégie de stimulation mixte, simultanée et séquentielle, appelée PPS et qui sera présentée plus loin.

L'implant Clarion CII, quant à lui, est équipé de télémétrie pour la gestion des électrodes et l'enregistrement des potentiels auditifs évoqués électriquement, présentant les mêmes possibilités que chez Nucleus concernant le système d'imagerie par réponse neurale (IRN). Cependant, le boîtier d'implant est en céramique contenant l'aimant qui n'est ni extractable ni compatible IRM. Nombre de paramètres associés à l'approche CIS, peuvent être ajustés afin d'optimiser les performances en reconnaissance de la parole pour chaque patient [29, 109]. Ces paramètres incluent :

- *Le rythme d'impulsion :*

Entre 100 imp/s et 2500 imp/s avec une durée de 33 μ s/phase [109].

- *L'ordre de stimulation :*

Réfère à celui avec lequel les électrodes sont stimulées. Ceci pour minimiser une possible interaction entre canaux. Une possibilité est de stimuler les électrodes dans l'ordre apex-vers-base, c'est-à-dire, électrode 1 stimulée en premier, puis électrode 2, ..., jusqu'à électrode 6. Cependant, cet ordre apex-vers-base ne minimise pas la séparation spatiale entre les électrodes stimulées séquentiellement. Alternativement, les électrodes peuvent être stimulées dans un ordre décalé, c'est-à-dire, dans l'ordre six-trois-cinq-deux-quatre-un, qui maximise la séparation spatiale entre électrodes stimulées. Comme pour le rythme d'impulsion, la préférence de l'ordre de stimulation varie d'un patient à un autre.

- *La Fonction de compression :*

La compression des sorties d'enveloppe transforme les amplitudes acoustiques en amplitudes électriques. C'est une transformation nécessaire, car la rangée des amplitudes acoustiques du langage parlé est considérablement plus étendue que la rangée dynamique des

implants des patients. Les auditeurs implantés peuvent avoir une rangée dynamique aussi faible que 6 dB, allant du seuil ou bas niveau audible (threshold, THR) au niveau maximal de confort sonore (most comfortable level, MCL) [23, 61].

Il est montré que le volume sonore, généré par un stimulus électrique en microampère, est analogue au volume sonore d'un stimulus acoustique en dB [38, 117]. La fonction de compression logarithmique de la forme $Y = A \log(x) + B$ est typiquement utilisée, où x est l'amplitude acoustique (sortie du détecteur d'enveloppe), A et B sont des constantes, et Y est l'amplitude électrique (comprimée). Les constantes A et B sont choisies telle que la rangée d'entrée acoustique $[x_{\min}, x_{\max}]$ est adaptée à la rangée électrique dynamique [THR, MCL], figure 2-10.

$$A = \frac{MCL - THR}{\log_{10}(x_{\max}) - \log_{10}(x_{\min})} \quad 2.1$$

$$B = THR - A \log_{10}(x_{\min}) \quad 2.2$$

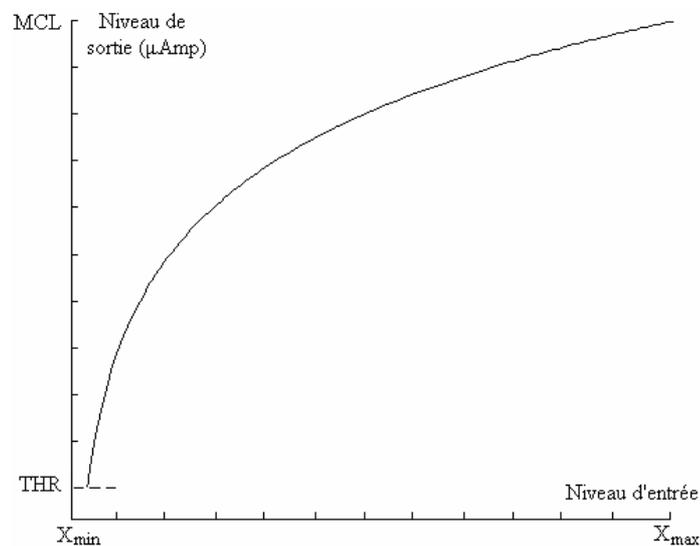


Figure 2-10: Compression logarithmique utilisée dans la stratégie CIS

Les valeurs de seuil, THR, et de maximum de confort, MCL, peuvent varier d'une électrode à une autre. La stratégie CIS est utilisée dans trois implants du commerce : Clarion, Nucleus, et Med-El. Ce dernier implant, est aussi désigné par COMBI-40 [46]. Son processeur est basé sur le DSP 56001 de Motorola [119], et peut être programmé à haut rythme avec les stratégies CIS ou ACE (cette dernière est présentée plus loin). L'implant utilise 12 électrodes avec séparation entre contacts de 2.4mm, ce qui permet une insertion profonde de 26.4mm dans la cochlée (deux tours complets). Cette longueur d'insertion est la plus profonde parmi les implants cochléaires commerciaux. Elle permet l'accès à un plus grand nombre de neurones dans la cochlée correspondant à une plage fréquentielle de 70Hz à 8500Hz approximativement. Théoriquement, une telle large plage contribue à une meilleure appréciation de la musique et compréhension d'une variété de langages parlés (incluant les langages tonals). De même, la motivation d'utiliser un espacement entre électrodes aussi large est d'augmenter le nombre de canaux de perception.

L'implant Med-El utilise la stratégie CIS avec huit électrodes monopolaires au rythme maximal de 1515 imp/s par canal [118]. Comme dans Clarion, l'électronique interne et l'aimant sont emboîtés dans du céramique, sauf que le système autorise une imagerie IRM. Les processeurs plus récents, CIS-Pro+ et Tempo+ (derrière l'oreille), offrent deux stratégies : la stratégie CIS+ utilisant la transformée d'Hilbert pour l'amélioration de la détection d'enveloppe, et la stratégie ACE de type "n-de-m" avec un rythme élevé d'impulsion.

De nouvelles stratégies visent actuellement à améliorer la performance perceptive des implantés en jumelant les points forts des différentes stratégies précédentes. Par exemple :

La stratégie PPS (Paired Pulsatile Stimulation) est une variante du CIS introduite en 1999, à la différence que les impulsions sont envoyées simultanément dans 2 canaux afin d'augmenter la vitesse de stimulation au double de CIS. Les paires d'électrodes étant suffisamment espacées, l'effet d'interaction entre canaux est donc minimisé. C'est donc une stratégie partiellement simultanée qui combine la stimulation séquentielle et la stimulation simultanée dans le but d'augmenter la quantité d'information transmise par unité de temps.

La stratégie ACE (Advanced Combination Encoders) est sensée allier le meilleur de CIS et de SPEAK. De CIS, on a retenu le rythme élevé de stimulation (constante mais paramétrable), et de SPEAK, un système de stimulation sélectionnant les sites d'activation nerveuse en fonction de l'énergie spectrale, c'est à dire du type "n-de-m". Dans cette stratégie, le signal est filtré dans m bandes fréquentielles et le processeur sélectionne, parmi les m sorties d'enveloppe, les n sorties les plus énergétiques afin de stimuler les n électrodes correspondantes. Elle fournit donc une profusion d'informations avec combinaison des aspects temporels et fréquentiels. On a constaté que le rythme de stimulation et le nombre de maximum choisis, peuvent être optimisés pour chacun des sujets. Par exemple, 8 maximums et 900 Hz par canal, ou 6 maximums et 720 Hz par canal, etc..

Nous réunissons dans les tableaux 2.2 et 2.3, les principales caractéristiques d'implants cochléaires multicanaux, d'ancienne et d'actuelle génération respectivement.

Stratégie	Représentation du signal	Stimulation	Nombre de canaux	Rythme par canal	Implant
CA	Formes d'onde filtrées	Analogique	4	Formes d'onde Continues	Ineraid
F0/F2	Formant F2, Voisement F0	Impulsionnelle	1	F0 ou aléatoire	Nucleus
F0/F1/F2	Formants F1 F2, Voisement F0	Impulsionnelle	2	F0 ou aléatoire	Nucleus
MPEAK	Formants F1 F2, Signaux d'enveloppe	Impulsionnelle	4	F0 ou aléatoire	Nucleus
SMSP	Signaux d'enveloppe, Maxima spectral	Impulsionnelle	6	250 imp/s	Nucleus

Table 2.2: Anciennes stratégies utilisées dans les implants cochléaires multi-canaux

Constructeurs	Cochlear (Australie)	Advanced Bionics (USA)	Med-El (Autriche)
Implants	Nucleus 22 / 24	Clarion 1.2 / S / CII – Pulsar	Combi 40+ / S / GB
Processeurs	Spectra / Sprint / Esprit	Platinum Sound / BTE	CIS-Pro+ / Tempo+
Stratégies & Rythmes	Speak : 180 à 300 imp/s CIS : 1200 imp/s ACE : 900 à 1200 imp/s	CIS : 833 imp/s SAS : 13000 éch/s PPS : 1666 imp/s	CIS+ : 1515 imp/s ACE : ?
Télémetrie IRN	Oui (Nucleus 24)	Oui (CII)	Oui
Boîtier hermétique	Titanium	Céramique	Céramique

Electrodes		22 / 24 (auto-courbure)	16 (Hi-Focus & CIC)	12 (espacement élevé)
Canaux		5 à 10 (Speak) / 8 à 12 (CIS et ACE)	8 (Clarion 1.2 et S) / 16 (Clarion CII et Pulsar)	8
Stimulation	Mode	BP+n (+ MP pour Nucleus 24) Séquentiel	BP (amélioré) + MP Simultané ou Séquentiel	MP Séquentiel
	Type	Impulsionnel	Impulsionnel ou Analogique	Impulsionnel
Possibilité IRM		Oui (Nucleus 24)	Non	Oui
Durée des piles		35 H (Nucleus 22) 50 H (Nucleus 24)	8 à 11 H	50 H

Table 2.3: Etat de l'art actuel des implants cochléaires les plus utilisés dans le monde

2.3.3. Résultats d'implantation avec les systèmes multi-canaux [4]

Des tests post-opératoires de l'implant, visent à la réhabilitation des patients :

- dans la reconnaissance de mots et de phrases à différentes vitesses d'expression.
- aux difficultés lexiques (mots phonétiquement similaires).
- en tests multisyllabiques à multi-orateurs, en présence de bruit.

Les systèmes d'implant cochléaire actuels conçus par Cochlear Corporation, Advanced Bionics, et Med-El donnent remarquablement des résultats similaires.

Des scores moyens de 35% à 45% en reconnaissance correcte de mots sans lecture labiale, ainsi que des scores de 65% à 80% en reconnaissance correcte de phrases, ont été mentionnés par des utilisateurs d'implants cochléaires du type :

- Nucleus, muni de la stratégie de traitement SPEAK (Hodges, 1999; Staller, 1997).
- Clarion, muni des stratégies de traitement CIS ou SAS (Osberger et Fisher, 2001).
- Med-El, muni de la stratégie de traitement CIS (Hochmair Desoyer, 1997).

La plupart des patients adultes ressentent maintenant une compréhension substantielle dès les trois premiers mois post-opératoires (Geir, 1999; Waltzman et Roland, 1999). En contraste aux adultes sourds post linguaux implantés, les enfants doivent utiliser le son en provenance de l'implant cochléaire pour acquérir le langage parlé. La vitesse de développement des qualifications auditives des enfants après implantation, semble augmenter en fonction d'amélioration technologique des implants cochléaires et l'âge d'implantation (Allum, 2000; Carrasco, 1999). Une implantation effectuée tôt, au plus jeune âge des enfants, donne de meilleures performances (Battmer, 1999; Miyamoto, 1997; Archbold, 1998). Cependant, il faut noter que le développement auditif est confondu avec les capacités du langage de l'enfant. (Blamey et al, 2001).

Les implants cochléaires sont vus comme des prothèses auditives dont le principal bénéfice est l'amélioration de l'audition. Cependant, lorsque ces dispositifs sont utilisés chez des enfants rendus sourds tôt dans leurs vies, les bénéfices sont substantiellement plus large dû à l'importance de l'audition dans l'acquisition du langage parlé. Connor et Zwolan (2000), ont trouvé que l'implantation au plus jeune âge résulte dans de meilleures productions de la parole.

Ces études ainsi que d'autres, constituent une forte évidence de l'apport d'expérience auditive améliorée fournie par l'implant cochléaire pour enfants supportant l'acquisition du langage parlé (Blamey, 2001; Bollard, 1999; Connor, 2000). Les capacités de communication des enfants porteurs d'implants cochléaires, ont été comparées avec celles d'autres enfants, du même âge

avec mêmes seuils d'audition, mais porteurs d'aides auditives. Les utilisateurs d'implant cochléaire ont généralement montré des résultats supérieurs (Riley, 1996; Miyamoto, 1989; Myres, 1991). Après 2 à 3 années d'utilisation d'implant, plusieurs études ont toutes rapporté une production précise et intelligible de la parole, significativement meilleure chez les enfants implantés en comparaison avec les enfants utilisant des dispositifs d'aide d'audition. (Sehgal, 1997; Tobey, 1991; Brenner, 1994; Woodworth, 1995).

2.4. Les constituants communs d'un implant cochléaire [110, 111, 112, 113]

La transformation de la parole en signaux de stimulation électrique adéquats, passe successivement par les constituants d'implant mentionnés sur la figure 2-11.



- Avec :
1. microphone
 2. processeur de signal
 3. bobine de transmission
 4. bobine de réception alimentant le récepteur implanté
 5. gaine multi électrodes
 6. mal-entendant porteur d'implant

Figure 2-11: Emplacement et constituants d'un implant cochléaire

Le récepteur doit permettre un maximum de flexibilité dans l'utilisation des paramètres du stimuli. Les concepteurs d'implants cochléaires ne sont pas encore sûrs de l'ensemble des paramètres requis ou de la stratégie de codage de la parole, pour permettre aux patients de comprendre parfaitement le langage ouvert [4, 31, 50, 61].

La plupart des systèmes d'implants cochléaires comme ceux de renommée mondiale (Nucleus, Clarion, et Med-El) présentent plusieurs similitudes :

1. une stimulation multicanal via des gaines d'électrodes à multiple contacts, insérées dans la rampe tympanique à partir d'une ouverture chirurgicale faite tout près de la fenêtre ronde. Le nombre des contacts (ou électrodes) varie cependant d'un implant à l'autre,
2. une possibilité d'implémentation d'une stimulation monopolaire, vu l'existence d'électrodes extracochléaires,
3. une communication transcutanée entre le transmetteur externe et les composants électroniques implantés via des bobines HF couplées. La transmission RF fournit à la fois, l'énergie d'alimentation du circuit électronique implanté et les signaux de contrôle du type et niveau de stimulation, donnant des sensations auditives interprétables. Des aimants (l'un étant implanté sous la peau) sont utilisés pour maintenir le contact entre les bobines de transmission et de réception,

4. une télémétrie utilisée pour vérifier l'intégrité des électrodes intracochléaires après leur implantation. Cette opportunité est importante car un dysfonctionnement de l'électronique implantée n'est pas toujours facile à détecter, surtout chez les jeunes enfants et individus avec une expérience auditive très limitée,
5. un choix parmi différentes stratégies ou algorithmes de traitement de la parole, portées en options. Ces stratégies représentent chacune une approche de conversion du signal acoustique en un signal électrique adéquat. Ceci inclut le nombre et location des électrodes à stimuler, ainsi que le type, vitesse et amplitude de la stimulation,
6. un établissement d'intervalles du niveau de stimulation allant du seuil au maximum d'écoute confortable du patient, et ce pour chacune des électrodes intracochléaires (dressement d'une carte du patient que l'audiologiste ajuste régulièrement).

Une autre similitude porte sur le coût des systèmes d'implants cochléaires, qui ne varie pas significativement d'un constructeur à l'autre. Cependant, les différentes firmes d'implants continuent leurs recherches et développements dans le but d'améliorer leurs produits.

2.5. Choix d'implants [4, 61]

Le choix d'un implant pour un patient dépend de plusieurs facteurs, incluant le centre hospitalier dont le patient appartient et sa préférence envers quel chirurgien et type d'implant. Avec la technologie courante des implants cochléaires, les résultats obtenus par les trois constructeurs, Nucleus, Clarion, et Med-El, sont similaires. Pour chaque type d'implant, certains patients obtiennent une compréhension substantielle de la parole sans aide labiale, alors que d'autres utilisent l'implant comme aide à la lecture labiale. La stratégie SAS a rencontré un succès limité chez certains patients, tandis que les stratégies SPEAK et CIS ont relativement réussi. Il apparaît qu'aucun système n'est effectif pour tous les implantés. C'est pour cette raison, que les constructeurs offrent actuellement plusieurs stratégies de traitement de la parole dans un même implant. Ceci permet à l'audiologiste et au patient de choisir la meilleure stratégie qui convient à ce dernier.

Les systèmes Nucleus emploient des stratégies telles que SPEAK, ACE, et CIS. Les systèmes Clarion utilisent principalement les stratégies CIS en mode monopolaire séquentiel et SAS en mode bipolaire simultané. Med-El (Medical Electronic) utilise la stratégie CIS la plus rapide de tous les implants. Des études récentes sont menées pour le développement de gaine d'électrode courbée, qui s'approche au mieux de la membrane basilaire et demandent ainsi moins de courant pour stimuler le nerf [113]. Ceci améliore éventuellement la spécificité spatiale de la stimulation.

2.6. Proposition d'une stratégie améliorée

La performance optimale d'une stimulation électrique du nerf auditif, dépend du signal sensoriel fourni qui doit émuler les caractéristiques de la réponse neurale naturelle.

Les filtres constitués des neurones, codifient le signal reçu à partir des cellules ciliées sous forme d'impulsions de décharge, voir chapitre 1. La réponse des fibres (décharges) est synchronisée sur la fréquence des formants de voyelles (information du lieu), ainsi que les transitions formantiques des consonnes (information de durée). C'est donc le signal fourni par les cellules ciliées internes qu'il faut reproduire par stimulation électrique, figure 1-11.

Les stratégies utilisées dans les implants cochléaires du commerce contiennent, pour la plupart, les deux types d'information: tonotopique et temporelle.

Une performance moyenne significativement améliorée a été constatée avec les trois implants Nucleus, Clarion, et Med-El, durant la décade passée. Les meilleurs utilisateurs d'implant cochléaire obtiennent maintenant des scores de 80% ou plus en reconnaissance de mots, sans lecture labiale, avec tout type d'implant confondu. Cependant, ce n'est pas tous les utilisateurs d'implant cochléaire qui jouissent de pareilles performances. Quel que soit le type d'implant porté, on trouve des patients qui obtiennent une reconnaissance limitée de mots ouverts [4]. Pour

ces individus, un large bénéfice est montré lorsque le son en provenance de l'implant cochléaire est combiné avec la lecture labiale.

Parmi toutes les stratégies de stimulation électrique de la cochlée présentées dans ce chapitre, nous avons essayé d'emprunter la méthode d'extraction des paramètres de la parole. En effet, cette méthode élimine la redondance du signal de la parole en se limitant à ses paramètres pertinents, tout en assurant un faible débit de transmission.

Par exemple, les stratégies F0/F1/F2, MPEAK, et SPEAK, incorporent un extracteur des formants de la parole. Des détecteurs de passage par zéro sont utilisés pour estimer le pitch F0 et les formants F1 et F2 de la parole, en sortie d'un banc de filtre. Cependant, cette méthode présente une limitation importante. Elle génère des erreurs dans l'extraction des formants, spécialement dans les situations où le signal de la parole est bruité. Cette limitation est inhérente aux algorithmes d'extraction [61]. En plus, la transmission de ces paramètres apporte plus d'information sur les basses fréquences (voyelles), que les hautes fréquences (consonnes). Ceci a poussé les constructeurs de se contenter de la localisation des énergies les plus élevées en sortie du banc de filtre et d'extraire l'information haute fréquence en utilisant des détecteurs d'enveloppe, à partir des bandes fréquentielles supérieures à 2 kHz.

Nous essaierons dans ce qui suit, d'adopter une nouvelle démarche inspirée de la norme GSM de la téléphonie portable, pour contourner les divers problèmes liés à l'extraction des paramètres de la parole.

2.6.1. Bilan des données physiologiques et psycho-acoustiques

La réhabilitation des surdités totales, profondes ou sévères par l'implant cochléaire doit se baser sur les données anatomo-physiologique de l'audition. Les constatations les plus importantes qu'on peut tirer du chapitre 1, se résument aux aspects suivants.

2.6.1.1. Intensité électrique de stimulation :

Pour la perception de l'intensité de la stimulation électrique, la rampe d'intensité est linéaire dans une rangée dynamique de 1 à 7 dB [61, 117]. Cette rangée d'intensité de courant étant étroite, elle présente cependant jusqu'à 60 pas discriminables [22, 23]. Ce qui rend possible l'utilisation de l'information intensité dans le traitement de la parole pour implant cochléaire.

2.6.1.2. Aspect de la théorie du temps :

Les caractéristiques temporelles du nerf auditif sont imposées par ses propriétés réfractaires [20, 23, 107]. La réponse d'une cellule ciliée via une microélectrode, figure 1-10, reproduit la forme complète du stimuli présenté à une fréquence $f \leq 1$ kHz, alors qu'elle ne fournit que l'enveloppe du stimuli pour une fréquence $f > 1$ kHz. Nous utiliserons donc le pitch F_0 comme rythme des impulsions électriques de stimulation.

2.6.1.3. Aspect de la théorie du lieu :

La membrane basilaire effectue une transformée de Fourier mécanique, produisant des activités maximales aux pics du spectre de l'information sonore. L'organisation tonotopique existant dans la cochlée [20, 58] a été démontrée par la simulation numérique de fonctionnement de la membrane, figure 1-7. D'autre part, chaque cellule ciliée manifeste une sélectivité en fréquence, où le moindre écart de sa fréquence caractéristique nécessite une augmentation de l'intensité pour reproduire une réponse détectable, figure 1-8. Nous proposons par conséquent, de détecter jusqu'à quatre formants de la parole, F_1 à F_4 , pour une stimulation en lieu de la cochlée. Les formants F_1 et F_2 suffisent à la représentation convenable des voyelles [22, 23], tandis que les formants F_3 et F_4 apportent l'information haute fréquence, nécessaire à la reconnaissance des consonnes [74].

2.6.1.4. *Découpage en tranches de la parole :*

Pour le lieu de stimulation variant dans le temps, des tests ont montré la capacité des patients à détecter un décalage en position d'électrode d'une durée de 20 ms [23]. Ainsi le lieu de stimulation convient pour le codage des trajets formantiques des consonnes (courtes transitions fréquentielles des formants), et les paramètres à stimuler doivent être rafraîchis au moins chaque 20 ms.

2.6.2. **Principe d'analyse spectrale**

L'analyse spectrale est basée sur le fait que l'énergie du son est concentrée en bandes fréquentielles, ou formants. Cette analyse peut s'effectuer selon 2 principes:

- avec des bancs de filtres, en plaçant en parallèles des filtres dont les bandes passantes sont étalées des graves aux aigus, et en mesurant l'énergie à la sortie de chaque filtre. Chaque filtre correspond à une rangée fréquentielle d'une électrode déterminée dans la cochlée. Les électrodes sont simulées proportionnellement à l'amplitude de la réponse des filtres passe-bande. Plus les filtres sont nombreux, plus l'analyse est fine. Autrefois, cette analyse par bancs de filtres était réalisée de manière analogique.
- par transformée de Fourier (FFT) : c'est une formule qui prend en compte des mesures échantillonnées du son à analyser. Elle permet d'en déduire le spectre du signal à court terme.
- nous proposons d'utiliser un troisième principe d'analyse spectrale basé sur le codage paramétrique de la parole, qui estime en plus des formants, leurs bandes, l'intensité et le pitch.

Ainsi, le processeur fournit des paramètres estimés de la parole, qui sont codés et présentés au patient implanté. Ces paramètres sont choisis parmi les plus utiles dans l'information véhiculée par le langage [22, 23], et sont :

<i>paramètre de la parole</i>		<i>paramètre électrique</i>		<i>perception auditive</i>
niveau de pression sonore	→	intensité du courant	→	son
fréquence fondamentale	→	rythme des impulsions	→	pitch
formants et bandes	→	position et nombre d'électrodes	→	tonalité

Le rythme de la voix est donc perçu comme vitesse de stimulation, et les fréquences des formants comme lieux de stimulation. La stimulation peut être simultanée ou non. Dans ce dernier cas, les stimulés entre canaux sont toujours séparés par un intervalle court de temps, de l'ordre de 0.8 ms [23]. La stimulation simultanée exige un certain écart entre électrodes adjacentes, avec l'utilisation du mode bipolaire à action plus sélective.

2.6.3. **Caractéristiques de la prothèse cochléaire suggérée**

La stratégie multiélectrode présentant de brèves impulsions à chaque électrode dans une séquence à recouvrement ou non, a donné de bons résultats lors des tests de reconnaissance de la parole, voir paragraphes précédents.

Ainsi, tenant compte des données physiologiques et psycho-acoustiques recueillies expérimentalement et selon les meilleurs résultats obtenus dans les différentes stratégies décrites précédemment; notre implant sera intra-cochléaire, multi-électrode à canaux multiples, qui exploite la solution combinant rythme et lieu de stimulation. L'approche retenue sera de remplacer la technique répandue, utilisant des bancs de filtrage du signal de la parole, par une autre estimant son spectre par codage paramétrique:

- Le spectre est évalué aux gammes fréquentielles des quatre premiers formants, F_1 à F_4 . La plage $F_0=[50 \text{ à } 1000 \text{ Hz}]$ sera attribuée au rythme du pitch.

- Un canal sera réservé à chaque formant. Les quatre canaux peuvent être activés en même temps ou multiplexés temporellement, utilisant ainsi des stimulations simultanées ou non.
- Les canaux F_2 , F_3 , et F_4 , dans l'échelle des aigus, seront desservis par les électrodes basales en stimulation bipolaire (localisée) du type BP+n utilisant des paires d'électrodes séparées par n électrodes neutres. n étant fonction de la bande du formant à stimuler.
- Le canal F_1 , dans l'échelle des graves, sera alimenté par les électrodes apicales en stimulation monopolaire (étendue). Cependant, une stimulation bipolaire du type BP+n, peut être retenue.
- L'intensité sonore est stimulée linéairement par l'amplitude du courant, tout en respectant la rangée dynamique du patient.

Pour cela, le stimulateur neural implantable se compose des parties suivantes:

1. un transmetteur contenant :

- un microphone directionnel (pour améliorer le rapport signal/bruit).
- un ADC pour l'échantillonnage et numérisation de la parole.
- un processeur de la parole (DSP) effectuant les traitements de:
 - préaccentuation de la parole (équivalente à l'amplification sélective de l'oreille externe, figure 1-11).
 - fenêtrage de la parole pour traitement à court terme (permet de considérer des tranches stationnaires de la parole sur 20 ms).
 - analyse spectrale de la parole par codage paramétrique à faible débit (codeur de source).
- un codeur de canal pour protection des paramètres de stimulation à transmettre.
- un lien de communication RF véhiculant données et énergie d'alimentation à l'implant, par le biais d'un transformateur HF transdermal.

2. un récepteur-stimulateur contenant :

- un circuit de conversion d'énergie RF reçue en alimentation continue symétrique.
- un démodulateur RF pour la voie des données.
- un décodeur des paramètres de stimulation : rythme, intensité, numéros de paires d'électrode par formant détecté, type de stimulation (simultanée ou non), etc.
- une source de courant biphasique contrôlé en intensité et fréquence jusqu'à 1 kHz, et alimentant les électrodes via des miroirs de courant.
- un ensemble d'électrodes délivrant les stimuli électriques aux sites neuraux appropriés.

De 8 à 20 électrodes sont nécessaires pour profiter pleinement de l'organisation des fibres nerveuses dans la cochlée [111, 112, 113]. Dans notre cas, la multi-électrode reste à choisir parmi celles disponibles dans le commerce.

Les données de la parole ainsi que l'énergie d'alimentation, sont tous deux transmis par couplage inductif utilisant une porteuse RF. La transmission est conçue de manière à ce qu'une puissance adéquate soit transférée sur une distance d'environ 10 mm au maximum, lorsque les bobines sont coaxiales avec un certain degré de désalignement qui reste tolérable à faible distance. Ce problème sera abordé dans le chapitre suivant.

Conclusion

L'implant cochléaire a pour but de shunter un organe de Corti déficient. Il va donc agir en aval de celui-ci en stimulant directement les terminaisons nerveuses du nerf auditif. Comme l'organe de Corti qu'il tente de remplacer, il assure la transduction de l'énergie, et le codage du

message acoustique. Il faut bien sûr que le nerf auditif soit fonctionnel. C'est le principe d'un test qu'on utilise en préopératoire, qui permet de stimuler l'oreille interne et d'enregistrer la réponse des centres auditifs. Cela permet d'affirmer une éventuelle intégrité du nerf auditif.

Depuis son invention, l'implant cochléaire a subi d'énormes améliorations. Les implants extra-cochléaires sont plutôt utilisés dans les cas de malformation anatomique comme une ossification de la cochlée ou l'imperméabilité de la cochlée à la pénétration d'une électrode. Pour les malentendants gardant une faible audition naturelle intacte, ce sont des implants intra-cochléaires multi-électrodes, qui permettent l'utilisation d'un courant plus faible et la possibilité d'avoir des informations plus complètes.

Les implants multi-canaux essaient de reproduire la physiologie cochléaire en répartissant tonotopiquement l'analyse fréquentielle du message sonore le long de la cochlée : un son d'une fréquence donnée stimule une zone précise dans la cochlée, donc certains neurones du nerf auditif, ce qui permet la différenciation des fréquences du son par le patient.

L'implant cochléaire est formé de deux parties : une partie implantée et une partie externe.

La partie implantée est formée d'électrodes qui sont reliées par une gaine et sont insérées dans la cochlée, et un récepteur situé sous le cuir chevelu.

La partie externe comprend un microphone et un processeur de traitement porté à la taille ou sur l'oreille. Le message est envoyé via l'antenne à la partie interne pour être appliqué au nerf auditif. Les stratégies de codage sont les techniques utilisées par le processeur vocal pour convertir les aspects fréquentiels, temporels et intensité du son en signaux que l'implant transmet au nerf auditif. Des résultats satisfaisants sont obtenus généralement avec des adultes post-linguaux. Pour les enfants pré-linguaux, les facteurs influençant les différences de résultats sont : l'âge d'implantation, et l'entraînement en communication orale. Une implantation faite tôt dans la vie tend à être associée à de meilleurs résultats, de même qu'une rééducation orale intensive favorise le développement d'une meilleure production linguistique.

Puisqu'il n'y a aucune stratégie de codage de la parole meilleure sur tous les plans, une optimisation de la compréhension des patients reste toujours à chercher.

Notre travail explore les potentialités de fournir plus de détails de la parole en les transmettant à l'implant sous formes de paramètres adéquats. Ceci est à assurer via des algorithmes appropriés. Pour cela, notre apport consiste à reprendre la démarche d'extraction des fréquences-clés de la parole tout en améliorant leur estimation. Dans ce contexte, une nouvelle méthode est inspirée de la norme GSM des cellules téléphoniques, en opposition à l'utilisation des détecteurs de passage par zéro comme c'est le cas dans l'implant Nucleus. Notre extraction des formants se poursuit jusqu'au quatrième formant, pour une meilleure perception des consonnes.

Comme les implants cochléaires Nucleus, Clarion, et Med-El, notre implant incorpore un système de transmission transcutanée pour connecter le processeur vocal externe avec le récepteur-stimulateur implanté.

CHAPITRE 3:

Alimentation RF transdermale

Dans le domaine des implants médicaux, les piles d'alimentation sont à remplacer périodiquement, nécessitant souvent pour cela une opération chirurgicale. Si des piles rechargeables sont employées, le moyen de les recharger s'effectue à l'aide d'un champ électromagnétique externe, évitant ainsi de nombreux désagréments aux patients.

Il est évident qu'un type d'alimentation par champ électromagnétique, seul et sans recours aux accumulateurs rechargeables, est particulièrement intéressant pour des implantations qui requièrent une faible puissance pour fonctionner. Ceci, réduit également l'encombrement de l'implant en supprimant les accumulateurs d'alimentation.

L'alimentation par champ électromagnétique est l'unique solution envisagée dans le cas des implants cochléaires, si on laisse de côté les anciennes alimentations percutanées [39, 53, 54, 55]. Ainsi, en communication de donnée avec l'implant ou en alimentation continue de celui-ci par champ externe, un transfert effectif d'énergie radio fréquence est requis. Ceci assure à l'implant une puissance d'alimentation suffisante et un meilleur signal sur bruit ou faible taux d'erreur des données reçues.

Le but de ce chapitre est donc d'étudier un transfert d'énergie RF optimisé, entre une bobine externe (émettrice) et une bobine implantée (réceptrice) afin d'alimenter l'implant cochléaire d'une part, et assurer la transmission de données d'autre part.

Une méthode analytique du champ rayonné par une bobine, à travers l'air libre puis à travers un tissu biologique est décrite. Deux types d'atténuation sont étudiés, afin d'évaluer la puissance disponible à l'implant : l'une liée à l'absorption du tissu biologique et l'autre au désalignement éventuel des bobines de transfert.

3.1. Transfert de puissance par couplage magnétique

Il est nécessaire de connaître l'atténuation apportée à l'énergie électromagnétique le long de sa propagation, commençant par le positionnement relatif des bobines et passant par le tissu biologique jusqu'au lieu d'implantation. Cette connaissance indique si l'énergie rayonnée est suffisante à l'alimentation par champ externe, du système implanté.

3.1.1. Calcul du potentiel magnétique d'une bobine circulaire

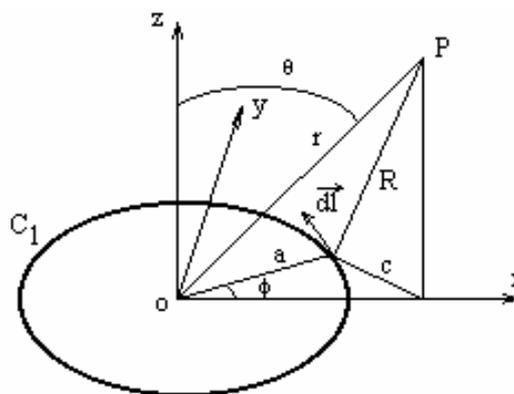


Figure 3-1: Potentiel magnétique à un point P d'une bobine

Soit une bobine coplanaire C_1 , à N_1 spires circulaires de rayon a et parcourue par un courant variable i , figure 3-1. Le potentiel magnétique en un point P , dans l'espace entourant la bobine est donné par [39, 69] :

$$\vec{A} = \frac{\mu_0}{4\pi} \oint_{C_1} \frac{N_1 \cdot i}{R} \cdot d\vec{l} \quad 3.1$$

où $\mu_0 = 4 \pi \cdot 10^{-7}$ H/m est la perméabilité de l'espace libre.

Par simplicité, qui ne modifie pas la généralité des résultats, on choisit le point P tel que:

$$\vec{r} = r \vec{u}_r = x \vec{u}_x + z \vec{u}_z \quad 3.2$$

$$R = \sqrt{z^2 + c^2} = \sqrt{z^2 + x^2 + a^2 - 2ax \cos \phi} = \sqrt{r^2 + a^2 - 2ar \sin \theta \cos \phi} \quad 3.3$$

$$d\vec{l} = a d\phi \vec{u}_\phi = a (-\vec{u}_x \sin \phi + \vec{u}_y \cos \phi) d\phi \quad 3.4$$

Pour une bobine considérée électriquement petite à la fréquence ω de travail, de façon à ce que le courant i est le même, en amplitude et phase, le long de C_1 et en tout point de la bobine, l'expression 3.1 devient alors :

$$\vec{A} = \frac{\mu_0 a N_1 i}{4\pi} \int_0^{2\pi} \frac{[\vec{u}_y \cos \phi - \vec{u}_x \sin \phi] d\phi}{\sqrt{r^2 + a^2 - 2ar \sin \theta \cos \phi}} \quad 3.5$$

Soit $\phi = \varphi - \pi$, $d\phi = d\varphi$, avec φ variant de $-\pi$ à $+\pi$ pour une variation de ϕ allant de -2π à 0 , mais pouvant se ramener au même intervalle de variation $[0$ à $2\pi]$.

Ce qui donne : $\cos \phi = \cos(\varphi - \pi) = -\cos \varphi$, et la première intégrale s'écrit alors :

$$I_1 = \int_0^{2\pi} \frac{\cos \phi d\phi}{\sqrt{r^2 + a^2 - 2ar \sin \theta \cos \phi}} = \int_{-\pi}^{\pi} \frac{-\cos \varphi d\varphi}{\sqrt{r^2 + a^2 + 2ar \sin \theta \cos \varphi}} \quad 3.6$$

Comme $\cos \varphi = 1 - 2 \sin^2(\varphi/2)$, ce qui rend le dénominateur de (3.6) égal à :

$$D = \sqrt{r^2 + a^2 + 2ar \sin \theta - 4ar \sin \theta \sin^2(\varphi/2)} = \sqrt{r^2 + a^2 + 2ar \sin \theta} \cdot \sqrt{1 - k^2 \sin^2(\varphi/2)} \quad 3.7$$

avec :

$$k = \sqrt{\frac{4ar \sin \theta}{r^2 + a^2 + 2ar \sin \theta}} \leq 1 \quad 3.8$$

L'intégrale I_1 devient :

$$I_1 = \int_0^{2\pi} \frac{(2 \sin^2(\varphi/2) - 1) d\varphi}{\sqrt{r^2 + a^2 + 2ar \sin \theta} \cdot \sqrt{1 - k^2 \sin^2(\varphi/2)}} = \frac{k}{2\sqrt{ar \sin \theta}} \int_0^{2\pi} \frac{(2 \sin^2(\varphi/2) - 1) d\varphi}{\sqrt{1 - k^2 \sin^2(\varphi/2)}} \quad 3.9$$

en multipliant (3.9) par k/k , et en faisant entrer le k^2 résultant dans le numérateur, ce dernier devient:

$$N = 2 k^2 \sin^2(\varphi/2) - k^2 + 2 - 2 = 2 - k^2 - 2 [1 - k^2 \sin^2(\varphi/2)] \quad 3.10$$

L'intégrale I_1 se scinde donc en deux autres intégrales :

$$I_1 = \frac{1}{2k\sqrt{ar \sin \theta}} \left[(2 - k^2) \int_0^{2\pi} \frac{d\varphi}{\sqrt{1 - k^2 \sin^2(\varphi/2)}} - 2 \int_0^{2\pi} \sqrt{1 - k^2 \sin^2(\varphi/2)} d\varphi \right] \quad 3.11$$

Les deux formes d'intégrales obtenues, sont appelées intégrales elliptiques de 1^{ère} et 2^{ème} espèce [21] et sont paramétrées par k et l'intervalle d'intégration ψ :

- intégrale de 1^{ère} espèce : $F(\psi, k) = \int_0^\psi \frac{d\varphi}{\sqrt{1 - k^2 \sin^2 \varphi}}$

- intégrale de 2^{ème} espèce : $E(\psi, k) = \int_0^\psi \sqrt{1 - k^2 \sin^2 \varphi} d\varphi$

Dans l'expression 3.11, l'intégration de 0 à 2π peut se ramener au double de l'intégration de 0 à π , et devient :

$$I_1 = \frac{1}{\sqrt{ar \sin \theta}} \left[\left(\frac{2}{k} - k \right) F(\pi, k) - \frac{2}{k} E(\pi, k) \right] \quad 3.12$$

La fonction $G(n, \varphi, k) = (1 - k^2 \sin^2 \varphi)^n$ est tracée pour φ allant de 0 à π , et ce pour $n = \pm 1/2$, figure 3-2.

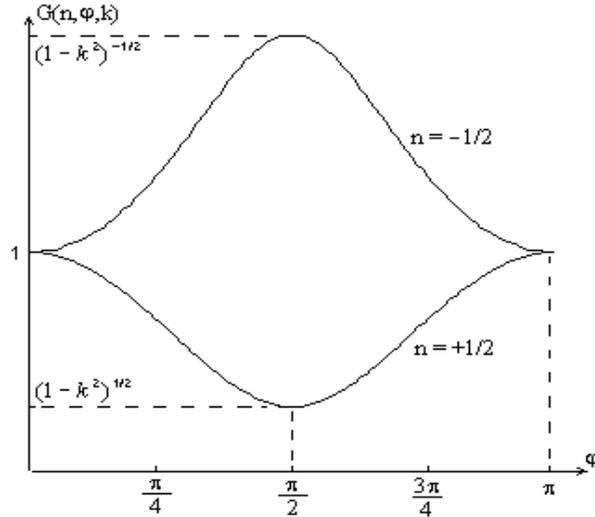


Figure 3-2: Une période de $G(n, \varphi, k)$, avec $n = \pm 1/2$

En exploitant les symétries de $G(n, \varphi, k)$, les intégrales $F(\pi, k)$ et $E(\pi, k)$ qui ne sont autres que les intégrales de la fonction $G(n, \varphi, k)$, peuvent s'écrire alors :

$$\begin{cases} F(\pi, k) = \int_0^{\pi} G(-\frac{1}{2}, \varphi, k) d\varphi = 2 \int_0^{\pi/2} G(-\frac{1}{2}, \varphi, k) d\varphi = 2 F(\frac{\pi}{2}, k) \\ E(\pi, k) = \int_0^{\pi} G(+\frac{1}{2}, \varphi, k) d\varphi = 2 \int_0^{\pi/2} G(+\frac{1}{2}, \varphi, k) d\varphi = 2 E(\frac{\pi}{2}, k) \end{cases}$$

L'expression 3.12 devient donc:

$$I_1 = \frac{2}{\sqrt{ar \sin \theta}} \left[\left(\frac{2}{k} - k \right) F\left(\frac{\pi}{2}, k\right) - \frac{2}{k} E\left(\frac{\pi}{2}, k\right) \right] = \frac{2}{\sqrt{ar \sin \theta}} S(k) \quad 3.13$$

où $S(k)$ est représenté en figure 3-3.

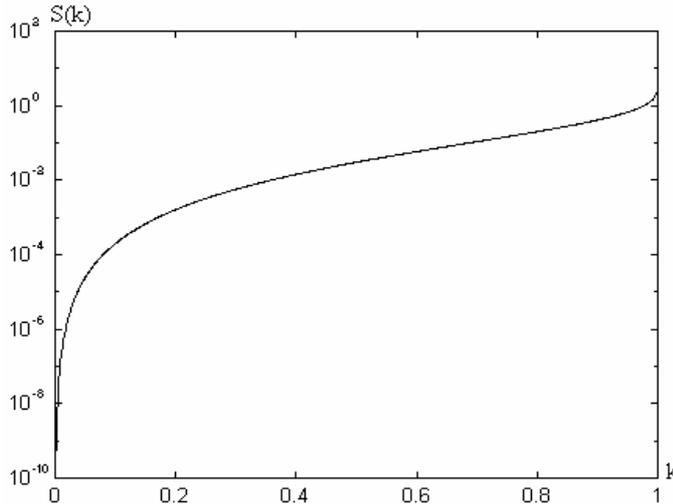


Figure 3-3: Courbe de S en fonction de k

En posant $-\sin \phi d\phi = d(\cos \phi)$ pour le deuxième intégral en (3.5), on obtient :

$$I_2 = \int_0^{2\pi} \frac{d(\cos \phi)}{\sqrt{r^2 + a^2 - 2ar \sin \theta \cos \phi}} = \frac{-1}{ar \sin \theta} \left[\sqrt{r^2 + a^2 - 2ar \sin \theta \cos \phi} \right]_0^{2\pi} = 0 \quad 3.14$$

Connaissant les résultats (3.13) et (3.14), et tenant compte de la symétrie en ϕ du potentiel magnétique obtenu, l'expression 3.5 devient :

$$\vec{A} = \frac{\mu_0 a N_1 i}{4\pi} \cdot \frac{2 S(k)}{\sqrt{ar \sin \theta}} \vec{u}_Y \Rightarrow \vec{A} = \frac{\mu_0 \sqrt{a} N_1 i}{2\pi} \cdot \frac{S(k)}{\sqrt{r \sin \theta}} \vec{u}_\phi \quad 3.15$$

3.1.2. Couplage entre bobines désalignées

On met à proximité de la première bobine C_1 , une deuxième bobine coplanaire C_2 à N_2 spires circulaires de rayon b . Dans le cas général, les deux bobines ne sont pas parfaitement alignées, et un désalignement quelconque des deux bobines couplées est représenté en figure 3-4. Il englobe les deux cas de désalignement, par translation Δ et rotation ψ .

Le repère $o'x'y'z'$, relatif à la bobine C_2 , donne par projection de ses vecteurs unitaires dans le

repère absolu $oxyz$, lié à la bobine C_1 :

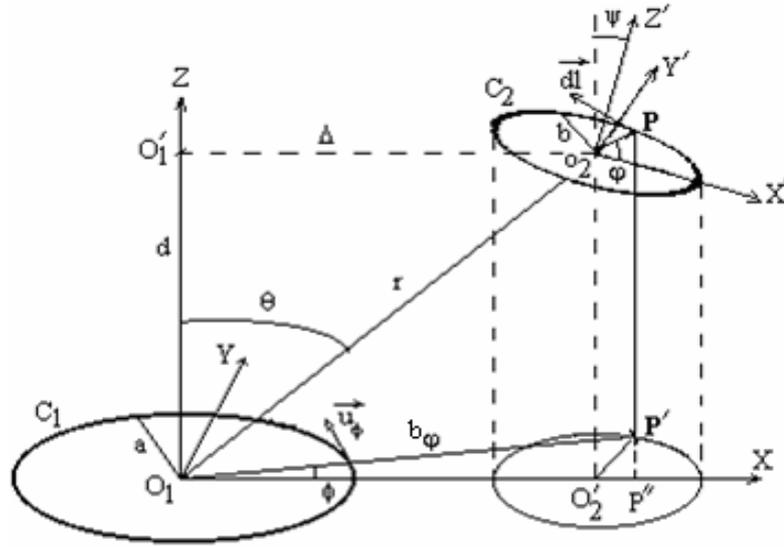
$$\begin{cases} \vec{u}_{X'} = \cos \psi \vec{u}_X - \sin \psi \vec{u}_Z \\ \vec{u}_{Y'} = \vec{u}_Y \end{cases}$$


Figure 3-4: Bobines couplées avec désalignement quelconque

Le flux ϕ_m du champ magnétique \vec{B} à travers la surface S de la bobine C_2 , peut s'obtenir à l'aide de la circulation du vecteur potentiel magnétique \vec{A} le long du périmètre de cette bobine:

$$\phi_m = \iint_S \vec{B} \cdot d\vec{s} = \oint_C \vec{A} \cdot d\vec{l} \quad 3.16$$

En utilisant le potentiel magnétique \vec{A} généré par la bobine C_1 et établi en (3.15), le flux magnétique à travers la bobine C_2 , s'écrit alors :

$$\phi_m = \frac{\mu_0 \sqrt{a} N_1 N_2 i}{2\pi} \oint_{C_2} \frac{S(k)}{\sqrt{r \sin \theta}} \vec{u}_\phi \cdot d\vec{l} \quad 3.17$$

Les vecteurs curvilignes élémentaires sont :

$$\begin{cases} d\vec{l} = b d\phi (-\sin \phi \vec{u}_{X'} + \cos \phi \vec{u}_{Y'}) = b d\phi (-\sin \phi \cos \psi \vec{u}_X + \cos \phi \vec{u}_Y + \sin \phi \sin \psi \vec{u}_Z) \\ \vec{u}_\phi = -\sin \phi \vec{u}_X + \cos \phi \vec{u}_Y \end{cases}$$

$$d'où : \quad \vec{u}_\phi \cdot d\vec{l} = b(\cos \psi \sin \phi \sin \phi + \cos \phi \cos \phi) d\phi \quad 3.18$$

$$\vec{r} = \vec{O_1P} = \vec{O_1O'} + \vec{O'O_2} + \vec{O_2P} \quad 3.19$$

$$\text{avec : } \vec{O_2P} = b \cos \phi \vec{u}_{X'} + b \sin \phi \vec{u}_{Y'} = b \cos \psi \cos \phi \vec{u}_X + b \sin \phi \vec{u}_Y - b \sin \psi \cos \phi \vec{u}_Z$$

$$\text{ainsi : } \vec{r} = (\Delta + b \cos \psi \cos \phi) \vec{u}_X + b \sin \phi \vec{u}_Y + (d - b \sin \psi \cos \phi) \vec{u}_Z \quad 3.20$$

Les projections de \vec{r} dans le plan O_1XY d'une part, et sur l'axe O_1Z d'autre part, donnent :

$$\begin{cases} r \sin \theta = \sqrt{(\Delta + b \cos \psi \cos \varphi)^2 + (b \sin \varphi)^2} = b_\varphi \\ r \cos \theta = d - b \sin \psi \cos \varphi \end{cases} \quad 3.21$$

En utilisant l'expression 3.20 et en mettant (3.21) dans (3.8), celle-ci devient :

$$k = \sqrt{\frac{4 a b_\varphi}{(a + b_\varphi)^2 + (d - b \sin \psi \cos \varphi)^2}} \quad 3.22$$

Le vecteur $\overrightarrow{O_2P}$, projeté dans le plan O_1XY , donne $\overrightarrow{O_2'P'}$:

$$\overrightarrow{O_2'P'} = b \cos \psi \cos \varphi \vec{u}_x + b \sin \varphi \vec{u}_y = \overrightarrow{O_2'P''} + \overrightarrow{P''P'}$$

Ce qui permet d'écrire :

$$\begin{cases} \sin \phi = \frac{\|\overrightarrow{P''P'}\|}{b_\varphi} = \frac{b \sin \varphi}{b_\varphi} \\ \cos \phi = \frac{\|\overrightarrow{O_1O_2'}\| + \|\overrightarrow{O_2'P''}\|}{b_\varphi} = \frac{\Delta + b \cos \psi \cos \varphi}{b_\varphi} \end{cases} \quad 3.23$$

Tenant compte de (3.21) et des expressions 3.23, introduites dans (3.18); le flux magnétique (3.17) devient :

$$\phi_m = \frac{\mu_0}{2\pi} N_1 N_2 i b \sqrt{a} \int_0^{2\pi} \frac{S(k)}{\sqrt{b_\varphi}} \left[\frac{b \cos \psi \sin^2 \varphi + (\Delta + b \cos \psi \cos \varphi) \cos \varphi}{b_\varphi} \right] d\varphi \quad 3.24$$

L'expression originale du coefficient d'inductance mutuelle s'écrit donc :

$$M = \frac{\phi_m}{i} = \frac{\mu_0}{2\pi} N_1 N_2 b \sqrt{a} \int_0^{2\pi} \frac{S(k)}{b_\varphi^{3/2}} (\Delta \cos \varphi + b \cos \psi) d\varphi \quad 3.25$$

Dans le cas d'alignement parfait des bobines ($\Delta = 0$, $\psi = 0$), nous avons $b_\varphi = b$, et (3.22) devient aussi une constante:

$$k = k_0 = \sqrt{\frac{4 ab}{(a + b)^2 + d^2}} \quad 3.26$$

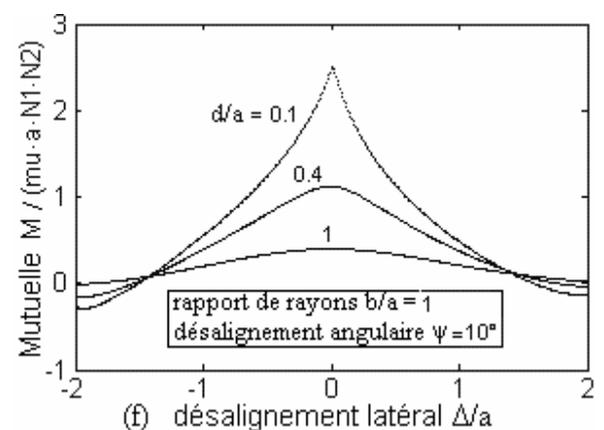
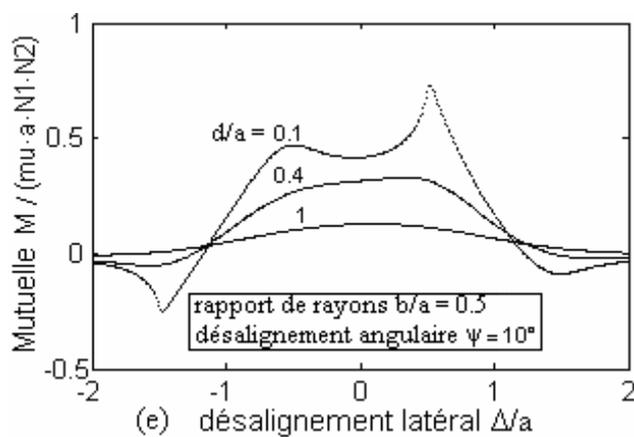
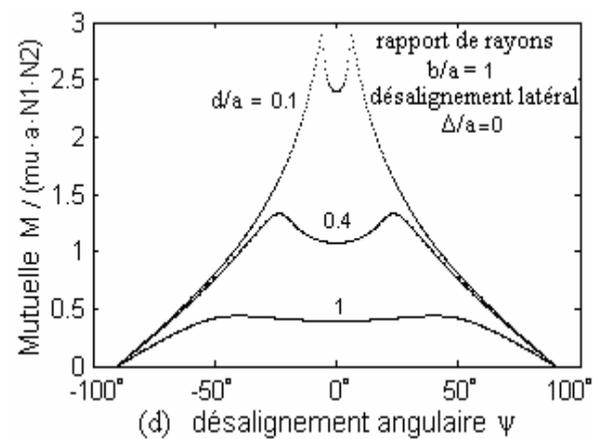
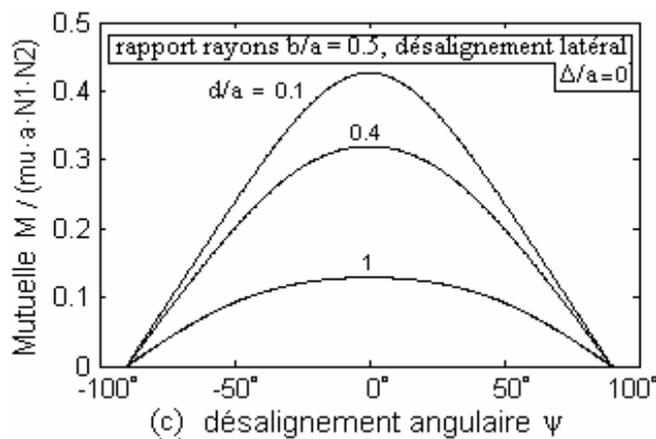
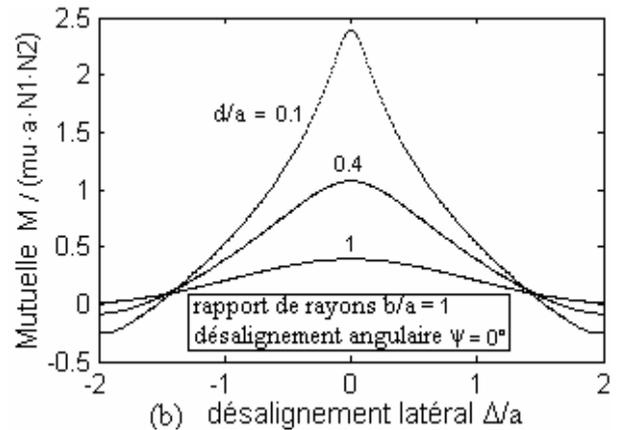
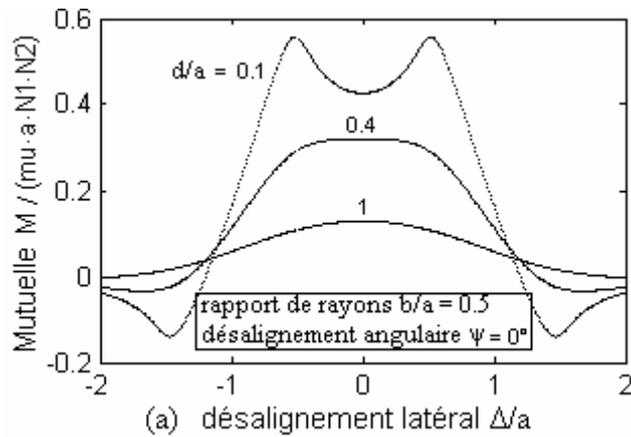
L'intégration de (3.25) devient simple et est donnée, comme dans [85, 88], par :

$$M_0 = \mu_0 N_1 N_2 \sqrt{a b} S(k_0) \quad 3.27$$

Dans le cas général avec désalignements combinés, l'intégrale du coefficient d'inductance mutuelle (3.25) peut être réalisée numériquement ou analytiquement de façon approximative. Dans les résultats suivants, une intégration numérique est adoptée afin d'obtenir une variété de courbes d'inductance mutuelle normalisée $M / (\mu_0 a N_1 N_2)$ en fonction du désalignement angulaire ψ (en degré) ou en fonction du désalignement latéral réduit Δ / a . Les figures les plus représentatives sont données en figure 3-5, où dans chacune d'elles, M est tracée sous trois distances longitudinales inter-bobines: $d/a = 0.1, 0.4$, et 1 . Dans les figures 3.5a et 3.5b, M est tracée en fonction du désalignement latéral uniquement, et ce, sous différents rapports de rayon des bobines, 0.5 et 1 , respectivement. Dans les figures 3.5c et 3.5d, M est tracée en fonction du désalignement angulaire uniquement, et ce, sous différents rapports de rayon des bobines, 0.5 et 1 , respectivement. Les figures 3.5e à 3.5h sont tracées sous les mêmes conditions que

précédemment, à part que des désalignements latéraux et angulaires simultanés sont considérés. Ainsi, en désalignement latéral, la figure 3.5e devient dissymétrique relativement à la figure 3.5a. Ce n'est pas le cas de la figure 3.5f, relativement à la figure 3.5b, qui devient plus étroite. En désalignement angulaire, les figures 3.5g et 3.5h deviennent dissymétriques relativement aux figures 3.5c et 3.5d, respectivement.

Des valeurs élevées de M sont notées en figure 3.5g relativement à la figure 3.5c. Cependant, des valeurs basses sont notées en figure 3.5h relativement à la figure 3.5d.



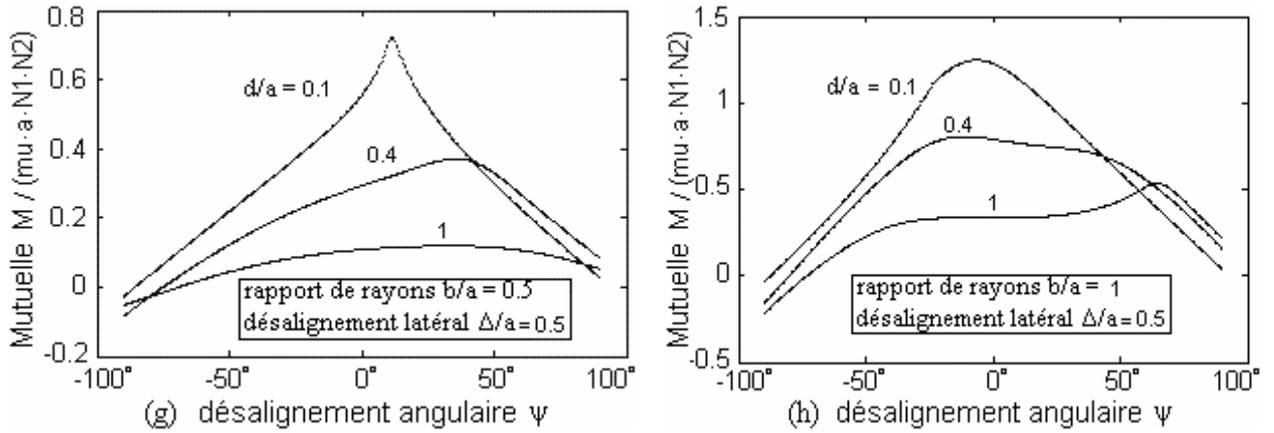


Figure 3-5: Inductance mutuelle normalisée en fonction des désalignements latéral (Δ/a) et angulaire (ψ) des bobines pour trois distances longitudinales inter-bobines (d/a)

3.1.3. Atténuation en puissance par désalignement des bobines

Soient L_1 et L_2 , les coefficients d'auto-induction des bobines de transmission-réception, et M , l'inductance mutuelle entre elles. En utilisant l'opérateur p de Laplace, la tension V_2 aux bornes de la bobine de réception, chargée par l'implant R , est la somme de la tension induite MpI_1 et la chute de tension aux bornes de la self L_2 causée par le courant induit I_2 . V_2 peut aussi être calculée à partir de la chute de tension RI_2 causée par l'implant, comme montré en figure 3-6.

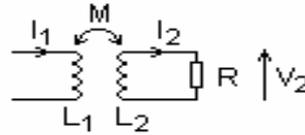


Figure 3-6: Bobines couplées et chargées par la résistance équivalente R de l'implant

$$\text{La tension reçue est : } V_2 = M p I_1 + L_2 p I_2 = R I_2 \quad 3.28$$

Qui devient en fonction du courant I_1 de la bobine de transmission :

$$V_2 = \frac{R M p}{R - L_2 p} I_1 \quad 3.29$$

Le facteur de transmission en puissance F_D du signal RF, due au désalignement des bobines, est le rapport des puissances reçues par l'implant en cas de bobines alignées et non alignées.

Ainsi, pour un même courant I_1 dans la bobine émettrice et la même impédance de charge (implant) présentée à la bobine réceptrice, on obtient :

$$F_D = \frac{V_2^2 \text{ avec désalignement}}{V_2^2 \text{ sans désalignement}} = \left(\frac{M}{M_0} \right)^2 \quad 3.30$$

$$\text{Avec (3.25) et (3.27), il vient : } F_D = \frac{b}{4\pi^2 S^2(k_0)} \left[\int_0^{2\pi} \frac{S(k)}{b^{3/2}} (\Delta \cos \varphi + b \cos \psi) d\varphi \right]^2 \quad 3.31$$

Cette formule originale donne le facteur de transmission en puissance sous désalignements combinés en translation et rotation des bobines de transfert RF.

L'atténuation en puissance dans ce cas, s'écrit simplement : $A_D = 1 - F_D$

3.2. Interaction des ondes électromagnétiques avec le tissu biologique

3.2.1. Propriétés électriques des tissus biologiques

L'utilisation d'une alimentation sans accumulateurs, basée sur l'énergie radiofréquence

transmise entre deux bobines, oblige à étudier le comportement du tissu biologique vis-à-vis de ces ondes électromagnétiques. On déduira par conséquent, l'atténuation en puissance causée par cette interaction.

Traversé par un champ électromagnétique, un tissu biologique se comporte comme un milieu avec pertes [18, 78, 79, 87]. En désignant respectivement par \vec{E} et \vec{H} les champs, électrique et d'induction magnétique complexes, variants à la fréquence ω ; la loi d'Ampère s'écrit alors [69] :

$$\vec{\text{rot}} \vec{H} = \vec{J} = \vec{J}_C + \vec{J}_D \quad 3.32$$

\vec{J} est la densité de courant, apparaissant suite à l'interaction du champ électromagnétique avec le tissu biologique. Cette densité de courant est constituée d'une composante dite de conduction ou Ohmique, responsable du mécanisme de perte d'énergie :

$$\vec{J}_C = \sigma \vec{E} \quad 3.33$$

et d'une composante dite de déplacement, représentant un stockage d'énergie :

$$\vec{J}_D = j\omega\epsilon \vec{E} \quad 3.34$$

σ est la conductivité du tissu biologique (inverse de la résistivité ρ), tandis que ϵ est sa permittivité diélectrique. Dans le plan complexe, ces deux courants définissent un angle de perte,

défini par :

$$\text{tg } \delta = \frac{|\hat{J}_C|}{|\hat{J}_D|} = \frac{\sigma}{\omega\epsilon} \quad 3.35$$

L'expression 3.32 peut donc s'écrire:

$$\vec{\text{rot}} \vec{H} = (\sigma + j\omega\epsilon) \vec{E} = j\omega\epsilon \left(1 - j \frac{\sigma}{\omega\epsilon}\right) \vec{E} = j\omega \hat{\epsilon} \vec{E} \quad 3.36$$

La permittivité complexe $\hat{\epsilon}$ caractérise donc un milieu avec des pertes de conduction, à qui, peuvent s'ajouter des pertes de polarisation du champ électromagnétique [69], aux fréquences atteignant les gigaHertz.

En utilisant (3.35), la permittivité complexe peut se mettre encore, sous la forme :

$$\hat{\epsilon} = \epsilon \left(1 - j \frac{\sigma}{\omega\epsilon}\right) = \epsilon (1 - j \text{tg } \delta) = \epsilon_0 (\epsilon' - j\epsilon'') \quad 3.37$$

avec la permittivité absolue : $\epsilon_0 = 8.8 \cdot 10^{-14}$ F/cm.

Les permittivités relatives: $\epsilon' = \epsilon / \epsilon_0$ et $\epsilon'' = \sigma / \omega\epsilon_0$ sont généralement mesurées indirectement à l'aide de ponts d'impédance RF. Elles sont fournies sous forme tablée en fonction de la fréquence, voir table 3-1, où les valeurs avec étoile sont des valeurs moyennes [79, 87].

Il existe des tissus biologiques équivalents ou fantômes, simulés par du gel et constitué de poudre de gélatine ajoutée par dose à de l'eau distillée bouillante. L'ajustage des valeurs de permittivité et conductivité qu'on veut atteindre, s'opère par addition progressive du chlorure de sodium NaCl [64].

tissus	T°	constante diélectrique ϵ'					constante diélectrique ϵ''						
		fréquence en MHZ						fréquence en MHZ					
		10	50	100	200	500	1000	10	50	100	200	500	1000
peau	37			65	57	47	44			140*	72	27	18*
graisse	37		12*	10*	6*	5*	6*		17*	12	6*	2,2	2*
sang	23			75*	69*	68*				216*	100*	49*	
muscle	37		91*	74*	56	53*	50*	1100*	282*		90*	41	23

Table 3-1: Constantes diélectriques selon la température et la fréquence de mesure

La permittivité et la conductivité des tissus biologiques sont dépendantes de la fréquence de l'onde électromagnétique incidente. Cette dépendance fréquentielle est caractérisée par la décroissance des constantes diélectriques avec la fréquence, table 3-1.

En conclusion, l'accroissement de la fréquence de l'onde électromagnétique incidente fait que le milieu organique stocke moins d'énergie et devient de plus en plus conducteur. Par conséquent, l'atténuation de l'onde sera plus prononcée.

3.2.2. Atténuation en puissance due au tissu biologique

Une nouvelle approche est adoptée ici, pour tenir compte de l'affaiblissement en puissance causé par la présence d'un tissu biologique entre deux bobines de transmission, couplées magnétiquement. La propagation du champ magnétique dans un milieu, défini par ses paramètres: permittivité ϵ , conductivité σ et perméabilité μ (celle du tissu biologique étant égale à celle de l'air μ_0 [87]), obéit à l'équation d'onde magnétique [69] :

$$\frac{1}{\mu_0} \nabla^2 \vec{B} = \sigma \frac{\partial \vec{B}}{\partial t} + \epsilon \frac{\partial^2 \vec{B}}{\partial t^2} \quad 3.38$$

où ∇^2 est le Laplacien.

Pour un champ magnétique à variation sinusoïdale : $\vec{B} = \vec{B}_0 e^{j\omega t} = \vec{B}_0$

l'équation (3.38) devient :

$$\nabla^2 \vec{B} = j\omega\mu_0(\sigma + j\omega\epsilon)\vec{B} = \hat{\gamma}^2 \vec{B} \quad 3.39$$

où $\hat{\gamma} \text{ (m}^{-1}\text{)} = \alpha + j\beta$ est la constante de propagation; α et β sont respectivement les constantes d'atténuation et de phase [69]. Nous démontrons que:

$$\alpha = \omega \sqrt{\frac{\mu_0 \epsilon}{2}} \left[\sqrt{1 + \text{tg}^2 \delta} - 1 \right]^{1/2} \quad 3.40$$

$$\beta = \omega \sqrt{\frac{\mu_0 \epsilon}{2}} \left[\sqrt{1 + \text{tg}^2 \delta} + 1 \right]^{1/2} \quad 3.41$$

avec $\text{tg} \delta$ donnée par (3.35) représentant les pertes du milieu.

Pour une propagation suivant l'axe oz par exemple, l'équation 3.39 donne :

$$\frac{d^2}{dz^2} \vec{B} = \hat{\gamma}^2 \vec{B} \rightarrow \vec{B}(z) = \hat{B}_0 e^{-\hat{\gamma}z} \vec{u}_z \quad 3.42$$

D'autre part, le milieu peut aussi être caractérisé par son impédance intrinsèque η , définie comme étant le rapport entre les deux champs de l'onde électromagnétique le traversant [69]:

$$\hat{\eta}(\Omega) = \frac{\hat{E}}{\hat{H}} = \frac{j\omega\mu_0}{\hat{\gamma}} \Rightarrow \frac{\hat{E}}{\hat{B}} = \frac{\hat{\eta}}{\mu_0} \quad 3.43$$

Cette impédance intrinsèque peut être décomposée en partie réelle et imaginaire :

$$\text{Avec :} \quad \hat{\eta} = \frac{\omega\mu_0}{\alpha^2 + \beta^2} (\beta + j\alpha) \quad 3.44$$

Dans son trajet, l'onde électromagnétique subit donc une réflexion partielle au contact de la peau. Le coefficient de transmission de l'onde, se propageant normalement du milieu 1 (air ambiant) au milieu 2 (tissu biologique), est défini par le rapport du champ transmis au champ incident.

En termes de champ électrique, cela devient [69]:

$$\hat{T}_E = \frac{\hat{E}_2}{\hat{E}_1} = \frac{2\hat{\eta}_2}{\hat{\eta}_1 + \hat{\eta}_2} \quad 3.45$$

Ce qui donne comme coefficient de transmission du champ magnétique, selon (3.43 et 3.44) :

$$\hat{T}_M = \frac{\hat{B}_2}{\hat{B}_1} = \frac{\hat{\eta}_1}{\hat{\eta}_2} \hat{T}_E = \frac{2\hat{\eta}_1}{\hat{\eta}_1 + \hat{\eta}_2} \quad 3.46$$

En utilisant les expressions (3.42, 3.45, et 3.46), le vecteur de Poynting [27] ou densité de puissance moyenne de l'onde transmise en fonction de celle incidente devient :

$$\bar{S}_2(\text{W/m}^2) = \frac{1}{2} \Re \left[\vec{E}_2 \wedge \frac{\vec{B}_2^*}{\mu_0} \right] = \frac{1}{2} \Re \left[\left(\hat{T}_E \vec{E}_1 e^{-\hat{\gamma}_2 z} \right) \wedge \left(\hat{T}_M^* \frac{\vec{B}_1^*}{\mu_0} e^{-\hat{\gamma}_2^* z} \right) \right] = \bar{S}_1 e^{-2\alpha_2 z} \Re \left[\hat{T}_E \hat{T}_M^* \right] \quad 3.47$$

Ce qui permet d'avoir le facteur de transmission en puissance de l'onde RF à travers un tissu biologique :

$$F_T = e^{-2\alpha_2 z} \Re \left[\hat{T}_E \hat{T}_M^* \right] = \frac{4e^{-2\alpha_2 z} \Re \left[\hat{\eta}_2 \hat{\eta}_1^* \right]}{|\hat{\eta}_1 + \hat{\eta}_2|^2} \quad 3.48$$

Les équations (3.44) permettent d'introduire, dans l'expression (3.48), les constantes de propagation :

$$F_T = \frac{4e^{-2\alpha_2 z} (\alpha_1 \alpha_2 + \beta_1 \beta_2)}{|\hat{\gamma}_1 + \hat{\gamma}_2|^2} \quad 3.49$$

L'expression du facteur de transmission en puissance de l'onde RF peut être généralisée au modèle de tissu à n couches biologiques superposées, d'épaisseurs d_i , et supposées avoir des caractéristiques très proches de façon à pouvoir considérer une adaptation entre elles :

$$F_T = \prod_{i=1}^n \frac{4e^{-2d_i \alpha_i} (\alpha_i \alpha_{i-1} + \beta_i \beta_{i-1})}{(\alpha_i + \alpha_{i-1})^2 + (\beta_i + \beta_{i-1})^2} \quad 3.50$$

L'atténuation en puissance dans ce cas, peut s'écrire : $A_T = 1 - F_T$

3.3. Confrontation avec l'expérience

Un banc expérimental, réalisé à l'Ensem-Cran de Nancy (en France) [64, 65], permet l'étude pratique du transfert d'énergie radio fréquence à 3 MHz par couplage magnétique.

Dans une expérience, un tissu biologique synthétique de 4cm de largeur, est inséré entre deux bobines couplées aux diamètres de 6.1 cm et 1.8 cm. Les caractéristiques de conductivité ($\sigma = 0.42 \text{ S.m}^{-1}$) et de permittivité relative ($\epsilon' = 120$) de ce tissu, sont étroitement celles du muscle humain. En manipulation, une puissance transmise de 10 W vers une charge de 2 k Ω , est utilisée. Les mesures de puissance reçue, sont effectuées à l'intérieur d'une cage Faraday sous différents états de désalignements et distances entre bobines, en l'absence puis en présence du muscle synthétisé. Ces puissances mesurées sont comparées à celles calculées à partir des expressions de facteurs de transfert en puissance établies, à savoir, les expressions de F_D et F_T .

Un bon accord est noté entre résultats pratiques et théoriques, concernant l'ordre de grandeur des valeurs numériques obtenues et les formes décroissantes d'atténuation en puissance [7, 8].

Par exemple, les facteurs de transfert d'énergie RF à travers le tissu biologique expérimental sont évalués à : F_T (mesurée) = 0.039, et F_T (calculée) = 0.045, soit une concordance de 87 %.

Ainsi, afin de stabiliser le rendement d'une connexion RF transcutanée en présence d'un désalignement possible entre bobines, il est nécessaire de satisfaire une bonne tolérance de désalignement. Par conséquent, l'évaluation d'une puissance suffisante à transmettre, peut être faite analytiquement par (3.31) ou graphiquement à partir des courbes d'inductance mutuelle comme en figure 3-5. En outre, l'amélioration des alimentations par champ doit aussi tenir compte de l'atténuation de la puissance apportée par les différentes couches du tissu biologique. Dans ce contexte, l'expression analytique approchée (3.50) peut servir, si l'on connaît les constantes de propagation des différentes couches constituant le tissu biologique.

3.4. Application au cas d'un implant cochléaire

3.4.1. Alimentation de l'implant par champ externe

Le système d'alimentation par champ externe à l'implant cochléaire est schématisé en figure 3-7. L'énergie RF, est amplifiée avant de l'appliquer à la bobine accordée du primaire. Le secondaire implanté, également accordé, est suivi d'un redresseur monoalternance rapide D, d'une capacité de filtrage C_0 à minimum de pertes et d'une Zener interdisant toute augmentation dangereuse de la tension continue obtenue.

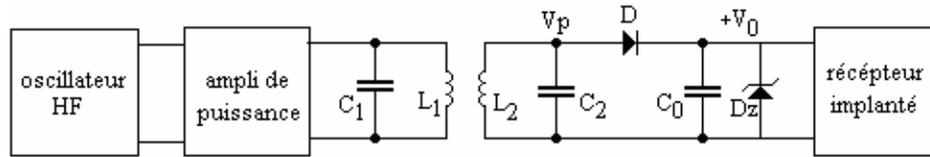


Figure 3-7: Principe d'alimentation transcutanée par ondes radio

Le choix des bobines L_1 et L_2 est aussi limité par le site d'implantation et ses contraintes anatomiques :

- Lieu d'implantation : os temporal (derrière l'oreille).
- Contraintes :
 - rayon typique de la bobine interne : $b = 12.5$ mm [53]
 - rayon maximal de la bobine externe : $a = 20$ mm [23]
 - distance maximale de séparation entre bobines : $d = 10$ mm [53].

Il a été précisé qu'entre 1 et 3 MHz, il y a un minimum d'absorption RF de la part des tissus biologiques [54], tout en assurant une miniaturisation de la circuiterie électronique implantée.

3.4.2. Recherche du coefficient d'induction mutuelle optimal

Nous avons réalisé un programme en Matlab, qui permet de chercher le maximum du coefficient d'induction mutuelle, normalisé par rapport aux nombres de spires: $M_0 = M / N_1 N_2$. La courbe obtenue, figure 3-8, donne un rayon optimal de la bobine externe $a = 18$ mm pour un rayon de la bobine interne $b = 12.5$ mm, se trouvant bien alignée à la distance de $d = 10$ mm. Ce rayon ne dépasse pas la valeur maximale admise (20 mm) pour un implant cochléaire, et il correspond à la valeur normalisée maximale de l'induction mutuelle $M_0 = 10.7$ nH.

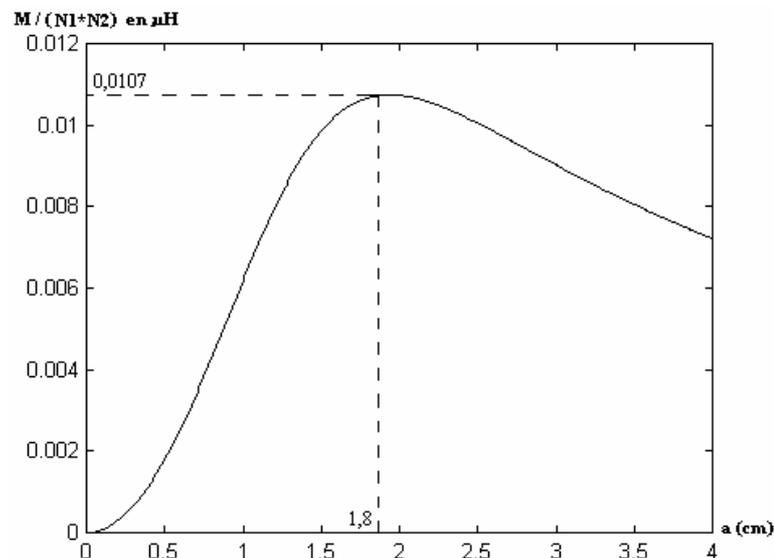


Figure 3-8: Courbe de $M/N_1 N_2$ en fonction du rayon a

D'autres programmes réalisés, nous permettent alors d'en déduire la tolérance en désalignement de ces deux bobines pour une diminution relative de l'inductance mutuelle M de 10 % de sa valeur maximale M_0 , soit $M=9.6$ nH.

Les courbes obtenues en figure 3-9 a et b, donnent les rangées d'alignement utiles:

- rangée axiale : $\Delta = 0.37 a = 6.6$ mm
- rangée angulaire : $\psi = 3\pi/16 = 33.75^\circ$

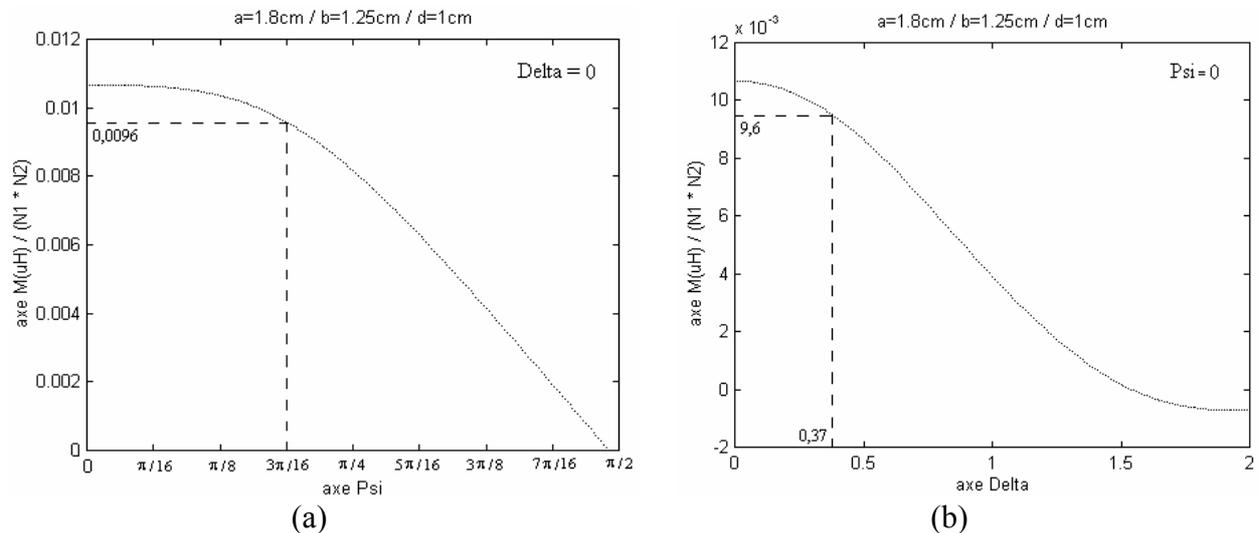


Figure 3-9: Inductance mutuelle en fonction du désalignement angulaire (a) et axial (b)

Mais pour les deux désalignements précédents réunis, l'inductance mutuelle résultante est fonction de leurs sens, figures 3-10 a et b :

- Δ et ψ de mêmes signes : $M = 11.7$ nH , c-à-d une augmentation de 9.3 % due à une compensation partielle des deux effets de désalignement.
- Δ et ψ de signes contraires : $M = 7.2$ nH , c-à-d une diminution de 32.7 % due à une addition des deux effets de désalignement.

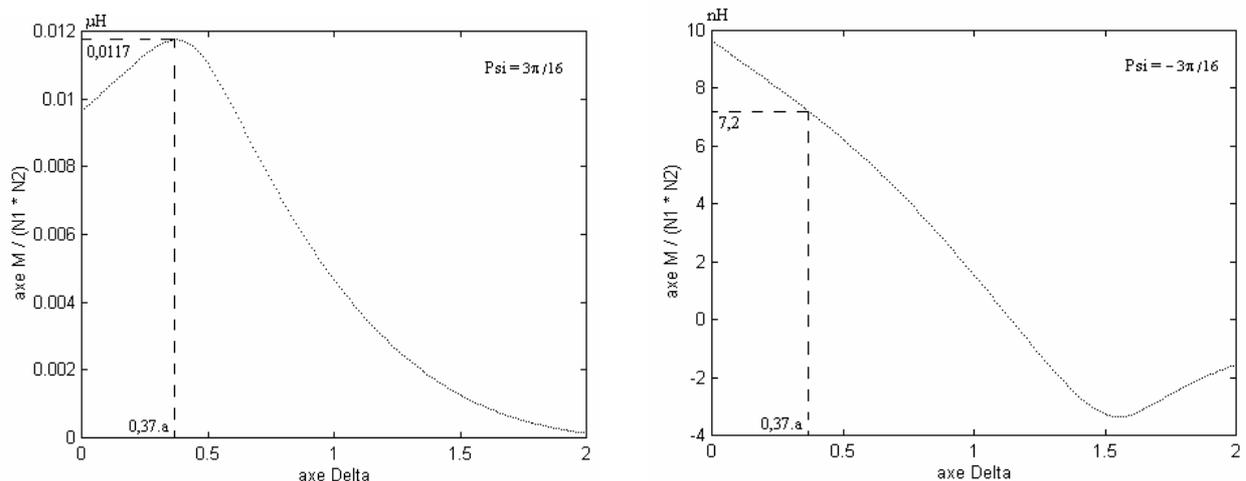


Figure 3-10: M en fonction du désalignement latéral avec $\Psi > 0$ en (a) et $\Psi < 0$ en (b)

Pour ces deux cas, la valeur moyenne de M reste approximativement 9.6 nH, et le facteur de transmission en puissance sous le désalignement des bobines (3.31), est alors de $F_D = 45\%$.

3.4.3. Estimation de la perte due au tissu biologique

Dans le cas de l'implant cochléaire, le tissu biologique séparant les deux bobines de

transmission est composé généralement de peau, graisse, sang, et muscle, et ce, de façon hétérogène. On supposera dans la suite, que le tissu biologique est formé de ces quatre ensembles mais non mélangés entre eux, c'est à dire de quatre couches d'épaisseurs égales, d_0 , homogènes et empilées le long d'une distance maximale $d = 4 d_0 = 1\text{cm}$, figure 3-11.

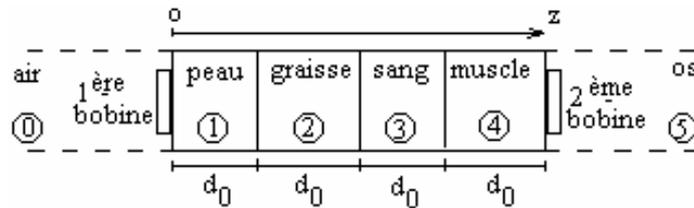


Figure 3-11: Séparation des bobines par du tissu biologique

Pour évaluer cette fois-ci, l'effet d'atténuation de l'énergie RF par ce tissu biologique, on utilisera les expressions (3.37), (3.40), et (3.41).

Pour cela, les angles de perte sont à calculer selon (3.35) qui devient : $\text{tg}\delta = \frac{\sigma}{\omega\epsilon} = \frac{\epsilon''}{\epsilon'}$.

Les permittivités relatives ϵ' et ϵ'' sont à déduire de la table 3-1, par extrapolation jusqu'à 3 MHz. L'estimation des permittivités à 3 MHz est effectuée, à l'aide de Matlab, par représentation polynomiale de degré permettant de passer sans erreur par toutes les données de la table 3-1, d'où la table 3-2.

	estimation polynomiale de ϵ'	estimation polynomiale de ϵ''
Peau	$\epsilon'(f) = -1.02e-7 f^3 + 2.10^{-4} f^2 - 0.13 f + 76.25$ d'où: $\epsilon'(3 \text{ MHz}) = 75.86$	$\epsilon''(f) = -1.34e-6 f^3 + 24e-4 f^2 - 1.29 f + 247.39$ d'où: $\epsilon''(3 \text{ MHz}) = 243.52$
Graisse	$\epsilon'(f) = -2.4e-10 f^4 + 3.59e-7 f^3 - 1e-4 f^2 + -0.027 f + 13.64$ d'où: $\epsilon'(3 \text{ MHz}) = 13.56$	$\epsilon''(f) = 1.67e-10 f^4 - 4.36e-7 f^3 + 4e-4 f^2 - 0.156 f + 23.77$ d'où: $\epsilon''(3 \text{ MHz}) = 23.31$
Sang	$\epsilon'(f) = 10^{-4} f^2 - 0.1 f + 83.83$ d'où: $\epsilon'(3 \text{ MHz}) = 83.53$	$\epsilon''(f) = 0.0025 f^2 - 1.9 f + 381.5$ d'où: $\epsilon''(3 \text{ MHz}) = 375.81$
Muscle	$\epsilon'(f) = 1.05e-9 f^4 - 2.35e-6 f^3 + 18e-4 f^2 - 0.57 f + 115.26$ d'où: $\epsilon'(3 \text{ MHz}) = 113.57$	$\epsilon''(f) = -3.72e-10 f^4 + 2.24e-7 f^3 + 2e-4 f^2 - 30e-3 f + 1.385$ d'où: $\epsilon''(3 \text{ MHz}) = 1295.8$

Table 3-2: Polynômes d'approximation des permittivités relatives

Un programme réalisé, permet de calculer à 3 MHz, les différentes constantes relatives au modèle adopté du tissu biologique, table 3-3 :

	α (Np/m)	β (rad/m)	$ \eta$ (Ω)	T_M	T_E
0- air	0	0.06	377.94	1.90	-
1- peau	0.59	0.81	23.69	0.49	0.12
2- graisse	0.16	0.28	73.53	1.59	1.51
3- sang	0.77	0.96	19.26	1.30	0.42
4- muscle	1.53	1.67	10.48	-	0.71

Table 3-3: Paramètres électriques du tissu biologique

Le facteur de transmission de la puissance véhiculée par l'onde électromagnétique, est trouvé notablement atténué à $F_T = 8 \%$ (3.50), après une traversée d'un centimètre du tissu biologique séparant les deux bobines de transfert. Pour une puissance d'entrée P_e , la puissance arrivant à la

bobine réceptrice est donc $F_T P_e$.

D'autre part, Le facteur de transmission en puissance a été évaluée à $F_D = 45 \%$ sous un désalignement simultané des bobines d'environ $\Delta=7\text{mm}$ et $\psi=30^\circ$, jugés être maximaux dans le cas d'implant cochléaire. Ainsi, la puissance utile arrivant à l'implant est $P_S = F_D F_T P_e$, avec $F_D F_T = 3.6 \%$. Ce facteur de transmission global sous les deux atténuations combinées, par désalignement des bobines de transfert et absorption du tissu biologique, sera pris en compte lors du dimensionnement du système de transfert HF transdermal à bobines couplées.

3.4.4. Calcul optimal des circuits accordés de l'alimentation

Le récepteur de la figure 3-7, peut être remplacé par une résistance équivalente R . Pour une amplitude HF, V_p limitée à V_0 , une puissance de $V_p^2 / 2R$ est utilisée. En négligeant la chute de tension aux bornes de la diode D , on a $V_p = V_0$. Le circuit équivalent est donné en figure 3-12.

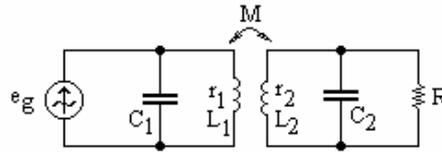


Figure 3-12: Circuit équivalent d'alimentation

Du fil en cuivre émaillé est utilisé dans les enroulements des bobines. En haute fréquence sa résistance est modifiée à cause des différentes pertes, notamment ceux de l'effet pelliculaire. On dit alors que la résistance du fil en HF est plus grande que sa résistance en continu. Des tables de coefficients correctifs de la valeur de la résistance, en fonction de la fréquence, existent [11]. On trouve pour une fréquence porteuse choisie égale à 2 MHz : $R_{AC} = 1.32 R_{DC}$.

De même, d'autres tables donnent la valeur de la résistance du cuivre émaillé en fonction de sa longueur et son diamètre [11]. Ainsi, pour un diamètre de 0.2mm choisi dans le bobinage des différentes selfs de notre montage, la résistance R_{AC} par mètre devient égale à 1.035 Ω/m . La résistance-série de L_1 s'écrit donc : $r_1 = 1,035.2\pi.a.n_1$, et celle de L_2 : $r_2 = 1,035.2\pi.b.n_2$.

Les bobines couplées seront réalisées aux dimensions optimales trouvées :

- rayon de la bobine externe L_1 : $a = 18\text{mm}$
- rayon de la bobine interne L_2 : $b = 12.5\text{mm}$.

Comme une self d'induction est proportionnelle au carré du nombre de spires, $L_1 = \alpha_1 n_1^2$ et $L_2 = \alpha_2 n_2^2$, les coefficients de proportionnalité mesurés avec des bobines d'essai réalisées aux dimensions précédentes, donnent: $\alpha_1 = 90 \text{ nH/spire}$, et $\alpha_2 = 70 \text{ nH/spire}$. Les capacités d'accord sont : $C_1 = 1 / L_1 \omega_0^2$ et $C_2 = 1 / L_2 \omega_0^2$.

Il est plus commode d'utiliser la représentation série des deux circuits accordés à ω_0 , comme illustré en figure 3-13. La résistance de charge en mode série devient égale à $R_L = R / Q^2$, où $Q = R / L_2 \omega_0$ est le coefficient de surtension du secondaire chargé par R .

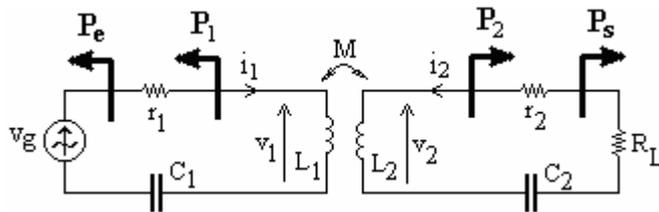


Figure 3-13: Représentation en série des circuits accordés

$$R_L \text{ peut encore s'écrire : } R_L = \frac{(L_2 \omega_0)^2}{R} \quad 3.51$$

Les coefficients de surtension, à la résonance, des deux bobines non chargées sont :

$$\begin{cases} Q_1 = \frac{L_1 \omega_0}{r_1} \\ Q_2 = \frac{L_2 \omega_0}{r_2} \end{cases} \quad 3.52$$

Le coefficient de qualité du secondaire chargé, est donc :

$$Q_S = L_2 \omega_0 / (r_2 + R_L) \quad 3.53$$

En utilisant des tensions et courants efficaces, la puissance développée au secondaire est :

$$P_2 = v_2 i_2 = - \left(r_2 + R_L + \frac{1}{jC_2 \omega_0} \right) i_2^2 \quad 3.54$$

La puissance délivrée à la charge s'écrit :

$$P_S = R_L i_2^2 = \frac{R_L}{r_2 + R_L + \frac{1}{jC_2 \omega_0}} P_2 \quad 3.55$$

Le système d'équation régissant les tensions aux bornes des selfs est donné par :

$$v_1 = jL_1 \omega_0 i_1 + jM \omega_0 i_2 \quad 3.56$$

$$v_2 = jL_2 \omega_0 i_2 + jM \omega_0 i_1 = - \left(r_2 + R_L + \frac{1}{jC_2 \omega_0} \right) i_2 = -(r_2 + R_L - jL_2 \omega_0) i_2 \quad 3.57$$

à partir de (3.57), on peut déduire la relation entre courants :

$$i_2 = - \frac{jM \omega_0}{r_2 + R_L} i_1 \quad 3.58$$

En reportant dans (3.56), il vient :

$$v_1 = \left(jL_1 \omega_0 + \frac{(M \omega_0)^2}{r_2 + R_L} \right) i_1 \quad 3.59$$

Ainsi, la charge du secondaire ramenée en série au primaire s'écrit :

$$R_e = \frac{(M \omega_0)^2}{r_2 + R_L} \quad 3.60$$

et le circuit équivalent devient comme en figure 3-14 :

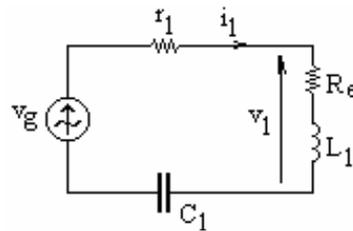


Figure 3-14: Charge ramenée au primaire

Le coefficient de qualité du primaire chargé, devient :

$$Q_P = L_1 \omega_0 / (r_1 + R_e) \quad 3.61$$

La charge ramenée en parallèle sur le primaire est alors :

$$R_c = Q_P^2 (r_1 + R_e) = L_1^2 \omega_0^2 / (r_1 + R_e) \quad 3.62$$

Comme la puissance émanant de la source à la résonance, est :

$$P_e = \frac{v_g^2}{r_1 + R_e} \quad 3.63$$

alors que celle fournie à la charge, ramenée au primaire, est :

$$P_1 = R_e i_1^2 = R_e \left(\frac{v_g}{r_1 + R_e} \right)^2 = \frac{R_e}{r_1 + R_e} P_e \quad 3.64$$

Le rapport de puissance entre secondaire et primaire devient :

$$\frac{P_2}{P_1} = \frac{v_2}{R_e i_2} \left(\frac{i_2}{i_1} \right)^2 = \frac{\left(r_2 + R_L + \frac{1}{jC_2 \omega_0} \right)}{R_e} \left(\frac{M \omega_0}{r_2 + R_L} \right)^2 = \frac{\left(r_2 + R_L + \frac{1}{jC_2 \omega_0} \right)}{r_2 + R_L} \quad 3.65$$

D'après les expressions (3.55), (3.64), et (3.65), une relation entre la puissance fournie par la source et celle parvenant à la charge, peut être déduite sous forme d'un rendement en puissance :

$$\eta = \frac{P_s}{P_e} = \frac{P_s}{P_2} \frac{P_2}{P_1} \frac{P_1}{P_e} = \frac{R_L R_e}{(r_2 + R_L)(r_1 + R_e)} \quad 3.66$$

En tenant compte des relations (3.51), (3.52), et (3.60), le rendement devient :

$$\eta = \frac{P_s}{P_e} = \frac{R M^2 \omega_0^2 Q_2^2 r_2}{(R + Q_2^2 r_2) [r_1 r_2 (R + Q_2^2 r_2) + R M^2 \omega_0^2]} \quad 3.67$$

Le rendement en puissance η , des bobines de transfert, est donc une fonction multivariable pouvant être optimisée en annulant sa différentielle totale [21]: $d\eta = 0$, ce qui revient à annuler toutes ses dérivées partielles.

En posant $M = N M_0$ où $N = N_1 N_2$, on constate que η est monotone en fonction de N et r_1 :

$$\frac{\partial \eta}{\partial N} \neq 0 \quad ; \quad \frac{\partial \eta}{\partial r_1} \neq 0$$

mais ne l'est pas en fonction des autres variables qui restent, à savoir, r_2 et Q_2 . Ces deux variables étant liées, il ne restera à optimiser η que par rapport à l'une d'elles, soit:

$$\frac{\partial \eta}{\partial (Q_2^2)} = 0 \quad \Rightarrow \quad Q_2^2 = \frac{R}{r_2} \sqrt{1 + Q^2} \quad 3.68$$

avec le coefficient :

$$Q = \frac{M \omega_0}{\sqrt{r_1 r_2}} \quad 3.69$$

D'où le rendement optimal en puissance (entre puissance fournie par la source P_e et celle parvenant à la charge P_s) :

$$\eta_{\text{opt}} = \left(\frac{Q}{1 + \sqrt{1 + Q^2}} \right)^2 \quad 3.70$$

En tenant compte du facteur de transmission du canal, le rendement final est :

$$\eta = P_s / P_e = F_D F_T \eta_{\text{opt}} \quad 3.71$$

Pour une puissance P_s nécessaire à l'implant, la puissance d'entrée requise est donc :

$$P_e = \frac{P_s}{F_D F_T} \frac{(1 + \sqrt{1 + Q^2})^2}{Q^2} \quad 3.72$$

En utilisant l'expression du coefficient d'induction mutuelle, $M = n_1 n_2 M_0$, et celle de la résistance série du primaire, $r_1 = 1,035.2\pi a n_1$; le carré de l'expression 3.69 devient :

$$Q^2 = \frac{(n_2 M_0 \omega_0)^2 n_1}{6.5 a r_2} \quad \text{de la forme } k_0 n_1 \quad 3.73$$

D'autre part, si V_c désigne l'amplitude de la porteuse amplifiée et disponible au primaire, figure 3-12, la résistance de charge au primaire R_c peut s'écrire, avec les expressions (3.72) et (3.73), comme :

$$R_c = \frac{V_c^2}{2P_e} = \frac{k_0 F_D F_T V_c^2}{2P_s} \frac{n_1}{(1 + \sqrt{1 + k_0 n_1})^2} \quad \text{de la forme } \frac{k_1 n_1}{(1 + \sqrt{1 + k_0 n_1})^2} \quad 3.74$$

La même résistance établie dans (3.62), devient après combinaison des expressions de M et r_1 ainsi que celle de l'induction du primaire, $L_1 = \alpha_1 n_1^2$, et les expressions (3.60) et (3.61) :

$$R_c = \frac{(\alpha_1 \omega_0)^2 n_1^3}{6.5 a + \frac{(n_2 M_0 \omega_0)^2 n_1}{r_2 + R_L}} \quad \text{de la forme } \frac{k_2 n_1^3}{k_3 + k_4 n_1} \quad 3.75$$

L'égalité entre (3.74) et (3.75) donne une équation du sixième degré en n_1 :

$$k_1^2 k_3^2 + 2k_1^2 k_3 k_4 n_1 + (k_1^2 k_4^2 - 4k_1 k_2 k_3) n_1^2 - 2k_1 k_2 (k_0 k_3 + 2k_4) n_1^3 - 2k_0 k_1 k_2 k_4 n_1^4 + k_0^2 k_2^2 n_1^6 = 0$$

Pour être réduite, la bobine implantée doit être de quelques spires seulement, soit $n_2 = 5$ spires. La tension Zener est choisie autour de 5V, alors que la tension de crête amplifiée de la porteuse qu'on espère recueillir au primaire est $V_c = 10V$ (afin de récupérer une alimentation bitension symétrique $\pm 5V$).

Un programme réalisé, permet le calcul et le choix des autres grandeurs à partir de représentations graphiques. Nous utilisons le transistor classique 2N2222 comme transistor d'amplification en puissance, vu ses bonnes caractéristiques en HF, mais sa puissance de dissipation maximale est de 0.5W. Ce qui limite la puissance P_p maximale au primaire à la même valeur, et une puissance maximale recueillie au secondaire à $P_s = 12mW$, figure 3-15. Nous limitons la puissance de dissipation du transistor et disponible au primaire aux environs de $P_p = 0.44W$, ce qui permet une puissance au secondaire de $P_s = 11mW$, soit $\pm 1.1mA$ sous $\pm 5V$.

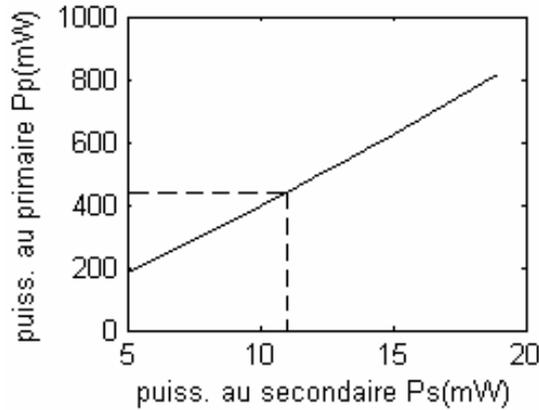


Figure 3-15: Puissance au primaire en fonction de la puissance au secondaire

Le nombre de spires au primaire n_1 et le rendement optimal de la liaison HF (sans atténuation du canal) sont tracés en fonction de la puissance P_s au secondaire, selon un nombre arrondi de spires ou non, figure 3-16. Ces deux grandeurs diminuent en fonction de P_s et valent à $P_s = 11 mW$: $n_1 = 7$ spires, et $\eta_{opt} = 69.44 \%$. Cette dernière valeur s'abaisse à 2.5 % si on tient compte de l'atténuation maximale du canal ($1 - F_D F_T = 96.4 \%$).

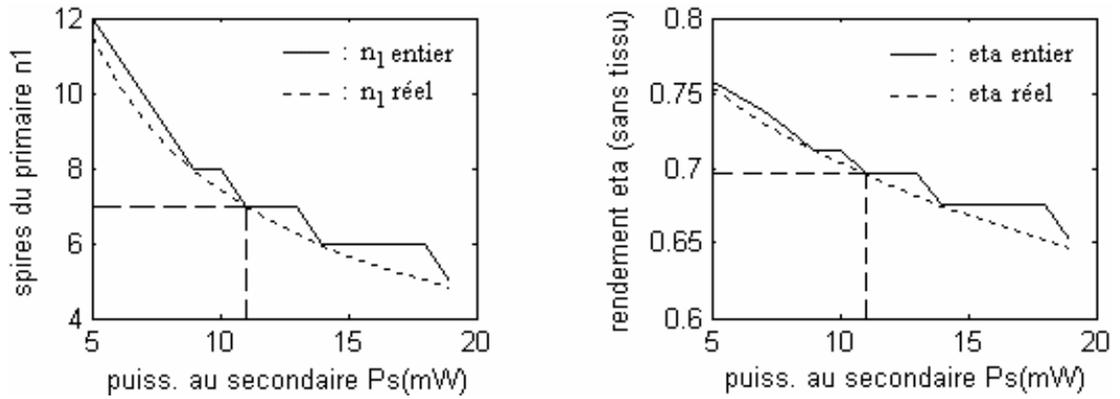


Figure 3-16: Nombre de spires du primaire et rendement optimal de la liaison en fonction de la puissance au secondaire

La charge R_c ramenée au primaire, ainsi que la charge résistive R équivalente du récepteur, sont également tracées en fonction de la puissance P_S du secondaire. Ces résistances valent $R_c = 236\Omega$ et $R = 1.14\text{ k}\Omega$, à la puissance du secondaire $P_S = 11\text{ mW}$, figure 3-17.

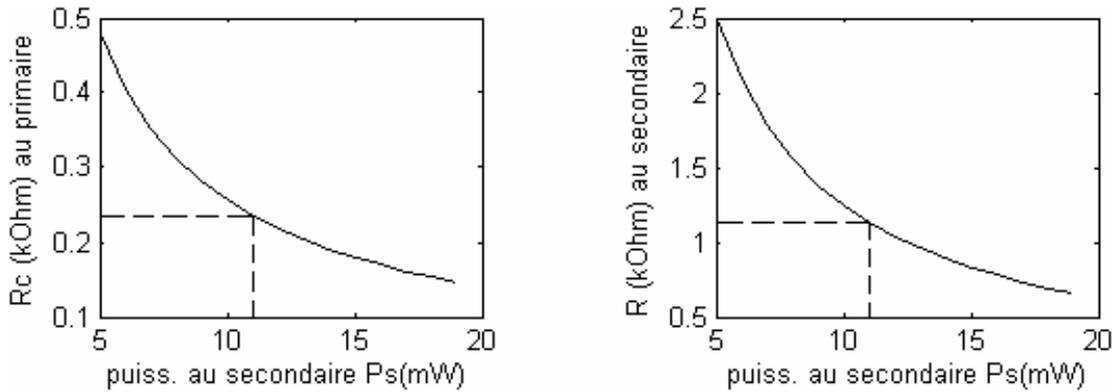


Figure 3-17 : Charge ramenée au primaire et charge du secondaire en fonction de la puissance au secondaire

Les coefficients de qualité du primaire et secondaire chargés, sont représentés en fonction de la puissance P_S du secondaire, figure 3-18. Leurs variations sont inversées l'une par rapport à l'autre, et à $P_S = 11\text{ mW}$ on a : $Q_P = 4.3$ et $Q_S = 26.4$. La bande passante à la réception (secondaire) est donc plus étroite qu'à l'émission (primaire) et détermine par conséquent, la bande passante de tout le système: $B = 2\text{ MHz} / 26.4 = 75.8\text{ kHz}$.

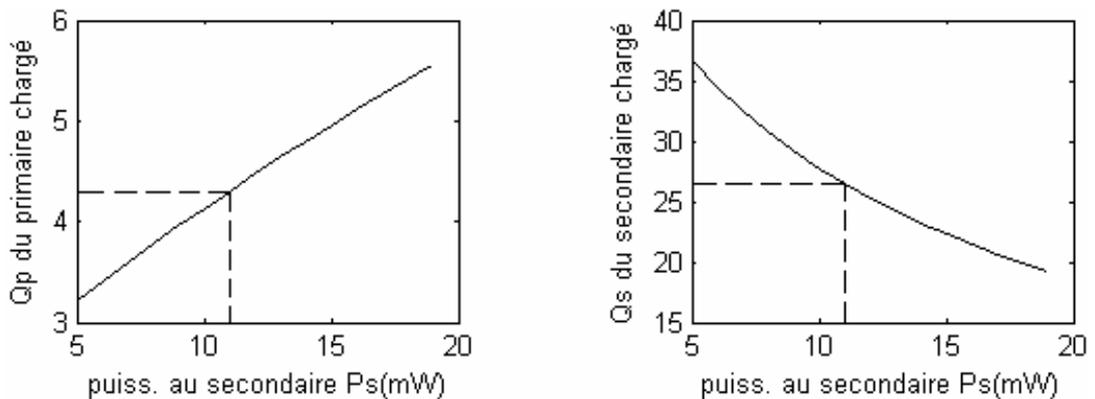


Figure 3-18: Coefficients de qualité du primaire et secondaire chargés en fonction de P_S

Les circuits d'accord calculés, sont donc : $C_1 = 1.5 \text{ nF}$ et $L_1 = 4.4 \text{ } \mu\text{H}$ à $n_1 = 7$ spires,
 $C_2 = 3.6 \text{ nF}$ et $L_2 = 1.8 \text{ } \mu\text{H}$ à $n_2 = 5$ spires.

Les bobines étant conçues manuellement, les capacités d'accord pratiques sont trouvées égales à $C_1 = 1 \text{ nF}$ et $C_2 = 3.3 \text{ nF}$.

Conclusion :

Dans le domaine biomédical, un transfert optimal de puissance vers les dispositifs implantés est requis. Ceci dans le but d'assurer un transfert suffisant en énergie/données, malgré l'atténuation due à l'absorption du tissu biologique et le désalignement éventuel des bobines de transfert.

Dans ce dernier cas, la bobine de transmission est maintenue en regard de la bobine de réception implantée, grâce à un aimant qui minimise leur désalignement.

Mais depuis peu, il existe des implants sans aimant permettant de réaliser des IRM avec succès, et le problème du désalignement réapparaît de nouveau.

L'analyse du transfert de puissance radio fréquence (RF), entre les deux bobines couplées, nous a permis d'évaluer la puissance réellement reçue. Pour cela, l'atténuation en puissance RF par un désalignement simultané, en translation et rotation des bobines de transfert, a pu être exprimée en une seule formule. Dans cette équation originale, différents diamètres de bobines sous une variété de désalignement latéral et angulaire, peuvent être considérés.

Une deuxième analyse portant sur l'atténuation en puissance RF, suite à une interaction avec le tissu biologique, a pu aboutir à une autre expression originale. Dans ce contexte, des modèles de tissu biologique à couches superposées plus ou moins complexes, peuvent être considérés sous une certaine approximation.

Ces deux analyses nous ont permis de choisir une meilleure conception du système de transfert à bobines, assurant une fluidité d'énergie sans rupture d'alimentation ou de données transmises aux circuits implantés.

CHAPITRE 4:

Analyse de la parole et sonagramme

Le transmetteur vers l'implant cochléaire, devant véhiculer l'énorme flux d'information de la parole, doit travailler en temps réel. D'où l'utilité de compresser l'information avant sa transmission.

De nombreux chercheurs ont été poussés à étudier et à comprendre le fonctionnement du système de production de la parole, et de mettre au point des outils mathématiques pour bien le décrire. Par exemple, une des hypothèses simples considère le signal de la parole comme étant une succession de signaux élémentaires stationnaires (sur des durées allant de 10 à 30 ms); les uns produits par les vibrations des cordes vocales, les autres étant par contre produits par un écoulement turbulent de l'air [14].

Pour l'analyse du signal vocal, la méthode classique utilise la représentation temps-fréquence en temps réel d'un signal audiofréquence afin de détailler simultanément ses informations fréquentielles et temporelles. Cela consiste à estimer son évolution spectrale au cours du temps en faisant appel à des outils tels que la transformée de Fourier à court terme ou glissante et la transformée en ondelettes.

D'autre part, les codeurs paramétriques permettent la transmission de la parole avec un débit réduit, ce qui passe tout naturellement par une prise en compte judicieuse des propriétés de production et de perception de la parole. Dans le souci de travailler avec des débits réduits du signal de la parole, une analyse approfondie de sa composition temporelle et surtout spectrale, s'impose afin de se limiter aux paramètres pertinents. Il s'agit alors de la compression de la parole pour des fins de transmission et de synthèse.

Depuis que les techniques de numérisation ont été rendues possibles par des composants électroniques plus complexes et plus rapides, la transmission à faible débit est devenue réalisable économiquement. Particulièrement, le codage par prédiction linéaire est une démarche séduisante tant par sa facilité de mise en œuvre que par sa capacité de description du signal vocal [14, 74].

Dans notre cas, deux différents outils d'analyse de la parole sont considérés pour aboutir à ses principaux paramètres. Des signaux de la parole préenregistrés, seront utilisés afin de garantir aux deux outils de traitement d'être alimentés par les mêmes échantillons d'entrée et de prétendre par conséquent, à une comparaison plus exacte de leurs sorties.

Le premier outil utilisé n'est qu'un sonagramme à large bande, écrit en turbo Pascal et s'exécutant sur un PC. Le deuxième outil, quant à lui, est un modèle prédictif de la parole, écrit en langage C et s'articulant sur une carte DSP.

4.1. Signal de la parole

4.1.1. Description phonatoire

La parole est un signal réel, continu, d'énergie finie et non stationnaire. Sa structure est complexe et variable avec le temps. Le signal de la parole couvre quasiment toute l'étendue du spectre audible. En pratique, on peut se limiter à la bande 50-5000 Hz.

L'unité élémentaire d'un son permettant la distinction des différents mots est le phonème. La représentation phonétique d'un texte dépend de la langue dans laquelle il est écrit.

Le tableau suivant donne une liste non exhaustive des phonèmes français [70].

PHONEMES								
VOYELLES		SEMI - VOYELLES	CONSONNES					
ORALES	NASALES		LIQUIDES	NASALES	FRICATIVES		OCCLUSIVES	
					voisées	non voisées	voisées	non voisées
[i]	[é]	[w]	[l]	[m]	[v]	[f]	[b]	[p]
[ô]	[ã]	[j]	[r]	[n]	[z]	[s]	[d]	[t]
[a]				[ŋ]	[ʒ]	[ʃ]	[g]	[k]
[u]								
[o]								
[y]								
[ɛ]								

Table 4-1: Exemple de phonèmes français

Pour extraire l'information par traitement de la parole, il faut connaître d'abord la structure de ce signal et la manière dont est codée l'information.

Le conduit vocal humain, commence par la glotte ou cordes vocales et se termine par les lèvres. En d'autres termes, il est constitué du pharynx et de la bouche ou cavité orale.

Plusieurs types de sons sont produits :

- Les sons voisés, où les cordes vocales vibrent (sans obstacle) à des oscillations de relaxation, excitant ainsi le conduit vocal par des impulsions d'air quasi-périodiques; exemple, les voyelles orales.
- les sons fricatifs, sont générés par un écoulement d'air turbulent à partir d'une constriction (fermeture partielle) en un certain point du conduit vocal; exemple, les consonnes fricatives. Les sons non voisés excitent le conduit vocal par un spectre ressemblant à celui d'un bruit.
- Les sons plosifs, sont créés par le relâchement brusque d'une pression maintenue par une occlusion (fermeture complète) en un certain point du conduit vocal; exemple, les consonnes occlusives.
- Enfin le conduit nasal est couplé acoustiquement au conduit vocal. Un son nasal est généré lorsque le conduit vocal est totalement fermé et que le rayonnement du son se fait par le nez; la cavité orale sert alors de cavité antirésonante capturant certaines fréquences; exemple, les consonnes nasales.

4.1.2. Caractéristiques de la parole

La parole se distingue des autres sons par des caractéristiques acoustiques ayant leurs origines dans les mécanismes de production. Les sons de la parole sont produits soit par les vibrations des cordes vocales (source de voisement), soit par l'écoulement turbulent de l'air dans le conduit vocal, ou lors du relâchement d'une occlusion de ce conduit (source de bruit).

Ainsi, tous les langages peuvent être décomposés en partie prosodique (rythme) et partie phonémique (prononciation), la partie phonémique étant divisée en son voisé (voyelles à formants dominants) et non voisé (consonnes ayant un spectre semblable au bruit) [14].

Un exemple de composition temporelle du signal de la parole est donné en figure 4-1, avec :

- Pseudo-périodique (D) pour sons voisés;
- Aléatoire (A) pour sons fricatifs;
- Impulsionnel (C) pour phase explosive des sons occlusifs. (B est du bruit).

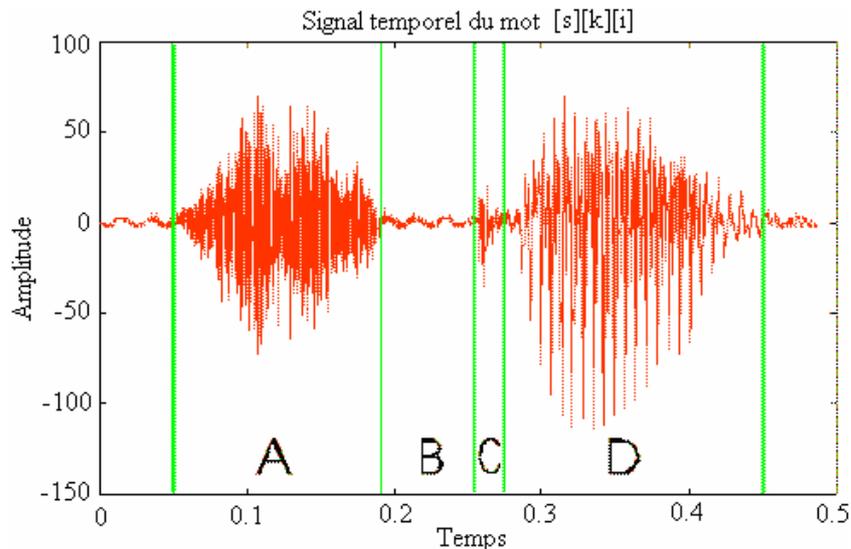


Figure 4-1: Représentation temporelle du signal acoustique de la parole

La durée d'une voyelle est de 200 à 300ms, et les consonnes ont des fréquences qui changent rapidement au bout de 20 ms [23].

L'intensité du son est liée à la pression de l'air en amont du larynx. La fréquence du voisement n'est rien d'autre que la fréquence du cycle d'ouverture/fermeture des cordes vocales, c'est le signal de la glotte. Le fondamental ou fréquence voisée (dit pitch) varie lentement. En général, il est limité à 120 Hz chez les hommes, et à 225 Hz chez les femmes [23].

Le spectre de la parole résulte du filtrage dynamique du signal glottique par le conduit vocal [74]. Les intensités de l'ensemble des harmoniques du son laryngé sont modulées par les résonateurs de la cavité pharyngo-buccale. Les fréquences de résonance du conduit vocal sont appelées formants [14], figure 4-2.

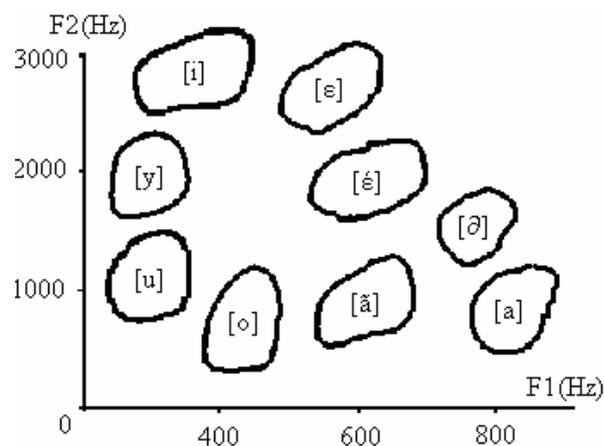


Figure 4-2: Les 2 formants des principales voyelles

Les informations linguistiques d'une voyelle sont essentiellement contenues dans les deux premiers formants, F1 et F2. La valeur d'un formant est définie par la fréquence correspondant à l'intensité maximum du formant. Cette valeur, assez constante chez le même individu, varie un peu d'un sujet à l'autre (d'où les zones entourées de la figure 4-2). Cette valeur particulière à chaque individu contribue à la reconnaissance des voix.

4.1.3. Modélisation de la parole

Plusieurs possibilités de représentations discrètes des signaux de la parole s'offrent

actuellement. Ces représentations peuvent être classées en deux groupes ; les représentations en forme d'onde et les représentations paramétriques.

Les représentations en forme d'onde, préservent l'enveloppe du signal analogique de la parole à travers des modulations numériques telles que PCM, LDM, etc...

Les représentations paramétriques, quant à elles, visent à représenter le signal de la parole comme la sortie d'un modèle de production vocale. Après échantillonnage et quantification de la parole, les paramètres du modèle sont extraits et sont, soit relatifs à la source sonore (paramètres dits d'excitation) soit relatifs aux signaux propres de la parole (paramètres liés au conduit vocal).

Parmi les importantes applications des représentations paramétriques de la parole est la réduction de la largeur de bande requise à sa transmission. Dans le cas de l'implant cochléaire, ce traitement de la parole offre une forme d'information plus adaptée à la stimulation électrique des fibres nerveuses spécialisées de la cochlée.

Le conduit vocal peut être représenté par un tube acoustique à section $A(x)$ non uniforme, figure 4-3 :

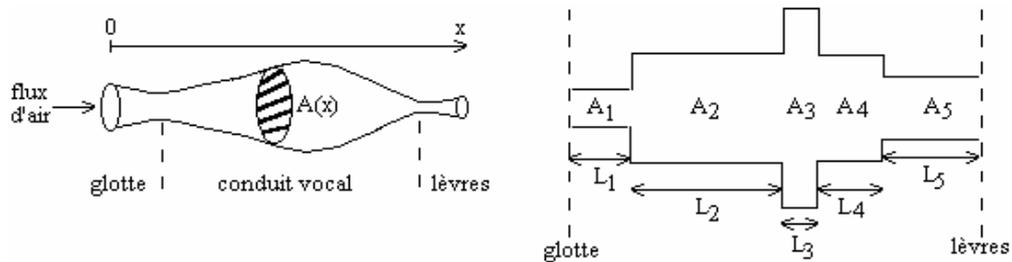


Figure 4-3: Schéma de l'appareil vocal et son approximation en tubes concaténés

Comme les différents sons sont formés par variation d'état du conduit vocal, les propriétés spectrales du signal de la parole varient dans le temps selon les variations de la forme du conduit vocal. Le spectre en fréquence est pondéré par la sélectivité fréquentielle du tube.

4.1.3.1. Théorie de la production de la parole :

Le modèle largement utilisé dans la production de la parole, est basé sur l'idée que le conduit vocal peut être représenté comme une concaténation de tubes acoustiques à pertes [57], figure 4-3. Les sections A_k des tubes, à aires constantes et de longueurs L_k , sont choisies de manière à approximer la fonction d'aire $A(x)$ du conduit vocal. Considérons la jonction entre le $k^{\text{ème}}$ et $(k+1)^{\text{ème}}$ tubes, figure 4-4.

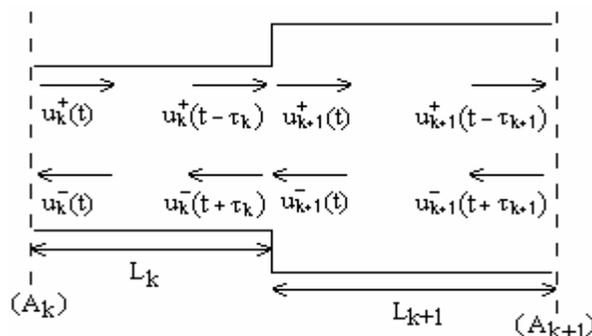


Figure 4-4: Jonction entre tubes avec pertes

Les équations de propagation des ondes acoustiques, au niveau de cette jonction, sont [74] :

$$u_{k+1}^+(t) = (1 + r_k) u_k^+(t - \tau_k) + r_k u_{k+1}^-(t)$$

4.1

$$u_k^-(t + \tau_k) = -r_k u_k^+(t - \tau_k) + (1 - r_k) u_{k+1}^-(t) \quad 4.2$$

où : $u_k^+(t)$ et $u_k^-(t)$ sont les ondes progressive et régressive se propageant dans le $k^{\text{ème}}$ tube.
 $\tau = L_k/v_S$ est le temps que met l'onde pour parcourir la longueur du $k^{\text{ème}}$ tube à la vitesse v_S du son. Tandis que r_k représente le coefficient de réflexion pour la $k^{\text{ème}}$ jonction:

$$r_k = \frac{A_{k+1} - A_k}{A_{k+1} + A_k} \quad 4.3$$

Comme les aires sont toutes positives,

$$\text{il vient :} \quad -1 \leq r_k \leq 1 \quad 4.4$$

4.1.3.2. Modèle de production de la parole :

Dans la figure 4-5, est représenté le conduit vocal de transmittance $V(z)$, excité par un générateur $u_G(t)$ d'impédance acoustique Z_G et chargé par l'impédance acoustique Z_L .

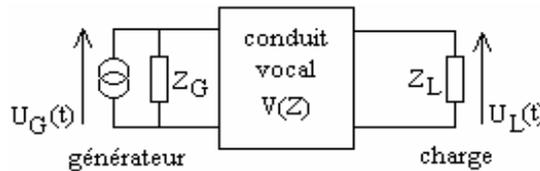


Figure 4-5: Modèle analogique de la production de la parole

$u_G(t)$ représente la vitesse du flux volumique d'air, dont la forme d'onde dans le cas d'une excitation voisée, est donnée en figure 4-6.

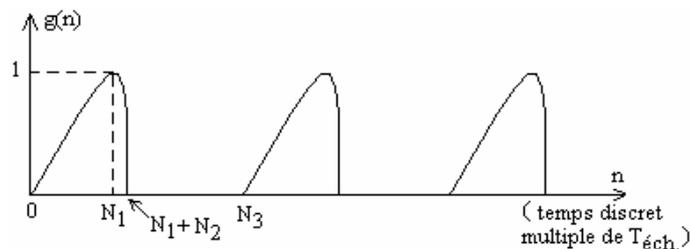


Figure 4-6: Forme des impulsions glottiques

Ce type d'onde peut être synthétisé par [74] :

$$\begin{cases} g(n) = \frac{1}{2} \left[1 - \cos\left(\frac{\pi \cdot n}{N_1}\right) \right] & 0 \leq n \leq N_1 \\ g(n) = \cos\left[\pi \left(\frac{n - N_1}{2 N_2} \right) \right] & N_1 \leq n \leq N_1 + N_2 \\ g(n) = 0 & N_1 + N_2 \leq n \leq N_3 \quad \text{avec} \quad N_3 T_{\text{éch.}} = T_{\text{pitch}} \end{cases} \quad 4.5$$

Le générateur modélise donc la vibration des cordes vocales dans la production de la parole voisée, par un train d'impulsions à la période du fondamental (ou pitch). Pour les sons non voisés, le générateur devient une source de bruit aléatoire.

L'impédance acoustique glottique Z_G est obtenue par linéarisation des relations entre pression et vitesse volumique d'air dans la glotte.

La charge Z_L , quant à elle, représente l'impédance du rayonnement des lèvres déterminée, en régime permanent, par le rapport entre pression et vitesse volumique au niveau des lèvres.

En général, les propriétés de l'excitation et du conduit vocal demeurent fixes durant des périodes de 10 à 30ms [14, 74]. La fonction de transfert du conduit vocal peut donc être celle d'un système linéaire. Le conduit vocal étant caractérisé par un ensemble d'aires de section, ou de façon équivalente par des coefficients de réflexion, sa fonction de transfert doit les faire intervenir. Son expression pratique, reliant la vitesse volumique d'air générée à la source à celle aboutissant aux lèvres, peut se mettre sous la forme simple d'une fonction tous-pôles [74] :

$$V(z) = \frac{U_L(z)}{U_G(z)} = \frac{G}{D(z)} \quad 4.6$$

où G et $D(z)$ dépendent de la fonction d'aire. Selon [74], le polynôme $D(z)$ s'écrit :

$$D(z) = 1 + \sum_{k=1}^N \alpha_k z^{-k} \quad 4.7$$

Avec N , nombre de sections du modèle à tubes avec pertes. Son choix dépend de la fréquence d'échantillonnage T du signal de la parole. La réponse fréquentielle du modèle à tubes avec pertes étant périodique, le modèle ne peut approximer le conduit vocal que dans la bande de fréquences $F \leq 1/2T$.

Au maximum, la période d'échantillonnage vaut $T=2\tau$ où τ est le temps de propagation dans une seule direction d'un tube du modèle. Ainsi pour N sections (N tubes) sur une longueur totale l , on a $\tau = l/v_s N$.

N étant aussi l'ordre du polynôme $D(z)$ (4.7), on peut avoir ainsi au maximum $N/2$ pôles complexes conjugués correspondant à des résonances dans la bande F .

En pratique, pour une longueur du conduit vocal estimée à $l=17.5\text{cm}$ et $v_s=35000\text{ cm/s}$, on obtient: $F = \frac{1}{2T} = \frac{1}{4\tau} = \frac{N v_s}{4l} = \frac{N}{2} 1000\text{Hz}$. Ce qui donne $N/2$ résonances (formants) par 1000Hz

de fréquence pour une longueur de 17.5cm du conduit vocal. Par exemple, pour $1/T=10\text{ kHz}$; la bande de base sera de 5 kHz et N doit être de 10.

Le modèle discret de l'appareil phonatoire, pouvant décrire le signal de la parole à partir de ses caractéristiques les plus importantes, est donnée en figure 4-7. Les paramètres propres au conduit vocal décrivent des résonances qui correspondent aux formants en cours. Ces paramètres du modèle sont supposés demeurer constants sur des intervalles typiques de 10 à 30 ms. Ce principe de quasi-stationnarité du modèle nous permettra, par la suite, de généraliser les méthodes d'analyse appliquées aux signaux aléatoires stationnaires.

Le changement du modèle d'excitation se fait par commutation entre générateurs d'excitation voisée et non voisée. Le contrôle du gain par A_V et A_N , permet de contrôler l'intensité de chaque type d'excitation [14, 74].

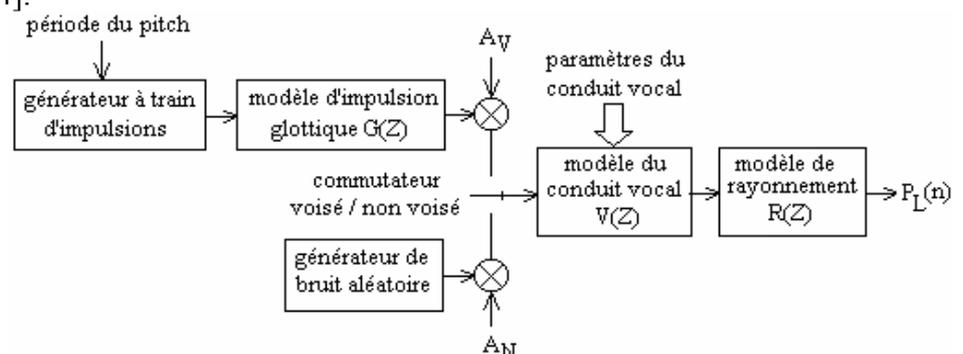


Figure 4-7: Modèle discret de la production de la parole

La limitation du modèle se ramène au manque de zéros (anti-résonance) de la fonction de transfert $V(z)$, comme il est requis théoriquement pour un son nasal.

Dans certains cas, il convient de combiner ensemble les modèles d'impulsion glottique et de rayonnement avec les composantes du conduit vocal, et de les représenter en une seule fonction de transfert : $H(z)=G(z)V(z)R(z)$, du type tous-pôles. Ce modèle sera adopté dans la prochaine analyse à prédiction linéaire.

4.2. Méthodes d'analyse de la parole

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

Le traitement est aussi utilisé pour réduire la redondance du signal vocal et permet ainsi de comprimer l'onde avant l'enregistrement, la transmission, ou l'extraction des paramètres pertinents pour la reconnaissance. Dans cette dernière perspective, les paramètres les plus couramment extraits sont issus d'une modélisation paramétrique autorégressive sous codage à prédiction linéaire [14]. Cette technique est retenue pour la stratégie d'implant cochléaire qu'on va développer.

Une autre méthode d'analyse est prévue comme outil de comparaison, afin de valider la méthode prédictive retenue. Elle se base sur les descriptions conjointes temps-fréquence, couramment utilisées en analyse de la parole. En particulier, la FFT à court terme générant spectres et sonagrammes en temps réel, est choisie dans notre cas pour confirmer les résultats obtenus par le modèle prédictif de la parole.

4.2.1. Analyse par prédiction linéaire

La méthode du codage à court terme par prédiction linéaire LPC [14, 74], est devenue une technique prédominante dans l'estimation des paramètres de base de la parole, tels que pitch et formants. En téléphonie cellulaire, elle constitue un standard de transmission à faible débit de la parole. L'importance de cette méthode est liée à la fois, aux estimations précises des paramètres de la parole qu'elle procure et sa relative rapidité en calcul.

L'idée de base de l'analyse à prédiction linéaire est qu'un échantillon de la parole peut être approximé par une combinaison linéaire d'échantillons passés de la parole. Par minimisation de la somme des carrés de différence (principe des moindres carrés) entre les échantillons présents de la parole et ceux prédits linéairement, un ensemble unique de coefficients de prédiction peut être déterminé, identifiant ainsi le signal de la parole.

4.2.1.1. Estimation du modèle autorégressif :

La modélisation ou identification d'un signal $x(n)$ consiste à lui associer un filtre linéaire qui, soumis à une excitation particulière reproduit ce signal le plus fidèlement possible.

La modélisation autorégressive (AR) d'un signal aléatoire est basée sur un filtre de transmittance de type tous-pôles, excité par un bruit blanc $u(n)$ de moyenne nulle et de variance unitaire [74].

La forme particulière du modèle discret de la parole évoquée en figure 4-7, est appropriée à l'utilisation de l'analyse LPC. Dans ce cas, l'effet du rayonnement aux lèvres $R(z)$, la forme de l'impulsion glottale $G(z)$, et la transmittance du conduit vocal $V(z)$, sont représentés ensemble par un filtre digital variable dans le temps, sous forme de la fonction $H(z)$. C'est à dire rapport de sortie du filtre $X(z)$ à son entrée $U(z)$, suivant les formes d'expressions (4.6) et (4.7) :

$$H(z) = \frac{X(z)}{U(z)} = \frac{\alpha}{A_p(z)} = \frac{\alpha}{1 + \sum_{k=1}^p a(k) z^{-k}} \quad 4.8$$

Ce système est excité par un train d'impulsions pour la parole voisée (V) et une séquence de bruit aléatoire pour la parole non voisée (NV). Ainsi les paramètres lentement variables dans le temps de ce modèle sont : la classification voisée /non voisée, la période de pitch du son voisé F_0 , le paramètre gain α , et les coefficients $a(k)$ du filtre digital, figure 4-8.

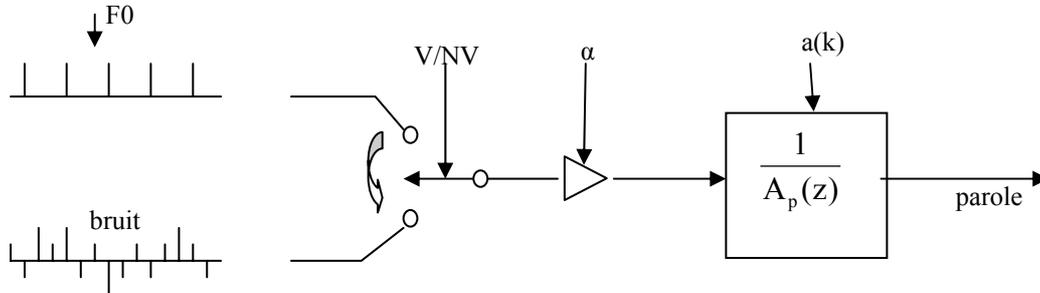


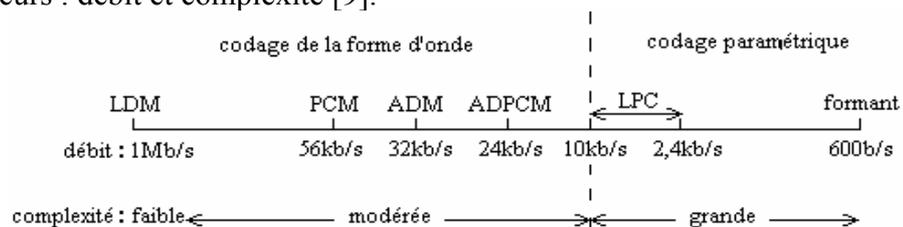
Figure 4-8: Modèle AR de la parole

L'avantage majeur de ce modèle est que ses paramètres peuvent être estimés à l'aide d'algorithmes rapides et efficaces utilisant la méthode LPC.

Le polynôme $A_p(z)$ représente la transmittance d'un filtre FIR appelé filtre inverse, qui, soumis aux échantillons de la parole $x(n)$, reproduit l'excitation $e(n) = \alpha u(n)$.

Le problème de l'estimation d'un modèle LPC, revient à déterminer les coefficients du filtre connaissant le signal de sortie, mais pas l'entrée. Il est par conséquent nécessaire d'adopter un critère, afin de faire un choix parmi une infinité de solutions possibles. Le critère classiquement utilisé est celui de la minimisation de l'énergie de l'erreur de prédiction. Cependant, la forme simplifiée en tous-pôles de la transmittance du filtre dite modélisation autorégressive, est restreinte à la représentation naturelle des sons voisés non nasals. Pour les sons fricatifs et nasals, la théorie acoustique fait appel à des pôles et des zéros dans la fonction de transfert du conduit vocal, afin de mieux représenter les phonèmes d'anti-résonance ou anti-formants. En revanche, on montre [14, 74] que pour un ordre p suffisamment élevé, le modèle tous-pôles assure tous de même une bonne représentation de tous les sons de la parole.

Le modèle AR est utilisé dans les vocodeurs LPC ou les vocodeurs à formants. Le codage LPC à 2400 bps a été normalisé par l'Otan [9]. Dans les vocodeurs à formants [14], le codeur détermine en plus, la position et la largeur de bande des 3 premiers formants. Au décodeur, la parole synthétique est obtenue par trois filtres accordés sur les formants. Le signal résultant est mis à l'échelle en fonction de l'énergie de la trame. On obtient avec cette technique un signal intelligible pour des débits jusqu'à 600 bps, mais la détermination des formants est une tâche un peu difficile. La figure 4-9 montre une comparaison des représentations digitales de la parole en fonction de deux facteurs : débit et complexité [9].



Où LDM : linear delta modulation
 PCM : pulse coding modulation
 ADM : adaptive delta modulation

Figure 4-9: Comparaison des méthodes de codage de la parole

De (4.8), les échantillons de la parole $x(n)$ sont liés à l'excitation $u(n)$ par une équation de différence :

$$x(n) = -\sum_{k=1}^p a(k) x(n-k) + \alpha u(n) \quad 4.9$$

Ainsi, chaque échantillon du signal original $x(n)$ est approché par une combinaison linéaire des p échantillons qui le précèdent. Cette prédiction n'est possible que si $x(n)$ est autocorrélé. Les coefficients $a(k)$ sont appelés coefficients de prédiction d'ordre p , et le signal $\alpha u(n) = e(n)$ qui est l'excitation, est aussi interprété comme erreur de prédiction (ou résidu) d'ordre p qu'il faut minimiser (excitation comme bruit blanc de moyenne nulle).

Soit :

$$e(n) = \sum_{k=0}^p a(k) x(n-k) \quad \text{avec } a(0) = 1 \quad 4.10$$

Nous supposons d'abord que le signal $x(n)$ est aléatoire et stationnaire. Les coefficients $a(k)$ sont donc indépendants du temps et leur estimation est basée sur la minimisation de la variance de l'erreur de prédiction ou énergie résiduelle de prédiction [14, 74].

Soit :

$$\sigma_e^2 = \sum_{n=-\infty}^{+\infty} e^2(n) = \sum_n \left[x(n) + \sum_{k=1}^p a(k) x(n-k) \right]^2 \quad 4.11$$

Trouver les $a(i)$ qui minimisent σ_e^2 , revient à annuler les dérivées partielles de σ_e^2 par rapport à ces coefficients eux même.

$$\frac{d\sigma_e^2}{da(j)} = 2 \sum_{n=-\infty}^{+\infty} \left[x(n) + \sum_{k=1}^p a(k) x(n-k) \right] x(n-j) \quad \text{pour } j=1 \dots P \quad 4.12$$

$$= 2 \left\{ \sum_{k=1}^p \left[a(k) \sum_n x(n-k) x(n-j) \right] + \sum_n x(n) x(n-j) \right\} \quad 4.13$$

$$= 2 \left\{ \sum_{k=1}^p a(k) \sum_m x(m) x(m+k-j) + \sum_n x(n) x(n-j) \right\} \quad \text{avec } m=n-k \quad 4.14$$

$$\frac{d\sigma_e^2}{da(j)} = 0 \Rightarrow \sum_{k=1}^p \left[a(k) \sum_m x(m) x(m+k-j) \right] = -\sum_n x(n) x(n-j) \quad \text{pour } j=1 \dots P \quad 4.15$$

Ce sont les équations de Yule-Walker.

On sait que pour un signal stationnaire, la fonction d'autocorrélation vérifie :

$$\phi_x(i) = \phi_x(-i) = \sum_l x(l) x(l+i) \quad \text{et } \phi_x(0) = \sigma_x^2 \quad 4.16$$

Les équations de Yule-Walker deviennent alors :

$$\sum_{k=1}^p a(k) \phi_x(k-j) = -\phi_x(j) \quad \text{pour } j=1 \dots p \quad 4.17$$

En posant :

$$\phi^{(p)} = [\phi_x(1), \phi_x(2), \dots, \phi_x(p)]^T \quad 4.18$$

$$\text{Et : } \Phi^{(p)} = \begin{bmatrix} \phi_x(0) & \phi_x(1) & \cdots & \phi_x(p) \\ \phi_x(1) & \phi_x(0) & \cdots & \phi_x(p-1) \\ \vdots & \vdots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ \phi_x(p) & \phi_x(p-1) & \cdots & \phi_x(0) \end{bmatrix} = \begin{bmatrix} \sigma_x^2 & \phi^{(p)\Gamma} \\ \phi^{(p)} & \Phi^{(p-1)} \end{bmatrix} \quad 4.19$$

Les équations de Yule-Walker (4.17) en notation matricielle, deviennent :

$$\begin{bmatrix} \phi_x(0) & \phi_x(1) & \bullet & \bullet & \bullet & \phi_x(p-1) \\ \phi_x(1) & \phi_x(0) & & & & \phi_x(p-2) \\ \bullet & & & & & \bullet \\ \bullet & & \bullet & & & \bullet \\ \bullet & & & \bullet & & \bullet \\ \phi_x(p-2) & & & & \bullet & \phi_x(1) \\ \phi_x(p-1) & \bullet & \bullet & \bullet & \bullet & \phi_x(0) \end{bmatrix} \begin{bmatrix} a(1) \\ a(2) \\ \bullet \\ \bullet \\ \bullet \\ a(p-1) \\ a(p) \end{bmatrix} = - \begin{bmatrix} \phi_x(1) \\ \phi_x(2) \\ \bullet \\ \bullet \\ \bullet \\ \phi_x(p-1) \\ \phi_x(p) \end{bmatrix} \quad 4.20$$

$$\text{Soit : } \Phi^{(p-1)} A = -\phi^{(p)} \quad 4.21$$

La matrice $\Phi^{(p-1)}$ est dite de Toplitz, puisqu'elle vérifie :

- la symétrie.
- Elle est définie positive.
- Les éléments situés sur une diagonale parallèle à la diagonale principale sont identiques.

Plusieurs algorithmes ont été proposés pour la résolution de ce système, afin de tirer le vecteur inconnu A relatif aux coefficients de prédiction $a(k)$. Le plus utilisé est celui de Levinson-Durbin. Mais nous avons opté pour l'algorithme de Schur qui est utilisé dans la norme GSM.

En revenant à l'expression (4.11), celle-ci s'écrit encore :

$$\sigma_e^2 = \sum_n \left[\sum_{j=0}^p a(j)x(n-j) \cdot \sum_{k=0}^p a(k)x(n-k) \right] \quad 4.22$$

$$= \sum_{j=0}^p a(j) \sum_{k=0}^p a(k) \cdot \sum_n x(n-j)x(n-k) \quad 4.23$$

Pour $n-j = m$, et $n-k = m+j-k$, on obtient :

$$\sigma_e^2 = \sum_{j=0}^p a(j) \sum_{k=0}^p a(k) \phi_x(j-k) \quad 4.24$$

En remplaçant (4.17) dans (4.24), on obtient la valeur minimisée de la variance de l'erreur de

$$\text{prédiction : } \sigma_{e_{\min}}^2 = \alpha_p = \sum_{j=0}^p a(j) \phi_x(j) \quad 4.25$$

Le signal de la parole étant fortement non-stationnaire, ce type de modélisation ne reste guère valable plus de 30ms. On notera donc que l'analyse LPC d'un signal de la parole, implique la résolution d'un système de 10 équations à 10 inconnues toutes les 10 ms (à 30ms au maximum).

D'autre part, si l'on réunit (4.21) et (4.25), on obtient :

$$\Phi^{(p)} \cdot \begin{bmatrix} 1 \\ A \end{bmatrix} = \begin{bmatrix} \alpha_p \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad 4.26$$

Comme $e(n) = \alpha u(n)$, et pour une excitation $u(n)$ du type bruit blanc de moyenne nulle et de variance unitaire, le gain de modèle devient :

$$\alpha = \sigma_{e \min} = \sqrt{\alpha_p} \quad 4.27$$

4.2.1.2. Méthode de résolution et stabilité du modèle :

La résolution du système (4.26) pour l'obtention des paramètres LPC (linear predictive coding) $a(k)$ du vecteur A , avec $k=1$ à P , peut s'effectuer par plusieurs méthodes [14, 74]. Les algorithmes sont de nature récursive sur l'ordre de la prédiction; où l'on construit les solutions pour des ordres successifs $m=1, 2, \dots$, jusqu'à P .

La méthode retenue dans notre cas est celle de Schur [14], qui présente moins de calcul avec une approche d'une relative simplicité. Les coefficients de prédiction $a_m(i)$, avec $(m, i) = 1, 2, \dots, P$, y sont complètement caractérisés par des coefficients de corrélation partielle ou Parcor :

$$k_m = a_m(m) = \sqrt{1 - \frac{\alpha_m}{\alpha_{m-1}}} \quad 4.28$$

Où $\alpha_m > 0$, est l'énergie résiduelle d'ordre m selon l'expression (4.25). Comme l'énergie résiduelle diminue avec l'ordre, c-à-d $\alpha_m < \alpha_{m-1}$, il en résulte que $k_m < 1$.

Le modèle autorégressif étant stable lorsque toutes les racines de $A_p(z)$ sont situées à l'intérieur du cercle unitaire, le critère de stabilité s'exprime tout simplement en fonction des coefficients parcor k_m : stabilité du modèle AR si et seulement si $|k_m| \leq 1$ pour $m=1, 2, \dots, P$.

Le modèle acoustique discuté au paragraphe 4.1.3, et le modèle LPC sont identiques lorsqu'on prend l'ordre P du filtre inverse égal au nombre N des tubes à pertes. Par comparaison [14], on a :

$$\begin{cases} D(z) = A(z) \\ r_m = k_m \end{cases} \quad 4.29$$

Les coefficients parcor représentent donc les coefficients de réflexion aux sections adjacentes du conduit vocal. Selon (4.4), on retrouve bien $|k_m| < 1$ pour tous m .

Ainsi l'algorithme de Schur, basé sur le calcul des coefficients parcor, fournit un modèle autorégressif stable. La forme simple de l'expression du critère de stabilité qui en découle, est d'un grand intérêt pour l'estimation du modèle autorégressif en temps réel.

4.2.2. Analyse conjointe temps-fréquence

Parmi les principales analyses, on dénombre [70, 77]:

4.2.2.1. Représentation par Ondelettes [67, 101] :

Une famille d'ondelettes peut être déduite par dilatation-translation d'une fonction carrée intégrable $\varphi(t)$:

$$\varphi_{a,b}(t) = \frac{1}{\sqrt{a}} \varphi\left(\frac{t-b}{a}\right) \quad 4.30$$

Ces fonctions peuvent être discrétisées si on prend par exemple : $a = 2^j$ et $b = n a$.

d'où :

$$\varphi_{j,n}(t) = 2^{-j/2} \varphi(2^{-j} t - n) \quad 4.31$$

$\varphi_{a,b}(t)$ et sa forme discrète $\varphi_{j,n}(t)$, ont même aspect que la fonction mère $\varphi(t)$ mais sont d'étendue temporelle proportionnelle au facteur d'échelle a (ou j), alors que l'étendue fréquentielle de leur spectre en est inversement proportionnelle.

La projection d'une fonction $f(t)$ sur la base $\varphi_{j,n}(t)$, c-à-d décomposition de f par combinaison linéaire des fonctions $\varphi_{j,n}(t)$, constitue la transformée en ondelette discrète, qui s'écrit :

$$W_{\text{ondelette}}[f(t)] = \int_{-\infty}^{\infty} f(t) \varphi_{j,n}^* dt \quad 4.32$$

Une représentation conjointe temps-fréquence, est appelée scalogramme.

Nous avons réalisé un logiciel délivrant le scalogramme de la parole analysée. Le scalogramme décrit les coefficients de projection, du signal vocal sur la famille d'ondelettes Spline d'ordre 1 [101]. Il s'agit donc d'une décomposition du signal en bandes fréquentielles, notées d_j sur l'axe vertical, en fonction du temps. Dans le cadre de la table 4-2, d_1 correspond à la bande fréquentielle la plus élevée et aussi la plus longue (donc moins de résolution en fréquence et plus de résolution en temps), tandis que d_9 correspond à la bande fréquentielle la plus basse et aussi la plus étroite (donc plus de résolution en fréquence et moins de résolution en temps).

Deux procédures ont été développées :

1. Procédure Filtrés :

Elle calcule, pour la base orthonormée d'ondelette spline linéaire, les deux filtres en quadrature qui lui sont associés, ainsi que leurs formes retournées [101]. Ces deux paires de filtres sont à réponse impulsionnelle RII, donc de longueur infinie. Mais pratiquement le nombre de coefficients à retenir s'est avéré limité à 8, à cause de leurs valeurs devenant négligeables par la suite.

2. Procédure Mallat :

Elle calcule la transformée en ondelette du signal vocal analysé en utilisant l'algorithme de Mallat [101]. Le calcul de cette transformée se ramène à filtrer les coefficients d'approximation $a^{j-1}(n)$ de l'échelle précédente $j-1$, pour obtenir ceux d'approximation $a^j(n)$ et de détail $d^j(n)$ de l'échelle suivante j . On rappelle que j est l'indice d'échelle désignant le numéro de bande fréquentielle analysée, alors que n est l'indice de translation temporelle qui suit l'évolution temporelle du signal d'entrée.

L'opération de filtrage des coefficients s'effectue par leur produit de convolution avec les deux filtres en quadrature retournés, tout en respectant la règle du sous-échantillonnage par deux.

On rappelle que la suite $a^0(n)$ n'est autre que les échantillons du signal d'entrée, qui sont traités par paquet de 512 échantillons.

Dans un scalogramme, c'est le module au carré des coefficients de détail $|d^j(n)|^2$, à différentes échelles j et en fonction du temps n , qui est généralement représenté. Mais on peut se contenter d'une représentation à valeurs absolues des coefficients réels, comme utilisé dans notre cas.

la répartition graphique des coefficients de projection en ondelettes, a été faite suivant des bandes fréquentielles en octave. Le tableau 4-2 indique cette répartition.

intervalle en fréquence f (Hz)	coefficient de projection représentés
2750-5500	$d^1(n)$ sur 128 pixels
1375-2750	$d^2(n)$ sur 64 pixels
688-1375	$d^3(n)$ sur 32 pixels
344-688	$d^4(n)$ sur 16 pixels
172-344	$d^5(n)$ sur 8 pixels
86-172	$d^6(n)$ sur 4 pixels
43-86	$d^7(n)$ sur 2 pixels

22-43	$d^8(n)$ sur 1 pixel
11-22	$d^9(n)$ sur 1 pixel
0-11	$a^9(n)$ sur 1 pixel

Table 4-2: Attribution de bandes fréquentielles aux coefficients d'ondelette

Malheureusement, les scalogrammes n'ont pas fourni un bon outil d'analyse de la parole. Ils ne sont pas bien adaptés à la représentation précise des trajets formantiques de la parole. Un scalogramme présente des intervalles fréquentiels qu'on fixe à priori. On peut les faire coïncider avec ceux des formants, qui sont d'ailleurs imprécis même si certains chercheurs fixent des intervalles pour les différents formants [74], exemple:

$F_0 = [30-100 \text{ Hz}]$, $F_1 = [0.2-1.3 \text{ kHz}]$, $F_2 = [0.55-3 \text{ kHz}]$, $F_{3,4} = [1.2-5 \text{ kHz}]$, $F_{\text{Nasal}} = [248-528 \text{ Hz}]$.

Cependant, un trajet formantique peut traverser des intervalles contigus et le numéro de son formant se voit ainsi changé dans le scalogramme, alors qu'il s'agit toujours du même formant. C'est la raison qui nous a poussé à utiliser la transformée de Fourier classique.

4.2.2.2. Représentation par Transformée de Fourier glissante :

La transformée de Fourier est un outil puissant permettant de déterminer le spectre en fréquences du signal temporel analysé. Son expression dans le cas d'un signal discret $x(k)$, est :

$$X(f) = \sum_{k=-\infty}^{+\infty} x(k) e^{-j2\pi f k} \quad 4.33$$

$X(f)$ est périodique à cause de la forme discrète du signal d'entrée. La fréquence de Nyquist ou fréquence maximale du spectre, pouvant être représentée à une fréquence d'échantillonnage F_e , est : $F_{\text{Nyq}} = F_e / 2$, ceci afin d'éviter le recouvrement spectral dû à la périodicité de $X(f)$.

Pour avoir une discrétisation en fréquence, on fait recours à l'échantillonnage dans le domaine fréquentiel par pas de $1/N$: on a donc $X(n) = X(n+N)$. Cette opération engendre une périodisation temporelle. Afin d'éviter le recouvrement introduit dans le domaine temporel, il faut limiter la durée de $X(n)$ à l'échantillon L , avec $L \leq N$ [77].

La transformée de Fourier discrète (TFD) permet le passage du signal discret $x(k)$ à sa représentation fréquentielle discrète $X\left(f = \frac{n F_{\text{Nyq}}}{N}\right) = X(n)$.

Ainsi, si N est le nombre d'échantillons de $x(k)$, sa TFD s'écrit :

$$X(n) = \sum_{k=0}^{N-1} x(k) e^{-j2\pi \frac{k}{N} n} \quad n = 0, 1, \dots, N-1 \quad 4.34$$

La quantité F_{nyq} / N représente la résolution en fréquence du spectre $X(n)$. La FFT est un algorithme efficace permettant d'accélérer le calcul de la TFD.

L'analyse spectrale de la parole présente des avantages au niveau de la perception, car l'oreille humaine effectue ce genre d'analyse. De plus, elle fait apparaître les propriétés auxquelles on attache une grande importance telles que les formants.

Ainsi d'après le principe de l'analyseur de spectre, la FFT à court terme (ou glissante, ou fenêtrée) est appliquée. Elle est obtenue en extrayant du signal temporel (ou audiogramme) une trentaine de millisecondes de signal vocal et en effectuant une transformée de Fourier sur ces échantillons. Le résultat de cette transformation mathématique est souvent présenté dans un graphique qui donne, en fonction de la fréquence, l'amplitude des composantes présentes dans le signal analysé. Il est souvent intéressant de représenter l'évolution temporelle du spectre à court terme d'un signal, sous la forme d'un *spectrogramme*. Le spectrogramme est la représentation temps-fréquence la plus courante. Les termes couramment employés de *Sonagraph* et *Sonagram*

sont des marques déposées [14]. Il est usuellement représenté avec le temps en abscisse, la fréquence en ordonnée, alors que les énergies sont notées en niveaux de gris ou en différence de couleur. Ce diagramme met en évidence l'enveloppe spectrale du signal, et permet par conséquent de visualiser ses trajectoires formantiques.

L'analyse temps-fréquence par FFT, utilise une fenêtre de pondération glissante. Le pitch étant responsable de nombreux lobes qui perturbent la lecture du spectrogramme, l'utilisation de fenêtre de pondération permet de s'en affranchir en assurant un lissage. Pour ne pas perdre d'information et assurer un meilleur suivi des non stationnarités, les fenêtres doivent se recouvrir. Elles ont généralement une longueur de 128 ou 512 points et le recouvrement est au minimum de 50 %.

Mathématiquement, la transformée de Fourier à court terme utilise une fenêtre de pondération g , symétrique et d'énergie finie :

$$X(n, b) = \text{TF}[x(k) g(k - b)] \quad 4.35$$

Le produit $x(k).g(k-b)$ représente la tranche du signal découpée par la fenêtre g à l'instant b . La famille de fenêtres d'analyse est obtenue par translation de g (variation de b).

Noté $S(n, b)$, le sonagramme est exprimé par [57] :

$$S(n, b) = |X(n, b)|^2 \quad 4.36$$

L'utilisation de fenêtres d'analyse conduit au compromis habituel [48, 57] entre la résolution spectrale, la résolution temporelle, l'amplitude des lobes secondaire, etc...

Le choix de la longueur de la fenêtre d'observation du sonagramme, doit être à la fois :

- Suffisamment grande pour une bonne résolution fréquentielle.
- Suffisamment courte pour une bonne résolution temporelle.

Quelques dispositions peuvent être prises pour améliorer la qualité des sonagrammes :

- Augmenter l'énergie des hautes fréquences afin de compenser le niveau plus faible des aigus, généralement par un filtre passe-haut dit de préaccentuation. Cela imite en fait les résonances du pavillon de l'oreille, voir paragraphe 1-1 du chapitre 1.
- Utiliser une échelle linéaire pour les basses fréquences (<1800Hz) et logarithmique pour les hautes fréquences. L'une de ces échelles mixtes est appelée échelle de Bark [14].

Remarques [14] : Le sonagramme permet de visualiser essentiellement les trajectoires et l'énergie des formants liés au signal de la parole. Les formants apparaissent sous la forme de bandes sensiblement parallèles à l'axe temporel. Chez un locuteur masculin, tous les formants sont inférieurs à 5000 Hz. Si les formants F1 et F2, voire F3, sont bien marqués, les formants F4 et F5 sont plus difficiles à retrouver dans le spectre. Ils peuvent d'ailleurs être regroupés. Le passage d'une consonne à une voyelle, ainsi que l'étape inverse, produit des transitions formantiques, c'est à dire des variations dans la position des formants. Le spectre d'un signal de type voisé possède en général plus de composantes en basse fréquence qu'en haute fréquence. Par contre, le spectre d'un signal non voisé présente une amplitude plus importante en haute fréquence (ce qui correspond à la perception que nous avons : les fricatives non-voisées sont des sons plus « aigus » que les voyelles).

4.3. Outil de comparaison par sonagramme

Deux types de traitement ont été apportés à la parole en vue d'extraire ses paramètres pertinents, notamment les formants. Un traitement réfère au codage LPC appliqué à notre implant, et l'autre utilise la FFT glissante pour visualiser les sonagrammes dans le but de valider le spectre obtenu par codage LPC. Le logiciel développé est de la programmation-système [3, 99] dans laquelle nous accédons directement à l'électronique de la carte son du PC.

4.3.1. Environnement requis du PC

L'outil générant les sonagrammes de la parole a été développé sur un ancien PC. Des restrictions existent donc pour son utilisation.

Les limitations du point de vue Hard, sont :

- La carte son doit être du type SoundBlaster ou équivalent avec un processeur audio-numérique DSP version 3.
- Le PC doit être muni d'un processeur 4-86 ou 3-86 avec co-processeur mathématique.
- L'horloge-système doit être supérieure à 100 MHz pour une restitution en temps réel de la parole avec ses graphes d'analyse.

Les limitations du point de vue Soft, sont :

- Le système d'exploitation MSDOS doit être complet, autorisant la programmation-système et l'accès à l'électronique de la carte son installée par le biais d'interruptions BIOS et DOS.
- MSDOS doit être autonome et non sous contrôle de Windows.
- Les fichiers audio, conçus ou traités, renferment des échantillons de la parole au format 8 bits, monophonique, à la cadence normalisée de 11 kHz.

Même dans les anciens PC, un enregistrement audionumérique se fait le plus souvent par une modulation par impulsions codées PCM à la fréquence de 44100 échantillons par seconde, où chaque échantillon est évalué sur 16 bits.

Dans notre cas, nous avons limité la bande fréquentielle du signal vocal à son minimum, soit à environ 3 kHz. Ce choix nous permet d'utiliser une fréquence d'échantillonnage au moins égale à 6 kHz. Un compromis, entre une bonne qualité d'échantillonnage de la voix et une représentation en temps réel du signal restitué, nous a poussé à opter pour la fréquence d'échantillonnage normalisée de 11 kHz.

Toujours dans la perspective de la limitation du temps de traitement du signal, le format retenu des échantillons est le PCM à 8 bits.

La plupart des microphones modernes sont du type à condensateur (ou à électret) depuis ceux du téléphone jusqu'à ceux vendus pour se brancher sur les PC multimédias. Le microphone que nous avons utilisé dans cette application, est de ce type.

4.3.2. Logiciel de traitement et d'enregistrement / restitution

Le logiciel en turbo-Pascal d'enregistrement-restitution avec traitement de la parole, est subdivisé en quatre unités avec leur programme principal. Une Description succincte du logiciel est donnée ci-après.

4.3.2.1. L'unité PCoutil.Pas :

Cette unité réunit les procédures de programmation du matériel du PC impliqué dans le cheminement du signal de la carte son vers le disque dur et réciproquement.

On dénombre deux contrôleurs du PC sur lesquels s'articulent les procédures de l'unité, ce sont le PIC et le DMA.

Parmi les fonctions essentielles de l'unité, on compte : l'autorisation-inhibition d'interruptions, l'installation de gestionnaire d'interruptions, la libération-masquage de canal DMA, la libération-allocation mémoire pour DMA, l'installation du mode de transfert DMA, etc...

4.3.2.2. L'unité detectSB.Pas :

Permet de détecter les caractéristiques de la carte son SoundBlaster installée. Les caractéristiques recherchées se reportent à : la version du DSP, l'adresse de base du port DSP et celui de la table de mixage, le canal DMA, et le numéro d'interruption IRQ.

4.3.2.3. L'unité Blaster.Pas :

Similairement à l'unité PCoutil.pas, cette unité réunit les procédures de programmation des contrôleurs propres à la carte son, à savoir, le DSP et la table de mixage ou mixer. Hormis les procédures relatives à l'ouverture-fermeture de fichiers, et lecture-écriture dans des fichiers, On distingue aussi, les principales fonctions de programmation de la carte son, suivantes : la lecture-écriture des données du DSP, la lecture-écriture du registre mixer, reset du DSP et du mixer, la suppression du tampon (DSP-DMA), le réglage de la fréquence en entrée-sortie du DSP, le réglage du volume en entrée-sortie, l'activation-désactivation du haut-parleur, l'activation-désactivation du filtrage passe-bas en entrée-sortie, l'activation-désactivation du mode stéréo, la détermination du nombre d'échantillons à transférer, l'enregistrement-restitution des échantillons.

4.3.2.4. L'unité Analyse.Pas :

Appelée par la procédure « DSP_RecPlay » de l'unité Blaster, elle permet d'effectuer les analyses et traitements suivants du son enregistré :

- représentation temporelle du son (mode oscilloscope).
- représentation fréquentielle du son, à l'aide de la transformée de Fourier glissante (mode analyseur de spectre).
- représentation du son sous forme d'un sonagramme (mode spectrogramme).

Deux unités sont à distinguer lors de la compilation du programme : Analyse1.pas et Analyse2.pas. La première unité prépare des sonagrammes à bandes étroites, en traitant les échantillons des fichiers audio sur des paquets successifs de 512 octets, d'une durée de 46.5 ms chacun.

La deuxième unité fournit quant à elle, des sonagrammes à larges bandes, où les paquets d'échantillons traités sont cette fois-ci à 128 octets de durée 11.6 ms chacun. Cependant, les 128 points sont complétés, de part et d'autre, par des zéros jusqu'à arriver à 512 points. Ceci afin de garantir une même étendue verticale des sonagrammes obtenus.

Dans les deux unités d'analyse, les échantillons bénéficient d'un prétraitement par accentuation et fenêtrage de Haming.

4.3.2.5. Le programme principal Son.Pas :

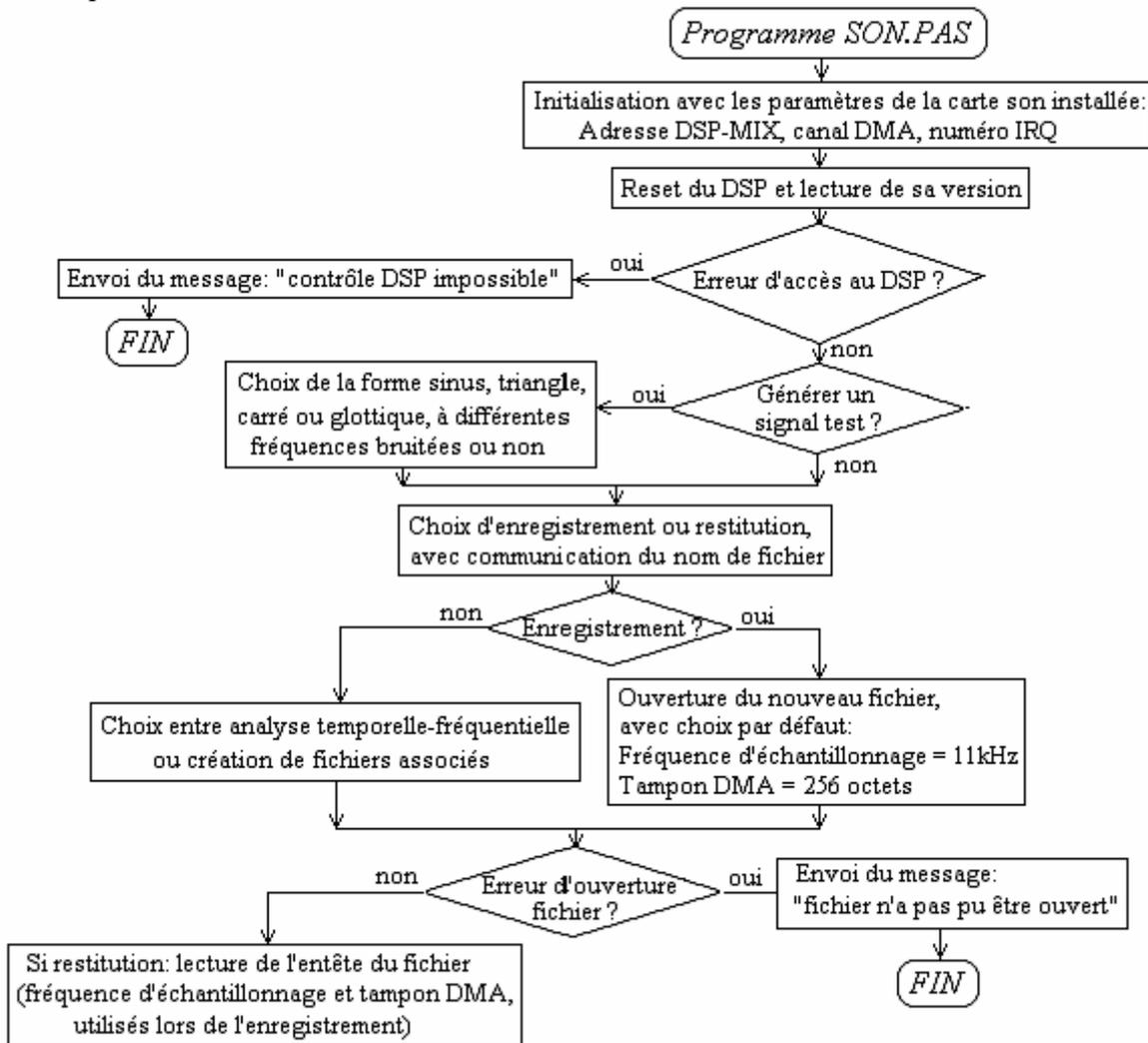
Il commence par détecter les caractéristiques de la carte SoundBlaster installée, et ce, grâce à l'unité detectSB.pas. Il offre par la suite à l'utilisateur, de visualiser certains signaux de test avec leurs spectres, comme le signal glottique de la parole par exemple, figure 4-9.



Figure 4-9: Allure temporelle et spectre d'une impulsion glottale à 40Hz

Le programme, représenté par l'organigramme de la figure 4-10, s'articule essentiellement sur la procédure « DSP-RecPlay » ou procédure d'enregistrement / restitution de la parole, appartenant à l'unité Blaster.pas.

Lors d'un enregistrement d'un fichier audio, sa taille est calculée à partir de la durée choisie. Le programme permet aussi de choisir lors de la restitution du fichier audio, entre ses différentes représentations : temporelle, fréquentielle ou sonagramme. Si le fichier audio est dépourvu de fichiers associés, il peut alors être traité et doté de fichiers d'analyse, temporelle et spectrale, qui sont affichés simultanément lors de sa restitution. La fréquence d'échantillonnage, du fichier restitué, peut être modifiée.



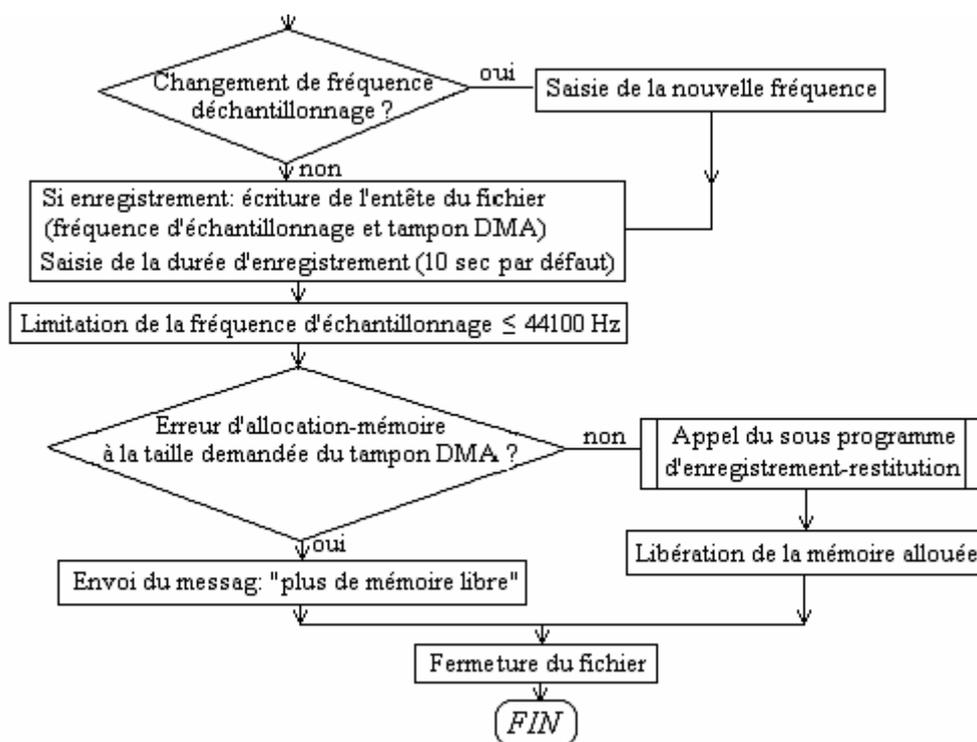


Figure 4-10: Organigramme principal Son.pas

4.3.2.6. La procédure d'enregistrement-restitution DSP_RecPlay :

Il s'agit de la procédure ultime, sur laquelle se base tout le logiciel. Figurant dans l'unité Blaster.pas, elle est utilisée par le programme principal. Cette procédure utilise à son tour et au choix, les procédures de l'unité d'analyse1.pas ou analyse2.pas, décrites dans l'organigramme de la figure 4-11.

Les étapes de programmation du couple de contrôleur DSP-DMA, sont respectées dans cette procédure. Il s'agit, avant tout enregistrement ou restitution d'un fichier audio, de programmer en premier, le canal DMA sur lequel est configuré la carte son avec :

- adresse de début du buffer (ou tampon) de transfert.
- longueur de transfert ou du bloc.
- mode de transfert choisi, ici « autoinitialize ».

Puis on passe à la programmation du DSP, en:

- fréquence d'échantillonnage choisie.
- taille des échantillons.
- exécution d'enregistrement-restitution à l'aide d'une commande appropriée.

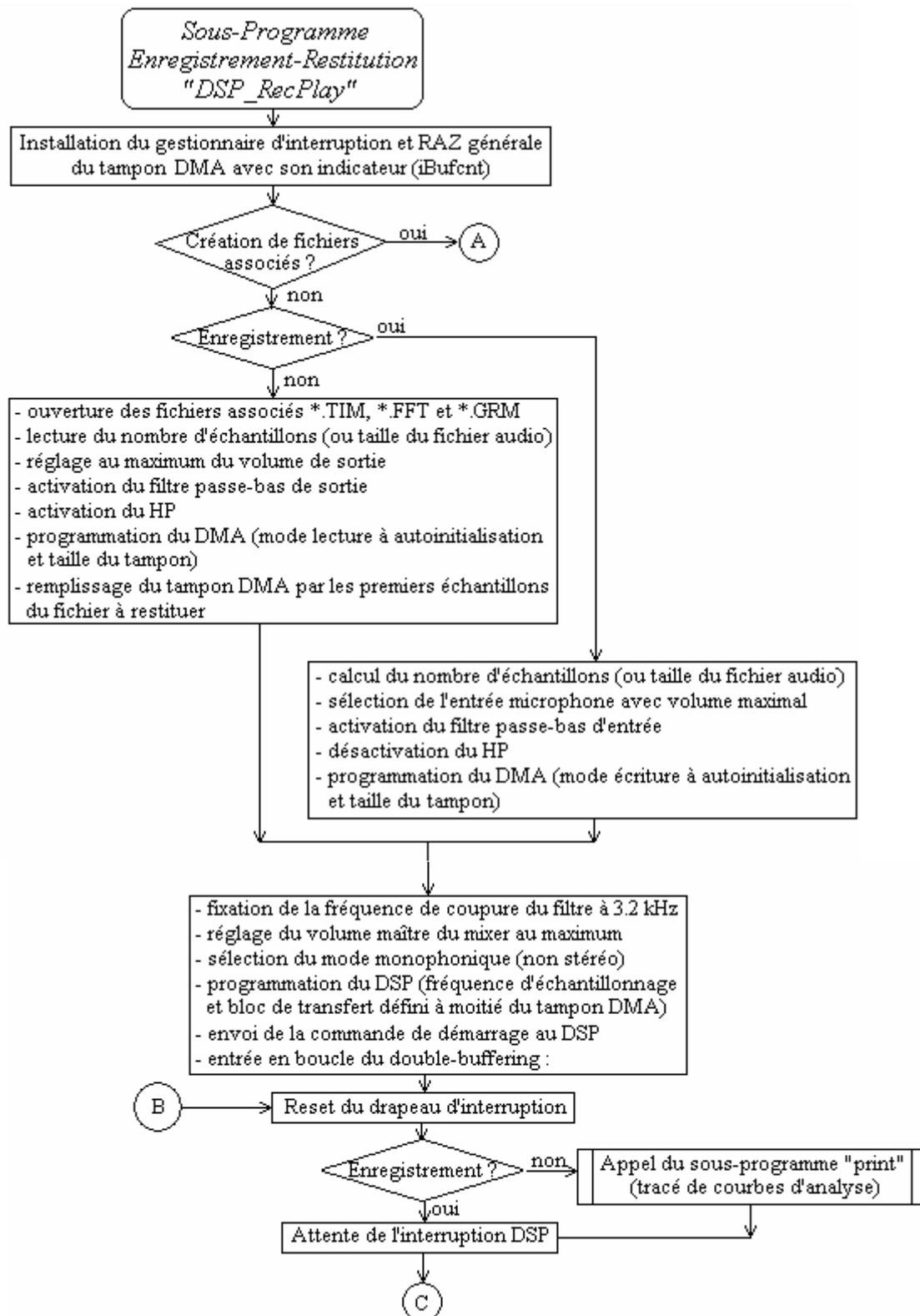
Le DMA, étant en mode Auto-Init, recycle indéfiniment entraînant de même le DSP qui lui est associé. Dans ce mode, le travail du couple DSP-DMA est donc déclenché une seule fois lors du lancement du DSP. Le procédé double-buffering (ou ping pong) est ensuite entamé à l'intérieur d'une boucle. Pour s'affranchir de la contrainte d'un buffer DMA limité en taille et de lire ou écrire n'importe quelle taille de fichier d'échantillons, on programme le DSP avec la moitié de la longueur et non pas la longueur tout entière. Ainsi, en double-buffering, le DSP est programmé sur la moitié du tampon DMA, et le couple DSP-DMA balaye ensemble la première moitié du tampon.

Une fois à sa fin, le DSP envoie une interruption au programme pour rafraîchir cette partie du tampon, tout en continuant le balayage de l'autre moitié. Ceci est dû à ce que le DSP suit automatiquement le bloc adressé par le DMA (adresse actuelle et longueur restante du bloc).

Une routine d'interruption accomplit sa tâche lorsqu'elle informe le programme que le DSP vient de terminer le traitement ou remplissage de la moitié du buffer, et qu'il faut lui préparer l'autre moitié. Si le DSP a lu, depuis le microphone des données, la routine inscrit le demi-buffer ainsi rempli dans le fichier d'échantillons sur disque dur. Si le DSP à jouer des données sur le haut-parleur, elle charge le demi-buffer avec le bloc de données suivant, provenant du fichier d'échantillons.

Quand le DSP a fini de traiter le deuxième demi-buffer, le contrôleur DMA revient au début du buffer grâce au mode *auto initialize* pour recommencer une nouvelle fois le transfert du premier demi-buffer. On a ainsi la garantie que le demi-buffer qui est chargé ou enregistré est toujours celui que le DSP n'est pas en train de traiter.

Il reste à préciser que le temps de traitement des échantillons (calcul FFT par exemple) est relativement long par rapport au temps de restitution du signal vocal en écoute. C'est pour cela que les calculs sont effectués juste après la phase d'enregistrement du fichier audio, et sont sauvegardés dans des fichiers associés utilisables lors de la phase de restitution. Ainsi, l'écoute du fichier audio et l'affichage de ses représentations spectrales et temporelles, se dérouleront au même rythme.



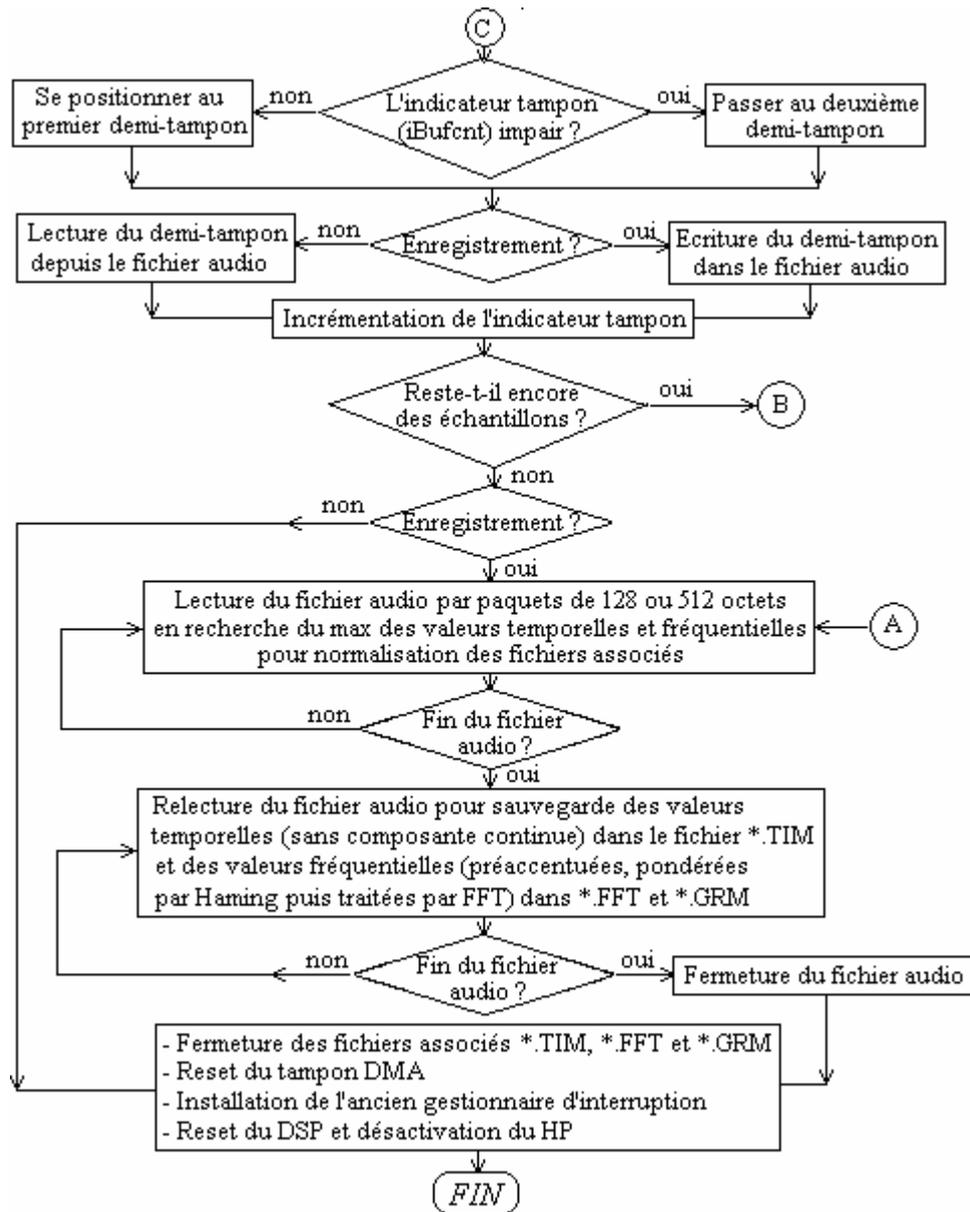


Figure 4-11: Organigramme DSP_RecPlay d'enregistrement / restitution

4.3.2.7. Procédures d'analyse du fichier audio :

Les procédures et fonctions suivantes, sont disponibles dans les deux unités complémentaires d'analyse : Analyse1.pas et Analyse2.pas.

* Procédure MaxValue :

Cette procédure recherche l'amplitude absolue maximale ou le module maximal des valeurs échantillonnées issues respectivement, du fichier audio ou de la procédure FFT. Toutes les autres valeurs, du fichier audio ou calculées par la procédure FFT, seront normalisées par rapport à ces deux valeurs maximales.

* Procédure SaveFiles :

Cette procédure calcule les valeurs normalisées précédentes et les ajuste pour un affichage ultérieur. Elles sont enregistrées dans des fichiers associés au fichier audio traité. Ce sont, le fichier d'extension « .TIM » pour les valeurs instantanées du signal vocal, le fichier

« .FFT » pour les valeurs fournies par la procédure FFT, et le fichier « .GRM » pour celles destinées au sonagramme.

** Procédure Print :*

Cette procédure permet de tracer deux graphiques et un diagramme propres aux sons enregistrés. Un graphique est relatif aux valeurs instantanées du signal vocal ou mode oscilloscope, donc issues du fichier associé d'extension « .TIM ».

L'autre graphique, représente le domaine fréquentiel du signal vocal ou mode analyseur de spectre, c'est le tracé des données du fichier associé d'extension « .FFT ».

L'axe fréquentiel s'étend jusqu'à la fréquence de Nyquist, qui est égale à la moitié de la fréquence d'échantillonnage. Dans notre cas, c'est : $f_N = f_{\text{éch.}} / 2 = 11000 / 2 = 5.5 \text{ kHz}$.

Si on échantillonne le signal à une fréquence supérieure à celle de Nyquist, il s'ensuit une sévère distorsion suite au repliement du spectre. Par conséquent, on filtre tout signal acoustique au-dessous de la fréquence de Nyquist avant de le numériser.

Nous représentons ici en figure 4-12, les différents fichiers associés .TIM et .FFT obtenus pour l'exemple de la voyelle [a].

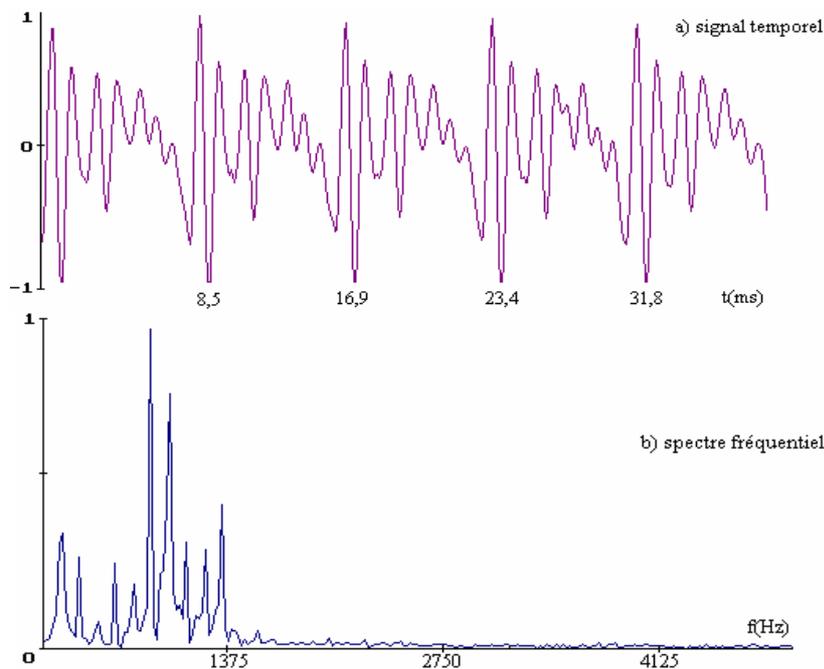


Figure 4-12: Chronogramme (a) et spectrogramme (b) de la voyelle [a] issus des fichiers TIM et FFT

En langage T. Pascal, nous avons une résolution graphique de 639 pixels sur l'axe horizontal d'un écran 14'' ou axe fréquentiel de notre application. Le nombre multiple de 2^n , le plus proche de 639, est $2^9=512$. Ce qui nous a conduit, d'ailleurs, à choisir cette valeur comme taille des échantillons manipulés par le DSP et d'allouer ainsi, un bloc de $2 \times 512 = 1$ koctets de mémoire-tampon pour le DMA dans le cas de l'unité Analyse1.pas .

Dans l'unité, Analyse2.pas, la taille allouée n'est que de $2 \times 128 = 256$ octets. La résolution spectrale en fréquence du signal audio, devient dans analyse1.pas:

$$\Delta f = \frac{f_{\text{éch.}}}{\text{taille_échantillons}} = \frac{11000}{512} = 21.48 \text{ Hz}$$

Alors qu'en analyse2.pas, on a : $\Delta f = \frac{11000}{128} = 85.94 \text{ Hz}$

L'indice de fréquence (numéro du pixel associé à la fréquence représentée) est donc :

$$i = \frac{\text{fréquence}}{\Delta f} = T \times \text{fréquence}$$

où $T = 46.5 \text{ ms}$ pour analyse1.pas et $T = 11.6 \text{ ms}$ pour analyse2.pas .

Ce temps représente la durée de la fenêtre d'observation ou durée du filtre de Haming utilisé. Ainsi, un temps d'observation long (46.5 ms) permet de mettre en évidence dans le sonagramme, les harmoniques de la parole sous formes de lignes horizontales équidistantes, on parle alors de sonagramme à bande étroite. Un temps d'observation réduit (11.6 ms), permet de se focaliser sur les formants de la parole, représentées sous forme de bandes verticales adjacentes, on parle alors de sonagramme à large bande.

Les deux types de sonagrammes peuvent être affichés séparément. Leur représentation, issue des valeurs traitées dans les fichiers associés d'extension « .GRM », est également limitée à la fréquence de Nyquist sur l'axe fréquentiel vertical, tandis que l'axe temporel horizontal est continuellement rafraîchi à l'aide d'un curseur balayeur. L'amplitude des raies fréquentielles est traduite par un ensemble de couleurs (ou différentes intensités du gris).

L'exemple de la figure 4-13, montre les sonagrammes à bande étroite (en haut) et à large bande (en bas) de la phrase : « As tu vu ce fameux lapin », traitée dans les fichiers associés GRM issus respectivement, des deux unités application1 et application2. L'exemple de la phrase réunit une grande diversité de phonèmes.

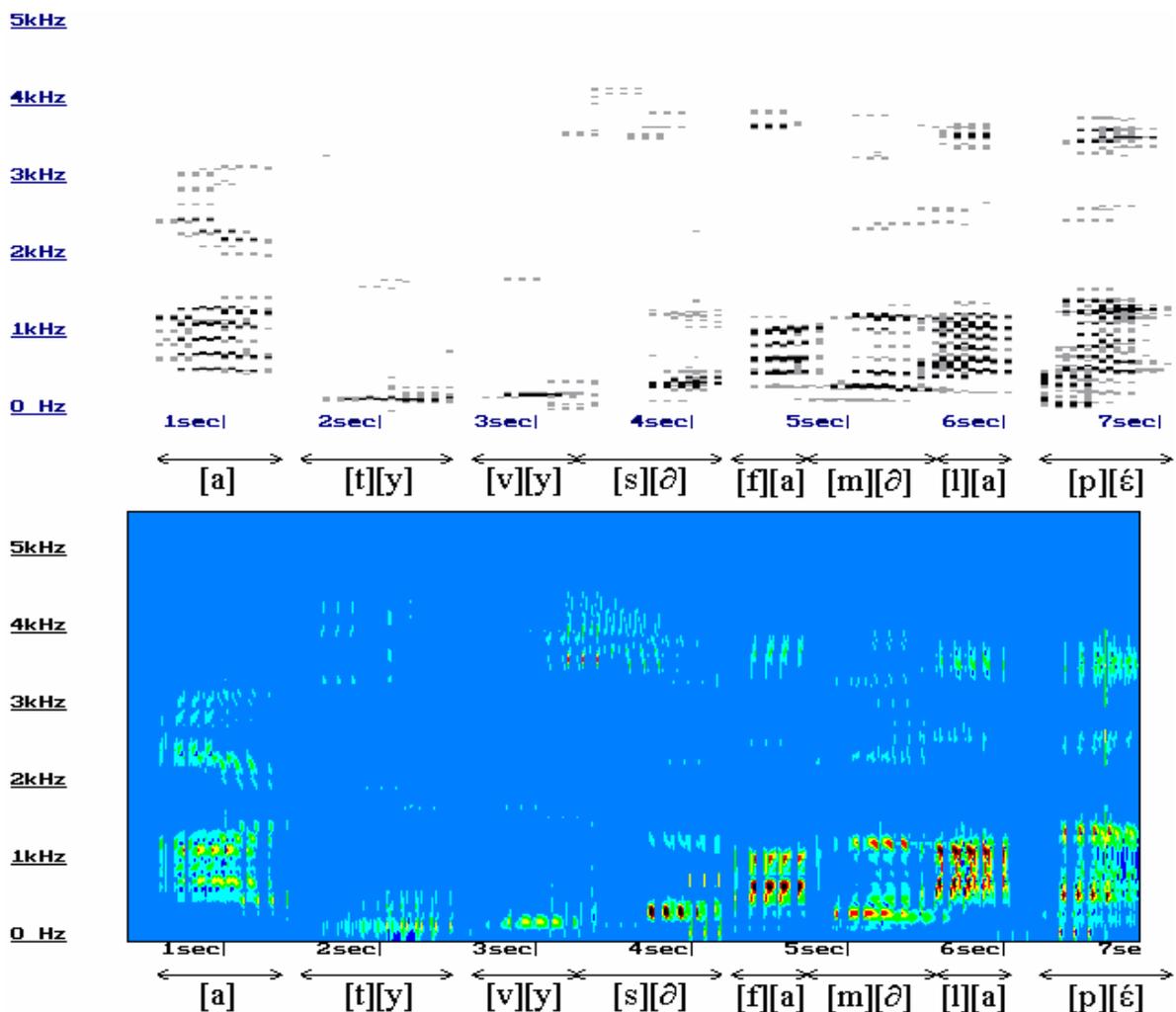


Figure 4-13: Sonagrammes à bande étroite (en haut) et à large bande (en bas) de la phrase : « As tu vu ce fameux lapin », du fichier GRM

* *Procédure ReverseBits* :

Cette fonction indispensable au traitement par FFT, trie les valeurs d'une séquence d'entrée (temporelle) suivant l'ordre binaire réfléchi (entrelacement temporel en papillon). Ceci permettra d'obtenir par la suite, une séquence de sortie (FFT) dans l'ordre naturel.

* *Procédure FFT* :

Cette procédure commence d'abord par une préaccentuation et pondération par la fenêtre de Haming du paquet d'échantillons à traiter. Puis calcule la transformée de Fourier discrète selon l'algorithme FFT qui convertit une fonction du temps à valeurs complexes échantillonnées, en une fonction à valeurs complexes de la fréquence, également échantillonnées selon (4.34). Comme le signal vocal échantillonné est une fonction à valeurs réelles, on fixe donc toutes les parties imaginaires d'entrée à zéro. La FFT fournit alors en sortie des valeurs redondantes de fréquences positives et négatives, c'est-à-dire que les valeurs correspondantes sont conjuguées complexes. Dans ce cas là, le premier élément de sortie contient la valeur moyenne des entrées, tandis que toute l'information dont on a besoin se trouve dans la première moitié des tableaux de sortie.

Les tableaux de valeurs, d'entrée et de sortie, pour cette procédure doivent avoir la même taille « n ». La valeur de cette taille devant être une puissance entière positive de 2, notre choix était limité à $2^9 = 512$ (voir procédure Print). Ainsi, il faut s'arrêter à l'indice $n / 2$ de la fréquence, correspondant à la fréquence de Nyquist (au-delà et jusqu'à la fréquence d'échantillonnage, s'étend l'axe fréquentiel négatif, conjugué de l'axe positif).

La taille des fichiers audio dépasse largement les 512 points que peut contenir les tableaux d'entrée-sortie. La transformée de Fourier à court terme utilisée dans notre programme, balaye le fichier audio par paquet de 512 échantillons (cas du sonagramme à bande étroite), ou 128 complétés à 512 par des zéros (cas du sonagramme à large bande), en utilisant une fenêtre d'analyse de Haming de durée $T = 46.5$ ou 11.6 ms, respectivement.

* *Procédure Test* :

Cette procédure permet de visualiser au choix, des signaux périodiques à différentes fréquences comme le signal sinusoïdal, triangulaire, carré ou même glottique, figure 4-9. Ces signaux à formes et spectres bien connus, constituent une référence pour juger de la qualité d'analyse temporelle et spectrale, apportée par notre programme.

Une autre généralisation de cette procédure vise à ajouter du bruit sinusoïdal, d'amplitude et de fréquence réglables, au signal de test choisi. La figure 4-14, montre l'exemple d'un signal sinusoïdal à 200 Hz contaminé par un signal parasite à 2 kHz.

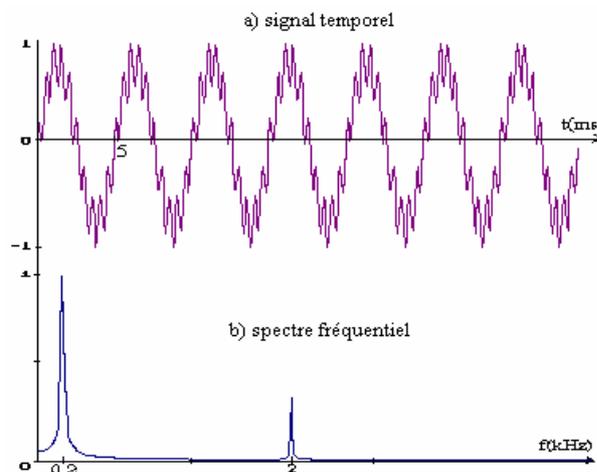


Figure 4-14: Spectre en fréquence d'un signal sinusoïdal parasité

conclusion:

L'objectif de ce chapitre est l'étude et l'analyse de la parole, commençant par son mécanisme de production et aboutissant à sa modélisation. Les transformées en ondelettes, malgré leur aspect prometteur, n'ont pas encore bénéficié d'analyse poussée jusqu'au stade de paramétrisation complète du signal de la parole, comme c'est le cas pour la transformée de Fourier. Cette dernière a été depuis longtemps considérée comme l'outil fondamental générant les spectres des signaux analysés. Dans le cas des signaux variants dans le temps comme la parole, la FFT à court terme constitue l'une des clés dans l'évaluation spectrale et la production de sonagrammes en temps réel. Dans notre cas, l'analyse par FFT glissante a été menée pour obtenir au choix, des sonagrammes à bande étroite ou à large bande. Cette dernière option, faisant apparaître les formants de la parole, a été sélectionnée afin de valider les résultats obtenus dans une analyse par modélisation de la parole. Cette comparaison spectrale sera menée au dernier chapitre.

La modélisation est un codage auto-régressif à court terme de la parole, qui est retenu pour notre stratégie d'implant cochléaire. Ce codage sert à une bonne paramétrisation du signal de la parole, et réduit le débit de transmission vers l'implant cochléaire en n'utilisant que les paramètres essentiels dans la stimulation électrique. La notion de débit est importante si l'on compare avec la transmission téléphonique classique sur une ligne RNIS qui exige un débit minimal de $8 \text{ kHz} \times 8 \text{ bits} = 64 \text{ kb/s}$. De même, la transmission ou l'enregistrement d'un signal audio de meilleure qualité, exige en principe un débit de l'ordre de $48 \text{ kHz} \times 16 \text{ bits} = 768 \text{ kb/s}$ (à multiplier par deux pour un signal stéréophonique). Alors qu'avec un codage paramétrique, il peut s'abaisser jusqu'à 1.2 kb/s .

CHAPITRE 5: Implémentation de la nouvelle stratégie d'implant sur DSP

5.1. Algorithmes de traitement

La prédiction linéaire comme vu dans le chapitre précédent, est une méthode de déconvolution directe qui sépare le signal d'excitation, de la fonction de transfert qui module ce signal. Cette fonction est souvent un modèle tous pôles autorégressif (AR) où chaque pôle correspond aux caractéristiques d'un formant, c'est à dire à sa fréquence et à sa bande passante. Le nombre de pôle de cette fonction donne l'ordre du modèle. L'ordre de prédiction P étant lié au nombre de formants N_F par la relation [14]: $P=2N_F+2$, il doit être adapté au nombre de formants. Un ordre de prédiction exagéré conduit à une modélisation trop fine du spectre, prenant en compte ainsi les harmoniques du fondamental. De ce fait, l'ordre est pris généralement égal à 10 ou 12 [14, 74]. Les trois premiers formants sont indispensables pour caractériser le spectre vocal; les deux suivants sont utiles pour une synthèse de qualité.

Pour un conduit vocal dont la longueur est de l'ordre de 17 cm, on peut observer avec $P=10$, jusqu'à 4 formants entre 100 et 5000 Hz.

L'estimation des formants et de leurs bandes passantes peut être obtenue par une recherche des maxima du spectre du modèle (méthode du « peak picking » [74]). Pour cela on applique la transformée de Fourier à la suite des coefficients $a_p(i)$ prolongée par des zéros jusqu'à N points, de façon à obtenir une puissance de 2 compatible avec la résolution spectrale souhaitée :

$$\{1, a_p(1), a_p(2), \dots, a_p(p), 0, \dots, 0\} \xrightarrow{\text{FFT}} A(k) = \sum_{i=0}^{N-1} a_p(i) e^{-j i k \frac{2\pi}{N}} \quad 5.1$$

Cette suite représente en effet la réponse impulsionnelle du filtre inverse $A_p(z)$. On prend ensuite, l'inverse des valeurs $|A(k)|$ pour en rechercher les maxima. La localisation d'un maximum peut être améliorée par une interpolation parabolique entre les valeurs de $1/|A(k)|$, [14].

Au lieu de calculer le spectre d'amplitude correspondant au modèle AR, et d'en déterminer les fréquences correspondant aux pics spectraux par détecteur de pics, la méthode la plus directe dans l'estimation des formants et leurs bandes passantes repose sur la factorisation du polynôme prédictif et l'utilisation de ses racines. Les résultats sont plus précis aux prix d'une certaine complexité d'extraction des racines. Nous avons suivi cette méthode qui sera exposée prochainement.

5.1.1. Prétraitements du signal de la parole

5.1.1.1. Recouvrement et pondération :

L'application du modèle autorégressif au signal de la parole est basée sur les résultats obtenus pour les signaux stationnaires. Comme le signal de la parole n'est pas stationnaire, on doit travailler sur des tranches de temps de durée limitées, avec recouvrement, figure 5-1.

Nous surveillons donc l'arrivée des échantillons de la parole sur des durées de 20 ms, voir paragraphes 2-6-1 et 4-1-2. Chaque paquet reçu, sera traité en chevauchement avec les deux paquets qui le précèdent. En d'autre terme, le traitement portera sur des tranches de 60 ms.

On pondère les échantillons du signal vocal par une fenêtre $w(n)$ symétrique. La fonction de pondération la plus courante est la fenêtre de Haming définie par :

$$w(n) = 0.54 - 0.46 \cos\left(\frac{2\pi n}{N-1}\right) \quad 0 \leq n \leq N-1 \quad 5.2$$

On montre [74] que la réponse fréquentielle d'une telle fenêtre présente une largeur de bande approximativement double de celle procurée par une fenêtre rectangulaire de même longueur :

$$B(\text{Hz}) = \frac{2F_e}{N} \quad 5.3$$

Où F_e est la fréquence d'échantillonnage du signal $x(n)$. En plus, la fenêtre de Haming présente

une atténuation plus prononcée en dehors de la bande passante par rapport à une fenêtre rectangulaire comparable.

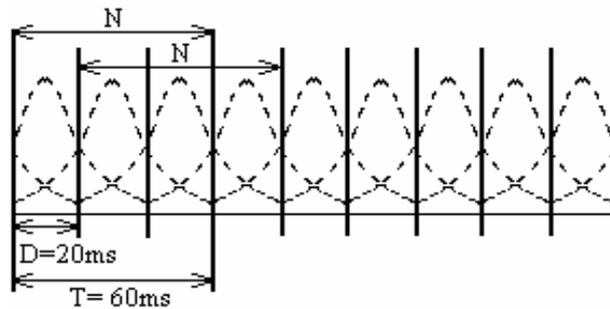


Figure 5-1: Découpage en tranches avec recouvrement et pondération

5.1.1.2. Préaccentuation :

Un autre prétraitement éventuel est la préaccentuation du signal. C'est une opération qui vise à accentuer la partie haute du spectre. Ce prétraitement assure une certaine compensation de la pente globale négative du spectre, observée généralement en fonction de la fréquence.

Cette pente négative du signal est due aux influences de la source d'excitation et du rayonnement des lèvres, lequel signal décroît de 6 dB/octave. Pour cela, on passe ce dernier par un filtre de préaccentuation, qui a pour fonction de transfert : $1 - \mu z^{-1}$, avec μ compris entre 0.8 et 1 [14].

L'équation de récurrence est de la forme : $\dots Y(i) = X(i) - \mu X(i - 1)$ 5.4

Où $X(i)$ représentent les échantillons de la parole.

Dans la norme GSM, la constante μ est égale à 0.86, et la réponse du filtre fréquentielle est montrée dans la figure 5-2.

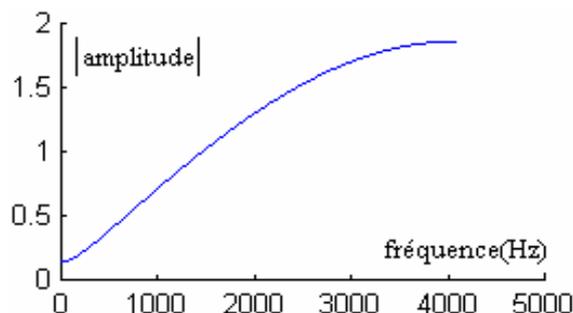


Figure 5-2: Réponse d'un filtre de préaccentuation

5.1.2. Traitements du signal de la parole

L'algorithme de la figure 5-4, présente l'analyse et le traitement apporté aux échantillons prétraités du tableau $Y[]$.

Cinq algorithmes jouant le rôle de détecteurs ou senseurs des paramètres de la parole, sont employés par le sous-programme modélisation dans le traitement des paquets d'échantillons prétraités reçus :

1. La segmentation Silence / Parole : Identifie la présence ou l'absence du signal vocal, ainsi que les silences occlusifs.
2. La mesure du gain du modèle ou valeur efficace vraie du signal vocal.
3. La mesure des formants et de leurs bandes.
4. La segmentation Voisé / Non-voisé : Permet de détecter la présence d'une excitation glottique.
5. La mesure de la fréquence fondamentale ou pitch.

dans fichier
analyse.c:

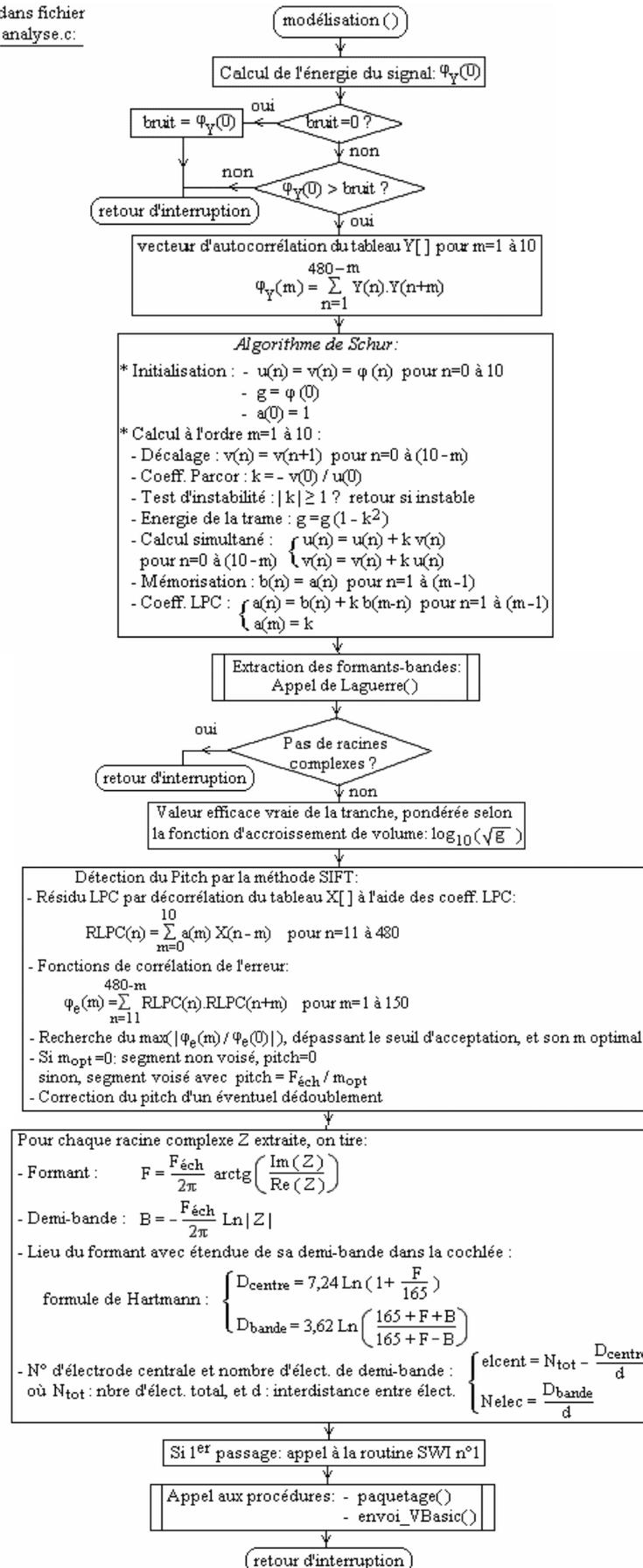


Figure 5-4: Organigramme de traitement de base de la parole

5.1.2.1. Senseur Silence / Parole :

Le senseur d'identification des segments Silence et Parole, se base sur la comparaison des énergies du signal avec celle du bruit. Ainsi une première mesure, pendant un temps d'une seconde, du maximum de l'énergie du signal environnant, est considérée être celle du bruit de fond. Cette quantité est à comparer à chaque nouvelle mesure pour décider si c'est encore du bruit (cas d'énergie au dessous de la première mesure), ou s'il s'agit d'un segment de la parole (cas d'énergie au dessus de la première mesure). La mesure d'énergie du signal est basée sur le calcul de la fonction d'autocorrélation des échantillons du tableau $Y[]$.

5.1.2.2. algorithme de Schur :

Après avoir calculer d'abord, les 11 fonctions d'autocorrélation à partir des 480 échantillons préaccentués et pondérés du tableau $Y[]$ et qui servent à initialiser entre autres, l'énergie g de la tranche de la parole analysée; l'algorithme de Schur est exécuté sous forme d'une boucle d'incrémentation de l'ordre m du modèle jusqu'à sa valeur finale $P=10$. Il s'agit d'une solution parmi d'autres, des équations de Yule-Walker établies au chapitre 4. Cet algorithme est montré dans l'organigramme de la figure 5-4.

L'opération la plus consommatrice en temps de calcul est l'établissement de la matrice $\Phi^{(P)}$. On commence donc par le calcul de la fonction d'autocorrélation à l'aide de la formule tronquée suivante [14] :

$$\phi(k) = \sum_{n=0}^{N-k-1} x(n) x(n+k) \quad k=0,1,\dots,p \quad 5-$$

5

Cette expression correspond à une estimation de la fonction d'autocorrélation dite à court-terme (on n'a pas divisé par le nombre d'échantillons N , ce qui ne modifie nullement les rapports en $\phi(k)$ comme dans le calcul des coefficients parcor).

Soient :

- la variable g représentant la variance minimisée ou l'énergie résiduelle de prédiction $\sigma^2 = \alpha_p = g$,
- les quatre vecteurs de coefficients :

$$\begin{cases} u = [u(0), u(1), \dots, u(p)]^T \\ v = [v(0), v(1), \dots, v(p)]^T \\ A = [a(0), a(1), \dots, a(p)]^T & \text{représen tant les } a_m(i) \\ B = [b(0), b(1), \dots, b(p)]^T & \text{représen tant les } a_{m-1}(i) \end{cases}$$

L'algorithme est comme suit :

1- initialisation : $g = \phi(0)$

$$u = v = [\phi(0), \phi(1), \dots, \phi(p)]^T$$

2- récursion sur l'ordre: pour $m = 1, 2, \dots, p$:

- pour $i=0, 1, \dots, p-m$:
 $v(i) \leftarrow v(i+1)$
- $k_m = -v(0) / u(0)$ { coefficient parcor }
 si $|k_m| > 1$: { instabilité : traiter une nouvelle tranche d'échantillons }
- pour $i=1, 2, \dots, m-1$: { non parcourue pour $m=1$ }
 $a(i) = b(i) + k_m b(m-i)$
- $a_m(m) = k_m$
 $g \leftarrow g (1 - k_m^2)$
- pour $i=1, 2, \dots, m$:

- $b(i) \leftarrow a(i)$ {mémorisation pour le prochain ordre m }
- pour $i=0, 1, \dots, p-m$:
 - $u(i) \leftarrow u(i) + k_m v(i)$
 - $v(i) \leftarrow k_m u(i) + v(i)$

3- résultats :

pour l'ordre final $m=p$, on obtient :

- les $a(i)$ avec $i=1$ à p et $a(0) = 1$: qui représentent les paramètres LPC (calculés par l'intermédiaire des coefficients parcor k_m).
- $\alpha = \sqrt{g}$: qui représente le gain du modèle autorégressif, fixant l'amplitude RMS du signal, voir l'expression (4.27).

5.1.2.3. Mesure de la valeur efficace vraie d'intensité :

A la fin de l'algorithme de Schur, nous obtenons par récurrence à l'ordre $P=10$, la variance minimisée de l'erreur g . Cette dernière fournit par extraction de sa racine carrée, le facteur de gain représentant la valeur RMS (valeur efficace vraie) de l'intensité de la tranche de parole analysée. Avant d'être transmise, l'intensité est pondérée par la fonction d'accroissement de volume (Loudness Growth Function) [38].

Comme : $RMS = \sqrt{g}$, et en combinant les expressions (2.1) et (2.2), tout en fixant $RMS_{min}=1$, il

vient :

$$I = \frac{MCL - THR}{\log_{10}(RMS_{max})} \log_{10}(RMS) + THR \quad 5-6$$

Où MCL et THR, sont les seuils en courant de stimulation assurant respectivement une perception minimale et maximale au patient implanté.

Les intensités I , MCL, et THR deviennent normalisées par rapport au courant maximal de stimulation toléré par le patient, lorsqu'on divise les deux membres de l'égalité de l'expression 5-6 par ce courant maximal.

Un programme réalisé, fournit un exemple d'une telle fonction pour une intensité I normalisé à 1 et un $RMS_{max} = 150$, figure 5-5.

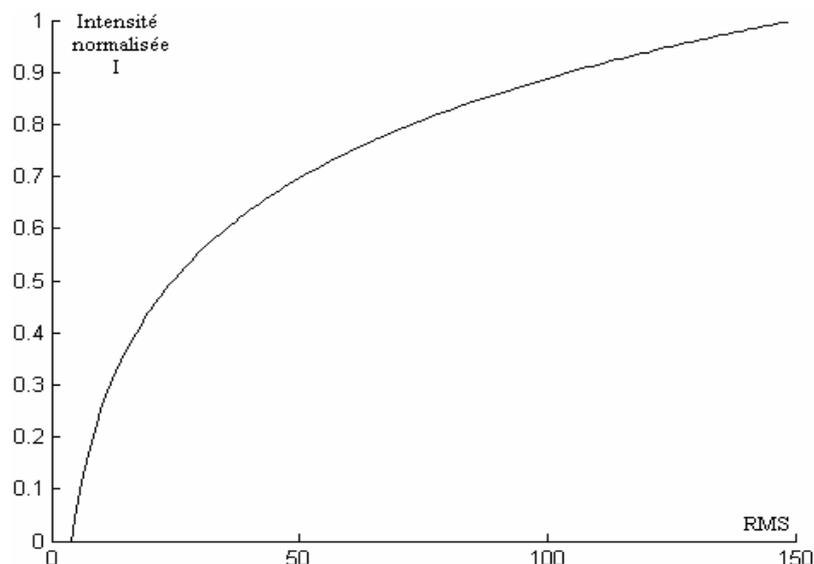


Figure 5-5: Allure d'une fonction d'accroissement de volume

5.1.2.4. Mesure des formants et bandes :

Il a été mentionné que pour un ordre de prédiction $P=10$ du polynôme prédicteur $A_p(z)$, on peut tirer jusqu'à 4 formants avec leurs bandes dans la gamme fréquentielle de 0.1 à 5 KHz. La méthode la plus directe et la plus précise dans l'estimation des formants et leurs bandes passantes, repose sur la factorisation du polynôme prédicteur afin d'extraire ses racines :

$$A_p(z) = 1 + \sum_{k=1}^p a_k z^{-k} = \prod_{k=1}^p (1 - z_k z^{-1}) \quad 5-7$$

Du moment que les coefficients LPC a_k sont réels, les racines z_k du polynôme $A_p(z)$ sont soit réelles, soit sous forme de paires conjuguées complexes. En passant au plan de Laplace, on a [14] :

$$(z_k, z_k^*) = e^{-\pi B_k T} e^{\pm j 2\pi F_k T} \quad 5-8$$

où T est la période d'échantillonnage du signal, et (F_k, B_k) sont les couples formant-bande.

de (5-8), on déduit :

$$\begin{cases} B_k = \frac{-1}{\pi T} \operatorname{Ln}|z_k| \\ F_k = \frac{1}{2\pi T} \operatorname{arctg}\left(\frac{\operatorname{Im}(z_k)}{\operatorname{Re}(z_k)}\right) \end{cases} \quad 5-9$$

Puisque le conduit vocal humain naturel est stable, on en déduit $|z_k| < 1$ [14]. En d'autres termes, toutes les racines du polynôme prédicteur doivent être à l'intérieur du cercle unitaire, selon le critère de stabilité.

5.1.2.5. Algorithme de Laguerre [72] :

Il y a deux principales méthodes sur la façon de chercher les racines de polynôme d'ordre élevé.

- *Méthode de la Valeur propre :*

Basée sur l'idée que les valeurs propres d'une matrice \mathbf{A} , sont les racines du polynôme caractéristique $P(x) = \det[\mathbf{A} - x \mathbf{I}]$. Ainsi, nous déduisons les racines d'un polynôme quelconque connaissant les valeurs propres de ses matrices associées. Typiquement, les routines basées sur les valeurs propres sont moins rapides d'un facteur 2, comparées à celle de Laguerre. Cette méthode est considérée comme alternative recommandée lorsque celle de Laguerre devient inefficace.

- *Méthode de Laguerre :*

Il s'agit d'une méthode qui converge à tous les types de racines: réelle, complexe, simple, ou multiple. La méthode de Laguerre est de loin, la plus fiable et la plus rapide des méthodes complexes existantes. Elle exige un calcul complexe, même en convergeant vers des racines réelles. Cependant, pour des polynômes à racines complexes, la convergence absolue n'est pas toujours prouvée. L'expérience montre, que la non convergence est beaucoup moins fréquente et qu'elle est souvent résolue par une simple cassation des cycles limites qui apparaissent (dans notre cas, ces cycles sont rares et toujours ignorés).

Pour dériver les formules de Laguerre, noter les relations suivantes entre le polynôme, ses racines et ses dérivés :

$$P_n(x) = (x - x_1) \cdot (x - x_2) \dots (x - x_n) \quad 5-10$$

$$\operatorname{Ln}|P_n(x)| = \operatorname{Ln}|x - x_1| + \operatorname{Ln}|x - x_2| + \dots + \operatorname{Ln}|x - x_n| \quad 5-11$$

$$\frac{d \operatorname{Ln}|P_n(x)|}{dx} = \frac{1}{x - x_1} + \frac{1}{x - x_2} + \dots + \frac{1}{x - x_n} = \frac{P_n'}{P_n} = G \quad 5-12$$

$$-\frac{d^2 \text{Ln}|P_n(x)|}{dx^2} = \frac{1}{(x-x_1)^2} + \frac{1}{(x-x_2)^2} + \dots + \frac{1}{(x-x_n)^2} = \left[\frac{P_n'}{P_n} \right]^2 - \frac{P_n''}{P_n} = H \quad 5-13$$

On assume que la racine x_1 que nous cherchons est à localiser à une certaine distance a de notre position actuelle x , alors que toutes les autres racines sont assumées être situées à une distance b : $x - x_1 = a$; $x - x_i = b$ pour $i = 2, 3, \dots, n$.

Ainsi, on peut écrire :

$$\frac{1}{a} + \frac{n-1}{b} = G \quad 5-14$$

$$\frac{1}{a^2} + \frac{n-1}{b^2} = H \quad 5-15$$

Ce qui fournit comme solution pour a :

$$a = \frac{n}{G \pm \sqrt{(n-1)(nH - G^2)}} \quad 5-16$$

Où le signe doit être choisi pour assurer la plus grande valeur au dénominateur. Comme le facteur sous racine carrée peut être négatif, a peut être complexe.

La méthode fonctionne itérativement : pour une valeur d'essai x , a est calculé. Puis $x-a$ devient la prochaine valeur d'essai. Ceci continue jusqu'à ce que a soit suffisamment petit. Remarquer que dans l'organigramme de la figure 5-6, a est noté dx .

Pour un polynôme donné d'ordre m , dont les coefficients peuvent être complexes, le premier coefficient $a[0]$ est le terme constant, alors que $a[m]$ est le coefficient de la puissance la plus

élevée de x :
$$P_m(x) = \sum_{i=0}^m a[i]x^i .$$

Dans notre cas, le polynôme prédicteur à factoriser est de la forme :
$$A_p(z) = \sum_{i=0}^P a[i]z^{-i} .$$

Ainsi pour $m=P$, chaque racine x de $P_m(x)$ trouvée, est l'inverse d'une racine z de $A_p(z)$.

La routine implantée, figure 5-6, met en application une version d'arrêt simplifiée due au critère d'Adams [72], qui répond d'une manière élégante au compromis d'une précision poussée, et du danger de réitérer pour toujours en présence d'erreur d'approximation.

En fixant l'erreur de calcul à 1%, les valeurs numériques des différentes racines extraites par la méthode de Laguerre, ont été trouvées suffisamment précises comparativement à celles évaluées par le logiciel Matlab. Le polissage des racines a été jugé non nécessaire et pouvant allonger encore le temps d'exécution du programme. C'est pour cette même raison que les rares cycles limites qui apparaissent, ne bénéficient pas d'une boucle de cassation de cycle et sont seulement ignorés.

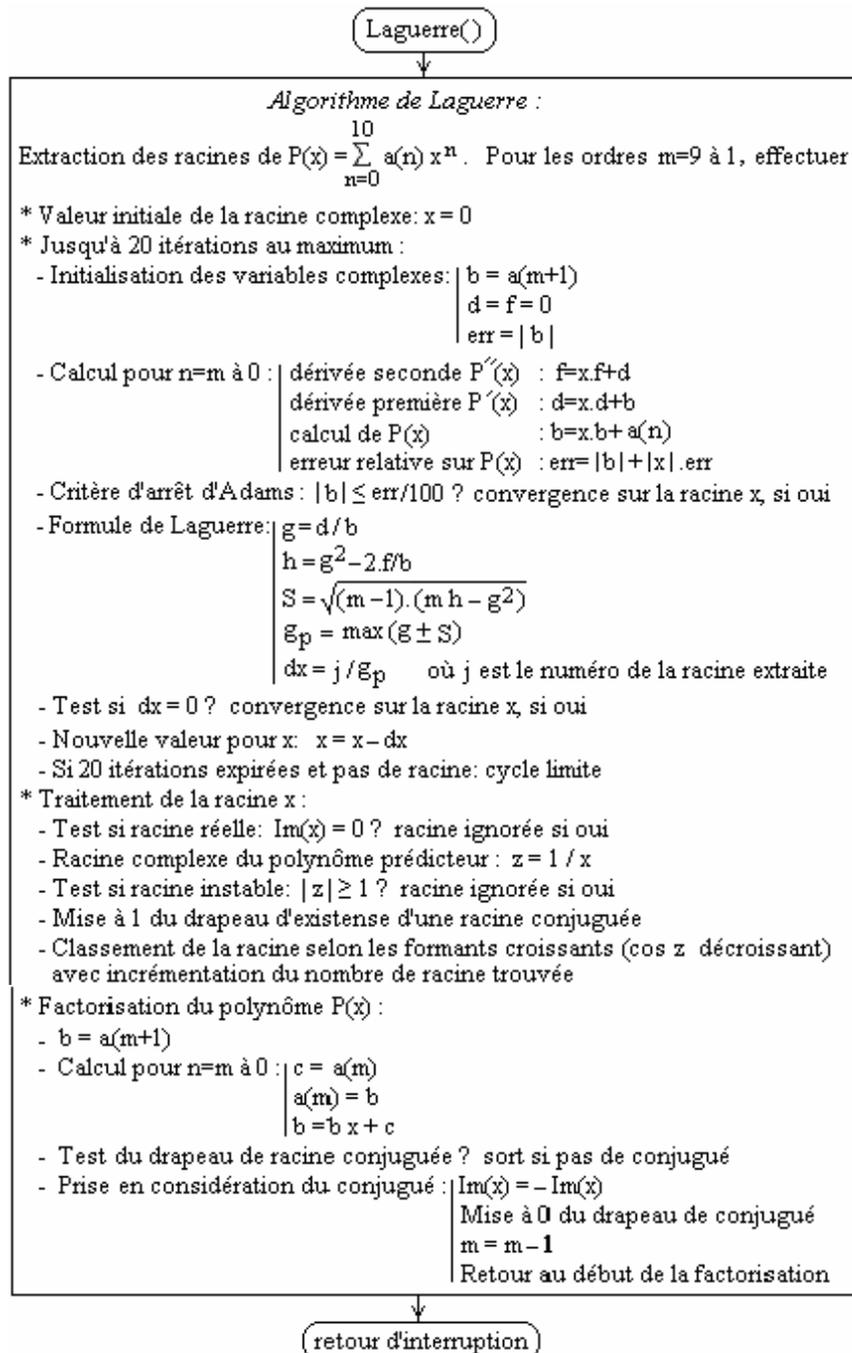


Figure 5-6: Algorithme d'extraction des racines complexes du polynôme prédicteur

Le programme d'application, appelle successivement la routine de Laguerre dans l'extraction de chaque racine x , pour ne retenir que les racines complexes inverses z dont le module est inférieur à 1. Une fois extraite, les racines sont classées selon l'ordre croissant des formants qu'elles véhiculent. En vue de reconnaître l'ordre d'un formant F_k ($k=1$ à 4 pour les quatre premiers formants), z_k de l'équation 5-7 est mise sous la forme :

$$z_k = z_{kr} + jz_{ki} = |z_k| \cos(2\pi F_k T) + j|z_k| \sin(2\pi F_k T) \quad 5-17$$

avec : $2\pi F_k T = 2\pi \frac{F_k}{F_{éch}} = [0 \text{ à } \pi]$ pour $F_k = \left[0 \text{ à } F_{Nyquist} = \frac{F_{éch}}{2} \right]$

Ainsi, pour :

1. $z_{ki} = 0 \rightarrow \begin{cases} \text{si } z_{kr} > 0 \Rightarrow F_k = 0 : \text{pas de formant} \\ \text{si } z_{kr} < 0 \Rightarrow F_k = F_{\text{Nyquist}} : \text{formant le plus élevé du spectre} \end{cases}$
2. $z_{kr} = 0 \rightarrow z_{ki} = |z_k| \Rightarrow F_k = F_{\text{Nyquist}} / 2 : \text{formant au milieu du spectre}$
3. autres cas : $\cos(2\pi F_k T) = z_{kr} / |z_k|$

Le classement des formants dans un ordre croissant, se ramène donc au classement des racines selon l'ordre décroissant de leur cosinus, figure 5-7.

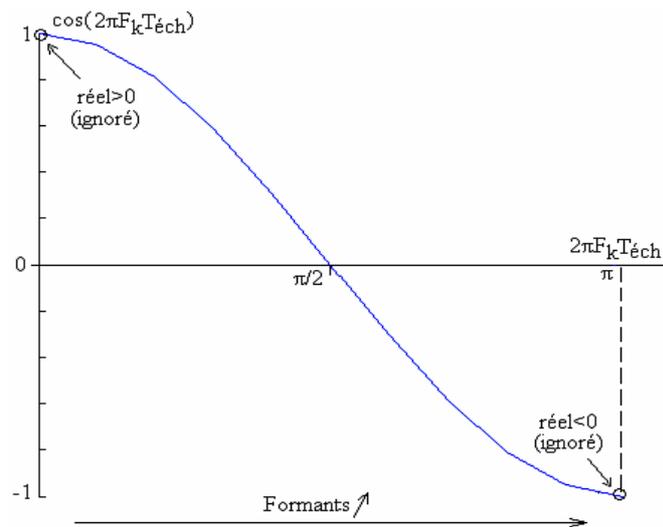


Figure 5-7: Classement en croissance des formants

5.1.2.6. Senseur de voisement et détection SIFT du Pitch :

Dans cette analyse, le signal résiduel est décorrélé afin de faire la séparation entre les signaux voisés et les non voisés. Le traitement s'effectue séquentiellement sur des trames de 480 échantillons.

La fréquence fondamentale F_0 (ou pitch) est un paramètre très important pour la synthèse de la parole, l'oreille est en effet très sensible à ses variations. La fréquence F_0 s'étend approximativement de 70 à 250Hz chez les hommes, de 150 à 400Hz chez les femmes et de 200 à 600Hz chez les enfants [10, 14]. Plusieurs méthodes existent pour sa mesure. Parmi les meilleures, on cite la méthode SIFT (Simplified Inverse Filter Tracking) [10, 14]. Elle fait appel à un filtre inverse $A(z)$ de faible ordre, généralement $P=4$, jugé suffisant pour modéliser le spectre du signal dans la gamme fréquentielle de 0 à 1 KHz, contenant généralement 1 à 2 formants. Dans cette méthode, Le signal d'entrée est filtré au-dessous de 900Hz et la fréquence d'échantillonnage est réduite d'un rapport de 5 par décimation. Ce signal est donc traité par le filtre inverse pour donner un résidu LPC (ou RLPC) débarrassé de la gênante contribution du conduit vocal. L'autocorrélation à court terme est appliquée à ce résidu afin de sélectionner le plus grand pic correspondant à la valeur du pitch.

Dans notre cas, la méthode SIFT a été modifiée en vue de simplifier son implantation sur DSP. Pour cela, on garde les mêmes échantillons utilisés dans l'extraction des formants, donc sans filtrage à 900 Hz ni décimation. Ceci est possible, car le filtre inverse calculé à partir de ces échantillons enlève, par déconvolution, la présence des quatre formants pour lesquels il a été dimensionné. Par conséquent la méthode de détection du pitch, retenue dans notre cas, est celle dérivée de SIFT à cause de sa relative simplicité d'implantation basée toujours sur la prédiction linéaire.

Ainsi, dans la suite du sous-programme modélisation(), figure 5-4, le signal obtenu après filtrage inverse des échantillons d'entrée est le signal d'erreur ou Résidu LPC :

$$RLPC(n) = \sum_{i=0}^{10} a(i) X(n-i) \quad \text{avec } n = 11 \text{ à } 480 \quad 5-$$

18

Où $a(i)$: sont les coefficients LPC du filtre inverse, et $X(i)$: sont les échantillons brutes du signal d'entrée.

Le signal résiduel a une forme de signal d'excitation qui parvient dans les cavités résonnantes de l'appareil phonatoire. Désormais, la structure du signal est théoriquement, presque identique à celle d'un bruit blanc (l'excitation produite par les poumons) sur lequel s'ajoute éventuellement un voisement (dû à la vibration des cordes vocales).

Pour mesurer le pitch, la fonction d'autocorrélation $\phi_e(m)$ du signal $RLPC(n)$ est exploitée [74]. Si le signal $RLPC$ est périodique de période n entière, alors $\phi_e(m)$ présente des maxima de même hauteur pour tout m multiple de n , figure 5-9. Donc l'autocorrélation permet de mettre en évidence les périodes du signal. En revanche, si le $RLPC$ est un bruit blanc, le $\phi_e(m)$ ne présente qu'un pic en $m=0$ car ce signal ne peut être corrélé qu'avec lui-même, figure 5-11.

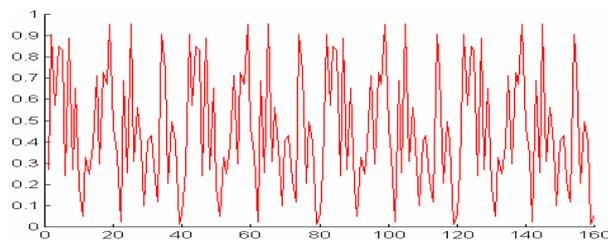


Figure 5-8: Signal RLPC de période 40 ms

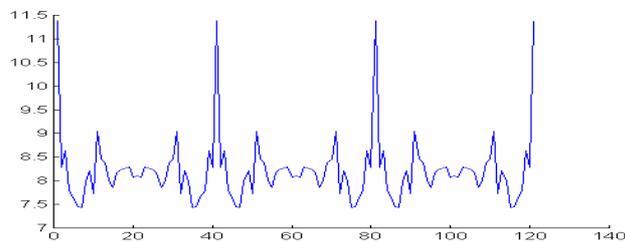


Figure 5-9: Fonction de corrélation du signal RLPC (périodique)

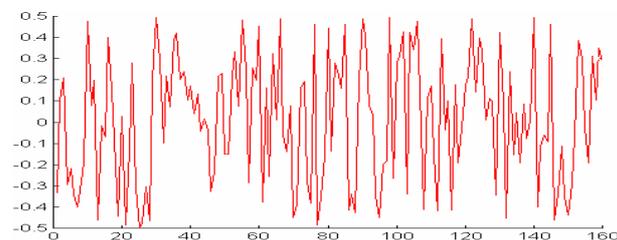


Figure 5-10: Signal RLPC comme bruit blanc

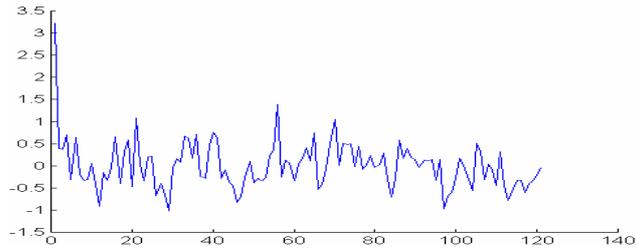


Figure 5-11: Fonction de corrélation du signal RLPC (bruit)

Dans la norme GSM [9], la détection des fréquences de voisement se limite de 67 à 200 Hz environ, ce qui correspond à des valeurs de m allant de 40 à 120 pour $F_{\text{éch.}}=8$ kHz. En effet, un pic à la position m correspond à l'intervalle de temps séparant deux impulsions de voisement. Ce temps vaut m fois la période d'échantillonnage, ou bien $m / F_{\text{éch.}}$.

Dans notre cas, nous avons étendu l'intervalle du pitch de 50 à 800 Hz. Pour pouvoir détecter le plus bas pitch à 50 Hz, le pas de calcul m de la fonction d'autocorrélation $\varphi_e(m)$, doit se faire

$$\text{jusqu'au rang: } m = \frac{F_{\text{éch.}}}{F_{\text{pitch}}} = \frac{8000 \text{ Hz}}{50 \text{ Hz}} = 160.$$

Ainsi, 160 fonctions d'autocorrélation sont calculées par la relation suivante :

$$\varphi_e(m) = \sum_{n=1}^{480-m} \text{RLPC}(n) \text{RLPC}(n+m) \quad \text{avec } m = 1 \text{ à } 160 \quad 5-19$$

Pour être accepté, un maximum (rapporté à $\varphi_e(0)$) doit dépasser un seuil que l'on fait varier avec m et dont le rôle consiste à éliminer un maximum non lié à la période du pitch. La loi proposée dans [14] est représentée en figure 5-12.

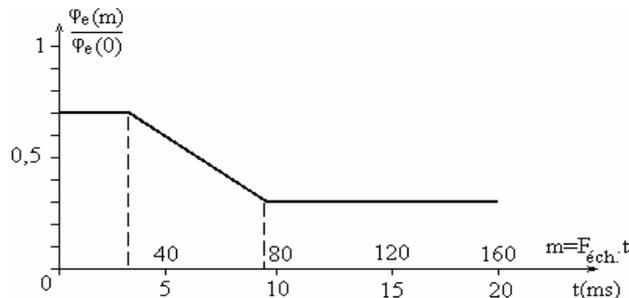


Figure 5-12: Seuil d'acceptabilité pour un maximum (SIFT)

La relation 5-19, détermine la valeur du pitch, en recherchant la valeur optimale de m qui maximise le rapport entre la fonction d'autocorrélation d'ordre m et celle d'ordre zéro du résidu RLPC :

$$\text{Pitch} = F_{\text{éch.}} / m_{\text{opt.}} \quad 5-20$$

Il s'agit par conséquent d'un segment voisé.

En revanche, si le RLPC est un bruit blanc, $\varphi_e(m)$ ne présente qu'un seul pic en $m=0$. Le segment est alors déclaré non voisé avec un pitch nul.

Toutefois, certaines mesures constatées du pitch arrivent au double de sa valeur moyenne. Cette erreur est fréquente dans les détecteurs de pitch [74], qui prennent le deuxième harmonique pour le fondamental. Une correction s'impose et l'idée sur laquelle elle se base, est la suivante :

Soit la variation du pitch au cours d'une conversation donnée : $\text{Pitch}(\text{Hz}) \in [P_{\text{min}}, P_{\text{max}}]$.

Initialement, on fixe P_{min} assez élevée, soit par exemple : $P_{\text{min}}=325$ Hz, avec un tampon du pitch P_{past} qu'on fixe initialement à : $P_{\text{past}} = 2 P_{\text{min}} = 650$ Hz. On suppose alors que $P_{\text{max}} \leq 2 P_{\text{min}}$, et l'algorithme résultant doit donc suivre la valeur minimale du pitch afin d'en limiter les valeurs maximales, voir organigramme en figure 5-13. Ainsi, l'identification précise de chaque période de voisement permet de suivre le fondamental lorsque celui-ci est sujet à de fortes variations.

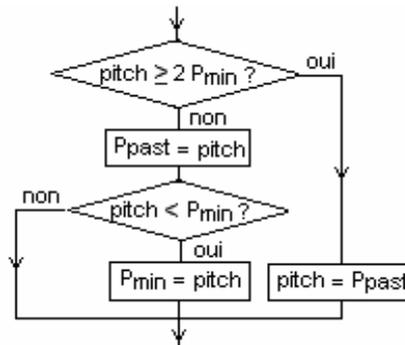


Figure 5-13: Correction éventuelle du pitch

5.2. Le processeur DSP de traitement

5.2.1. Choix du DSP

L'application pratique du modèle discret de la parole, implique un traitement en temps réel. L'implantation des algorithmes sur un microprocesseur à usage générale, risque de ne pas satisfaire à cette contrainte. Depuis quelques années, l'utilisation d'un DSP (Digital Signal Processor) dans les systèmes électroniques devient indispensable, par exemple dans les chaînes hi-fi, les téléphones cellulaires, etc...

Ce processeur est tout particulièrement adapté au traitement numérique du signal, domaine dans lequel un microprocesseur classique ou un microcontrôleur n'excelle pas. Pour le développement de notre application, nous avons utilisé au début le Starter-Kit TMDS3200051 de chez Texas-Instrument [90], qui se connecte au PC via un port RS232.

Le Kit comporte le DSP TMS320C50 à virgule fixe [97], et un circuit d'interface analogique AIC du type TLC32040. Le DSP est conçu pour exécuter 20 millions d'instructions par seconde (MIPS) au format de 16 bits. Il peut réaliser en un seul cycle machine (50 ns), des opérations simultanées telle que: accumulation, produit, et décalage d'un pointeur. Ce DSP comporte deux ALU, l'une à 16 bits et l'autre à 32 bits munie d'un multiplieur parallèle, un ensemble de 30 registres à 16 bits, une pile à 8 niveaux, un pipeline à 4 niveaux (quatre instructions simultanées par cycle), des périphériques internes (16 E/S parallèles, Timer programmable à 16 bits, Port série synchrone full-duplex, TDM ou port série à multiplexage temporel, 10 Kmots de 16 bits de RAM, 2 Kmots de 16 bits de ROM, etc...), et un jeu de 132 instructions.

Le convertisseur AIC est un convertisseur ADC/DAC à 14 bits de résolution. Il est relié au DSP par une liaison série et il est constitué d'un filtre passe-bande et d'un filtre anti-repliement, qui limitent la fréquence de conversion à un maximum de 70 kHz. La bande fréquentielle utile du signal, est limitée entre 300 Hz et 3.6 kHz. Du point de vue logiciel, le kit est fourni avec le compilateur DSK5A (assembler-linker) et le débogueur DSK5D.

Cependant, la majorité des calculs de notre application s'effectuent sur des nombres réels. Nous avons essayé de donner toute la précision nécessaire au calcul des coefficients LPC. Même si le DSP est à 16 bits en virgule fixe, un format à 32 bits a été utilisé : un mot à 16 bits était réservé à la partie entière du coefficient, tandis que l'autre renfermait sa partie fractionnaire. Les deux parties signées, formaient le nombre réel à 32 bits. Malgré le développement de plusieurs routines en assembleur pour réaliser des opérations sous ce format, telle que la multiplication, la division, la racine carrée, etc..., le calcul itératif dans la boucle principale de Schur divergeait (à la dixième itération) par manque de précision.

Nous nous sommes résigné d'utiliser un DSP plus puissant, travaillant à virgule flottante sur 32 bits. L'application est donc développée sur une carte de Texas Instrument, bâti autour du DSP TMS320C6711. IL utilise l'architecture VelociTI™ et la VLIW avancé (Very Long Instruction Word) de la famille C6000 [92], donnant une meilleure flexibilité et haute performance. Un pipeline à 4 phases permettant une exécution en parallèle de huit instructions par cycle. Ceci équivaut à 1,2 milliards d'opérations en virgule flottante par seconde à la fréquence horloge de

150 MHz [92]. Cette génération de DSP est munie d'un ensemble d'outils de développement software, incluant un compilateur C/C++ et un optimisateur d'assembleur, intégré dans un environnement de développement CCS (code composer studio).

5.2.2. Architecture du DSP

Le C6711 est constitué de trois parties principales : CPU, mémoires, et périphériques, figure 5-14.

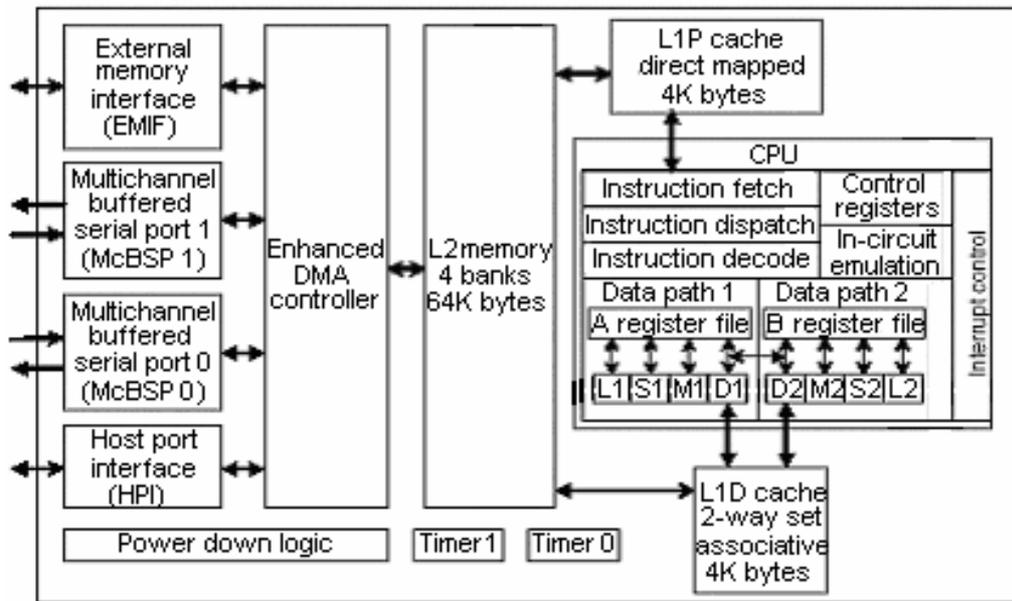


Figure 5-14: Constitution interne du DSP TMS320C6711

Nous nous intéressons aux périphériques internes utilisés dans notre réalisation, à savoir, les deux ports série McBSP (multi-channel buffered serial port), les deux Timers, et le DMA amélioré (EDMA). Ces périphériques sont importants dans la construction des systèmes de communication. Certaines données sur ces périphériques sont fournies ci-après.

5.2.2.1. Timer :

C'est un compteur programmable à 32 bits. L'horloge du Timer peut être dérivée à partir d'une source externe ou interne ($F_{CPU} / 4$). Les registres à 32 bits d'un Timer sont au nombre de trois [93].

5.2.2.2. McBSP :

Le port série multicanal bufférisé est un port série standard, composé d'un transmetteur et d'un récepteur autonomes. Le McBSP présente un chemin pour les données (DX pour la transmission et DR pour la réception), et un autre pour les signaux de contrôle (horloges CLKX/R et synchronisations de trames FSX/R). Le C6711 communique avec le McBSP à travers dix registres de contrôle à 32 bits [93].

La communication du C6711 avec son périphérique McBSP se fait de trois façons :

- scrutation (polling) du McBSP à travers RRDY et XRDY.
- interruption vers CPU à travers RINT et XINT.
- événement vers EDMA à travers REVT et XEVT.

tel que: RRDY, indique la présence d'une donnée à lire par le CPU ou EDMA, en véhiculant éventuellement l'événement REVT ou l'interruption RINT.

XRDY, indique que la donnée a été transmise et que c'est prêt à nouveau, il véhicule éventuellement l'événement XEVT ou l'interruption XINT.

5.2.2.3. EDMA (DMA amélioré) :

le C6711 exécute le transfert de données interne et externe, via le CPU ou le contrôleur EDMA. Typiquement le transfert d'un bloc de données ou celui des requêtes de périphériques se font à travers l'EDMA. Il soulage ainsi le CPU des tâches fastidieuses.

Seize canaux sont gérés individuellement par les événements (interruptions) correspondants, en déclenchant les transferts de données associés. L'attribution des canaux est donnée en table 5-1.

Canal EDMA	Événement	Description
0	DSP_INT	Interruption du hôte au DSP
1	TINT0	Interruption Timer 0
2	TINT1	Interruption Timer 1
3	SD_INT	Interruption EMIF-SDRAM
4	EXT_INT4	Interruption Externe pin 4
5	EXT_INT5	Interruption Externe pin 5
6	EXT_INT6	Interruption Externe pin 6
7	EXT_INT7	Interruption Externe pin 7
8	EDMA_TCC8	Interruption Transfert Complet de Code 8
9	EDMA_TCC9	Interruption Transfert Complet de Code 9
10	EDMA_TCC10	Interruption Transfert Complet de Code 10
11	EDMA_TCC11	Interruption Transfert Complet de Code 11
12	XEVT0	Événement transmission McBSP0
13	REVT0	Événement réception McBSP0
14	XEVT1	Événement transmission McBSP1
15	REVT1	Événement réception McBSP1

Table 5-1: Événements de synchronisation des canaux EDMA

A part les registres de contrôle qui sont au nombre de huit, les paramètres de transfert sont stockés dans des tables en mémoire interne du EDMA. Ainsi, PaRam est une table de configuration des canaux sur 2 Ko, répartie sur 85 entrées (une entrée représente 6 mots à 32 bits). Chaque entrée contient les paramètres de transfert propres à un événement (parmi les seize) ou les liaisons pour un autre événement, table 5-2.

31	16 15	0	
Options (OPT)			Mot 0
Adresse Source (SRC)			Mot 1
Nombre de paquets (FRMCNT)		Nombre d'éléments (ELECNT)	Mot 2
Adresse Destination (DST)			Mot 3
Index de paquet (FRMIDX)		Index d'élément (ELEIDX)	Mot 4
Recharge du nombre d'élément (ELERLD)		Adresse de liaison (LINK)	Mot 5

Table 5-2: Paramètre pour un événement EDMA

Les six mots de chaque entrée, sont :

OPT : options de transfert, fixant les niveaux de priorité des événements, la taille des éléments à transférer, le transfert à 1 ou 2 dimensions (1D/2D), les modes de mise à jour des adresses source/destination, le mode de synchronisation, la possibilité de linkage, et la validation d'interruption de fin de transfert.

SRC : adresse de départ pour le transfert de données avec une mise à jour programmable.

ELECNT : nombre d'éléments par paquet (transfert 1D ou 2D).

- FRMCNT : nombre de paquets par bloc-1 (transfert 1D ou 2D).
- DST : adresse d'arrivée pour le transfert de données avec une mise à jour programmable.
- ELEIDX : adresse d'index des éléments dans un paquet (transfert 1D).
- FRMIDX : adresse d'index des paquets dans un bloc (transfert 1D ou 2D).
- LINK : adresse de PaRam contenant les paramètres de liaison.
- ELERLD : valeur à charger à la fin de transfert de chaque paquet (transfert 1D).

Deux types de transfert de données existent :

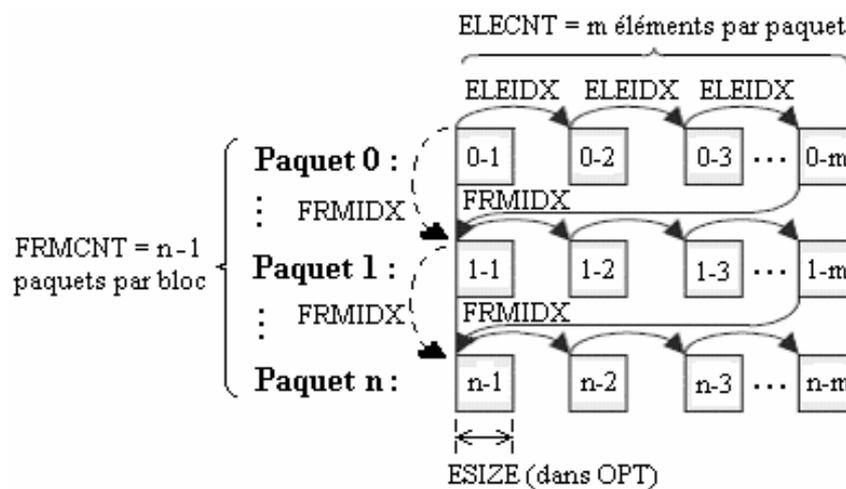


Figure 5-15: Types de transfert d'éléments

- Transfert à une dimension (1D) : se fait par paquet (à éléments espacés ou contigus). Pour une synchronisation par élément, on a ELECNT=0 à la fin du transfert d'un paquet. Pour un transfert multi paquet, il faut recharger ELECNT par ELERLD. C'est la synchronisation par paquet. Dans la figure 5-15, FRMIDX est désignée par des flèches continues pour un transfert d'un élément par événement. Alors que FRMIDX est désignée par des flèches discontinues pour un transfert d'un paquet (frame) par événement.

- Transfert à deux dimensions (2D) : se fait par bloc (à éléments toujours contigus, sans ELEIDX). Dans la figure 5-15, FRMIDX est désignée par des flèches discontinues pour un transfert d'un paquet ou 'array' par événement. Alors que FRMIDX est désignée par des flèches continues pour un transfert d'un bloc par événement.

L'enchaînement des canaux par événement peut être réalisé par deux procédés :

- Procédé par LINK, table 5-2 : les 16 bits de poids faible d'adresse de liaison LINK dans PaRam, spécifient l'adresse des paramètres à charger ou à recharger au prochain événement. Lier une entrée à elle-même exécute une autoinitialisation pour des transferts répétitifs.

- Procédé par chaînage : si le linkage recharge les paramètres du canal courant par les paramètres de liaison, le chaînage ne modifie pas les paramètres du canal mais fait passer au canal enchaîné. Ce dernier doit être sélectionné parmi les quatre canaux spécifiés par TCC=8, 9, 10 ou 11 dans OPT du canal appelant. Ainsi, on peut enchaîner successivement jusqu'à 4 canaux à la fin de transfert d'un canal donné suite à un événement unique.

5.2.3. Interruptions

Il y a huit registres de contrôle pour gérer les interruptions. Le DSP supporte 16 interruptions à priorité. L'interruption à plus haute priorité est INT_00 qui est dédiée à l'interruption RESET, alors que l'interruption la plus basse en priorité est INT_15. Les quatre premières interruptions (INT_00–INT_03) sont non masquables. Les autres interruptions (INT_04–INT_15) sont masquables et correspondent par défaut aux sources désignées dans la table 5-3. L'ensemble des périphériques du C6711 procure jusqu'à 32 sources d'interruptions, alors que le CPU n'en peut gérer que 12. Le sélecteur d'interruptions permet de choisir et d'établir l'ordre de priorité des interruptions.

Numéro d'INT.	Interruption	Source d'interruption
INT_00	RESET	–
INT_01	NMI	–
INT_02	–	Réservée
INT_03	–	Réservée
INT_04	EXT_INT4	Interruption externe au pin 4
INT_05	EXT_INT5	Interruption externe au pin 5
INT_06	EXT_INT6	Interruption externe au pin 6
INT_07	EXT_INT7	Interruption externe au pin 7
INT_08	EDMA_INT	Interruption EDMA (canal 0 à 15)
INT_09	XINT0	Interruption McBSP0 (transmetteur)
INT_10	SD_INT	Interruption EMIF-SDRAM-Timer
INT_11	RINT0	Interruption McBSP0 (récepteur)
INT_12	XINT1	Interruption McBSP1 (transmetteur)
INT_13	DSP_INT	Interruption HPI (vers DSP)
INT_14	TINT0	Interruption Timer0
INT_15	TINT1	Interruption Timer1
–	RINT1	Interruption McBSP1 (récepteur)
–	–	Réservée

Table 5-3: Interruptions du DSP (sources par défaut)

5.2.4. La carte DSK-C6711

5.2.4.1. Caractéristiques principales :

Le C6711 DSK est une plate-forme à interface parallèle avec l'hôte PC. Elle permet le test et développement efficace des applications basées sur le processeur TMS320C6711. Parmi ses caractéristiques, on cite :

- Un DSP TMS320C6711 à 150 MHz.
- Un contrôleur de port parallèle (PPC) interfacé à un port parallèle standard (EEP/SPP) du PC.
- 16 Mo de SDRAM (synchronous dynamic random access memory) à 100 MHz.
- 128 Ko de mémoire flash, programmable et effaçable.
- Un émulateur JTAG via le port parallèle avec support externe XDS510.
- AIC à 2 canaux de 16 bits (1 Codec pour la voix, 1 ADC/DAC pour les données).
- 1 connecteur d'extension mémoire J₁.
- 1 connecteur d'extension de périphériques J₃.

C'est ce dernier connecteur, J₃, qui est utilisé pour câbler notre transmetteur-récepteur à la carte DSP.

5.2.4.2. Codec TLC320AD535 :

Il s'agit d'un circuit à deux canaux indépendants, voix/donnée, avec les caractéristiques suivantes :

- Faible consommation sous 5 ou 3.3 V.
- Fréquence d'échantillonnage jusqu'à 11 kHz. Fixée à 8 kHz sur le kit.
- Bande passante à -3 dB, de 3.6 kHz.
- Tension d'entrée pic à pic de 3 V maximum.
- Format à 16 bits en complément à 2, disponible en série.
- Les voies ADC et DAC utilisent des modulateurs σ - Δ à haute résolution et faible bruit.
- Possibilité de test en boucle interne des voies ADC-DAC.
- Polarisation maximale du microphone de 5 mA sous 2.5 à 1.5 V.
- Mode stéréo supporté.
- Préamplification de 0 à 20 dB du microphone.
- Amplification programmable de -36 dB à $+12$ dB sur les voies ADC et DAC.

Le codec utilise deux types de communication :

- Communication primaire, se reportant au transfert des données converties.
- Communication secondaire, assurant la lecture ou l'écriture dans ses registres de contrôle.

Le premier bit D_0 est alloué au contrôle du type de communication : $D_0 = 0$ sélectionne la communication primaire, tandis que $D_0 = 1$ sélectionne la communication secondaire. Ainsi, les mots sont au fait de longueur 15 bits, avec D_0 dédié à la demande d'une communication secondaire.

Six registres de contrôle sont à programmer durant les communications secondaires : les registres n°1 et 2 sont réservés au canal « donnée », et les registres n°3 à 6 sont réservés au canal « voix ».

5.3. Logiciels d'accompagnement

Cette section discute les outils logiciels accompagnant la carte de développement à DSP, et qui sont intégrés dans son environnement de développement CCS (Code Composer Studio) [89]. Ces outils constituent le système opératoire DSP/BIOS et ses bibliothèques [95, 96]. Le DSP/BIOS est un ensemble d'API (Application Programming Interface) et d'outils pour le développement d'applications en temps réel sur DSP. Les API constituent des fonctions exécutables en temps réel et pouvant être utilisées efficacement dans des fichiers sources en langage C/C++.

La génération automatique de code peut être effectuée par un outil DSP/BIOS spécial de configuration, présentant une interface graphique. La bibliothèque CSL (Chip Support Library) [91] est une catégorie de cette interface graphique. Elle permet de configurer aisément les périphériques internes en programmant leurs registres, voire leurs champs en bits, dans des pages de propriétés. Le code de configuration résultant automatiquement, est inséré dans un fichier source en C prêt à être employé dans l'application.

Dans le cas d'un port McBSP ou composant Timer par exemple, l'interface inclut deux sections :

- Un gestionnaire de configuration : qui génère des objets de configuration (en nombre illimité) à partir d'une page de propriétés.
- Un gestionnaire de ressource : qui permet de sélectionner un périphérique et lui associer un objet de configuration. Pour un objet de configuration donné, l'option « Show Dependency » du menu principal montre quel périphérique l'utilise.

Le DSP/BIOS gère quatre niveaux distincts d'amorçage d'exécution (à priorités décroissantes) : les routines de service d'interruptions hardware, les interruptions software, les tâches, et les fonctions IDLE de fond (boucles sans fin avec économie d'énergie).

Parmi les autres catégories constituant l’outil graphique de configuration DSP/BIOS, et qui sont utilisées dans notre application d’implant, on trouve :

5.3.1. Catégorie «Scheduling»

Dans le Module «HWI-SR Manager», Voie «HWI_RESERVED1», est programmée la fonction `_RTDX_Poll` pour la communication RTDX (real time data exchange). Les canaux RTDX sont unidirectionnels dans leurs transferts de données, de la carte DSP vers le PC ou l’inverse. Le transfert de données est asynchrone. Le DSP appelle les fonctions de la librairie RTDX pour envoyer ou recevoir des données du PC via l’interface JTAG (liaison d’émulation), et ce, de façon transparente à l’application du DSP. Dans la plupart des applications, le RTDX utilise des interruptions pour accomplir le transfert des données. Ce transfert s’opère en temps réel pendant que l’application s’exécute. Sur PC, la librairie RTDX opère en conjonction avec l’environnement CCS de la carte DSP. Le choix des logiciels d’affichage sur PC inclut : Visual Basic, Visual C++, Excel, LabView, Matlab. Le CCS contrôle le flux de données entre le hôte (sur PC) et la carte DSP, figure 5-16.

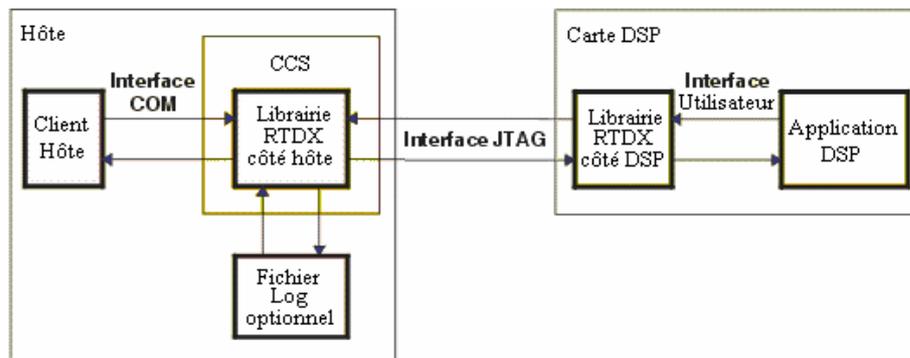


Figure 5-16: Flux de données par voie RTDX

Pour entretenir un flux de données du DSP vers le hôte, il faut déclarer un canal RTDX en sortie et utiliser des routines spécifiques appartenant à ‘User Interface’ pour y écrire les données à envoyer. Ces données sont immédiatement enregistrées dans un buffer cible défini dans la librairie RTDX (côté DSP). Les données du buffer sont alors envoyées vers le hôte via l’interface JTAG. La librairie RTDX (côté hôte) reçoit les données à partir de l’interface, et les enregistre soit dans un buffer en mémoire, soit dans le fichier `Log_RTDX` (dépendant du mode d’enregistrement spécifié sur le hôte). Le transfert de données à partir de l’application DSP vers la librairie RTDX du côté hôte est effectué en temps réel. Les données enregistrées peuvent être retrouvées par n’importe quelle application hôte permettant à l’utilisateur de visualiser ses données d’une manière plus représentative (graphique par exemple).

Similairement, pour que le DSP reçoive des données de la part du hôte, il faut déclarer un canal RTDX en entrée et postuler une demande de données en utilisant des routines définies dans ‘User Interface’. La requête de données est enregistrée dans un buffer cible et envoyée au hôte via l’interface JTAG. Les données sont transférées rapidement vers un buffer en mémoire au sein de la librairie RTDX du côté hôte. Lorsque cette librairie reçoit une requête de lecture à partir de l’application DSP, les données bufférisées lui sont envoyées via l’interface JTAG. Les données sont écrites dans la location requise par le DSP en temps réel sans interférer avec l’application. Le hôte notifie la librairie RTDX (côté DSP) lorsque l’opération est terminée.

Un contrôle graphique de la configuration RTDX, existe. Elle permet de valider ou inhiber la liaison RTDX, et d’accéder à la page des propriétés de configuration RTDX.

D’autres modules existent dans la catégorie ‘scheduling’, tels les modules d’interruptions qui sont de deux types :

Interruptions Hardwares :

Les événements hardware se déroulant dans l'environnement du DSP, peuvent déclencher une routine de service d'interruption (ISR). Les interruptions hardware sont utilisées pour des événements à temps critique, et le ISR doit être déclenché directement et traité rapidement. Par conséquent, il doit être court.

Le module HWI [95] est un gestionnaire d'interruption hardware renfermant des objets HWI prédéfinis, qui sont au nombre des 16 numéros d'interruptions de la table des vecteurs d'interruptions hardwares de la table 5-5. Il est possible de modifier les valeurs par défaut de la table des vecteurs d'interruption dans le but de réarranger les priorités des interruptions, de sélectionner une source d'interruption non utilisée, de valider/inhiber l'exécution d'interruptions, ou de choisir la polarité des interruptions externes sur les pins 4 à 7.

Avec l'outil de configuration, on peut assigner une routine à n'importe quelle interruption hardware. Cependant, certaines routines sont assignées automatiquement à certaines interruptions comme le cas de l'interruption du Timer0, configurée pour exécuter une fonction d'incrémenter temporelle du module CLK. Les routines d'interruptions peuvent être écrites en assembleur ou en C. Dans ce dernier cas, l'outil « HWI dispatcher » est fourni pour utiliser la même pile du système en incluant des macros d'encapsulation du code [95].

Interruptions Softwares :

Le module SWI [95] est un gestionnaire d'interruption software. Une interruption software est un objet qui encapsule une fonction à exécuter avec une priorité. Le module SWI sauvegarde automatiquement les registres du processeur, avant de servir une interruption soft à haute priorité interrompant une autre à basse priorité. Les registres sont restaurés en fin d'exécution afin d'achever la routine précédente à basse priorité, à condition que d'autres interruptions soft à plus hautes priorités n'ont pas été postées dès lors.

Vu le temps de réponse des SWIs, elles ne doivent être utilisées que pour des événements à longs traitements. Une routine ISR software est appelée lorsque un SWI est postée. L'appelant peut utiliser quelques options additionnelles dans la boîte-aux-lettres (mailbox) du ISR. Ces options sont vérifiées avant le postage du SWI. Ceci rend un ISR software plus flexible dans sa réponse à une interruption SWI.

Le gestionnaire des SWI valident ces propriétés afin qu'elles puissent être opérationnelles.

5.3.2. Catégorie « Instrumentation »

Le module «LOG-Event Manager» contient le gestionnaire de message Log, qui s'occupe de l'affichage d'informations sur les événements. On peut ainsi utiliser avantageusement l'API «LOG_printf» plutôt que l'instruction standard «Printf» du langage C, du point de vue rapidité d'exécution.

Conclusion :

L'énergie résiduelle et les coefficients de prédiction sont obtenus après les étapes de prétraitements et de traitements à court terme du signal de la parole. Le nombre relativement limité de ces paramètres prédits, suffit pour une représentation précise de l'enveloppe du spectre vocal. Le spectre du signal de la parole dépend de la transmittance du conduit vocal, du rayonnement à partir des lèvres (éventuellement des narines) et du spectre de l'excitation (impulsion glottique ou bien passage turbulent de l'air dans le conduit vocal).

L'estimation des formants principaux et de leurs sélectivités, constitue une tâche primordiale dans l'extraction des paramètres utiles à la stimulation cochléaire à partir d'un codage au plus faible débit de transmission.

Les paramètres jugés utiles à la stimulation cochléaire, sont : le pitch, l'intensité, et les quatre premiers formants avec leurs bandes. Dans leur extraction, un prétraitement est appliqué en premier lieu, opérant sur des tranches de 480 échantillons avec chevauchement sur 320 échantillons. Le traitement de chaque tranche de la parole par DSP, vient en second lieu, pour déterminer les paramètres du modèle autorégressif, et qui sont utilisés dans la stimulation de l'implant cochléaire. Les traitements sont répétés chaque 20 ms. Ils sont accomplis en temps réel, par un DSP à 32 bits du type TMS320C6711 à virgule flottante.

Dans le cadre de la transmission d'une information digitale, nous proposons de développer et concevoir un dispositif à la fois simple et fiable. Nous essayons dans la suite de ce chapitre de jumeler théorie et pratique de façon à emprunter les techniques les plus adaptées. Pour cela, notre première étape est consacrée à la conception et à la réalisation du circuit transmetteur répondant aux besoins de notre stratégie de stimulation cochléaire. La deuxième étape est quant à elle, réservée à la conception et réalisation du circuit récepteur. Une troisième étape s'occupe de la gestion software des données à transmettre et à recevoir, et ceci, via la carte DSP.

6.1. Conception et réalisation du transmetteur d'implant

6.1.1. Cahier des charges

L'envoi de données sans fil, du transmetteur vers le récepteur implanté, doit se faire sous les conditions suivantes :

1. Codage en ligne ou modulation en bande de base, permettant au récepteur une récupération facile de donnée avec une extraction aisée d'horloge.
2. Protection du message transmis à faible débit, fixé a priori à 4800 bauds.
3. Porteuse HF bien stable entre 1 et 3 MHz (minimum d'absorption du tissu biologique), avec amplificateur de sortie à haut rendement énergétique et faible encombrement.
4. Un transfert de puissance HF adéquat à travers un canal avec perte (tissu biologique). Pour cela, sera utilisée une liaison par couplage inductif sans noyau dans lequel peu de spires sont à bobiner du côté récepteur.
5. L'étage de démodulation au niveau du récepteur doit être de conception la plus simple possible (minimum d'encombrement).
6. Un générateur de données pseudo-aléatoires est à utiliser pour les différents tests et mesures (taux d'erreur binaire, diagramme en œil, etc ...).
7. Un compromis est à satisfaire entre une conception globale simple des deux circuits de transmission-réception et une liaison performante.

La qualité de transmission d'un système numérique [40] se réfère à la :

- Probabilité d'erreur par bit.
- Occupation spectrale du signal.
- Complexité du récepteur.

Pour ces trois caractéristiques, le cahier des charges prévoit aux points 1, 2 et 3, une puissance élevée de transmission liée à un codage avec protection des données transmises. Ceci afin d'augmenter le rapport signal/bruit au niveau du récepteur et par conséquent, réduire au maximum le taux d'erreur binaire.

On prévoit au point 6 du cahier des charges, un générateur de données aléatoires dans le but d'estimer le taux d'erreur du système de transmission, une fois conçu.

Le meilleur choix du codage ou modulation en bande de base, vise aussi à réduire l'encombrement spectral du signal et augmenter ainsi les chances d'une meilleure utilisation du canal de transmission.

Les points 4, 5, et 7, mettent l'accent sur le récepteur. Une réalisation simple et moins encombrante est exigée, tout en assurant une liaison fiable avec le transmetteur.

De façon générale, il s'agit selon le cahier des charges, de mettre en œuvre un système de transmission-réception simple et sans protocole de liaison.

Le schéma synoptique du système de transmission numérique qu'on va adopter sera selon le modèle de la figure 6-1.

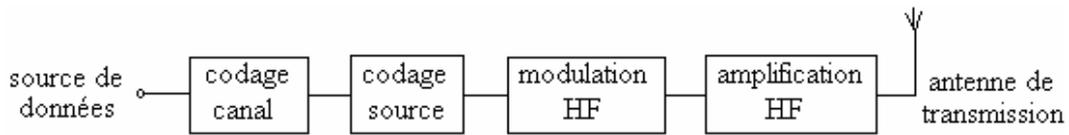


Figure 6-1: Schéma général d'une transmission HF de données

Dans ce modèle, la source émet un message numérique sous forme d'une suite d'éléments binaires. Le codage de canal est un codage détecteur d'erreurs ou correcteurs d'erreurs qui consiste à introduire de la redondance dans le signal émis, en vue de le protéger contre les perturbations et l'atténuation du canal de transmission. Le codage de la source est un codage en ligne dont la fonction est d'associer une suite d'impulsions physiques convenables à la suite d'éléments binaires. Il peut s'agir aussi d'un codage en bande de base, en vue de son adaptation avec un modulateur HF analogique. La modulation a pour rôle d'adapter le spectre du signal transmis au canal de propagation. L'amplification permet d'assurer la puissance requise à l'implant.

6.1.2. Codage de canal

Les systèmes de communication numériques modernes se servent typiquement de codes détecteurs d'erreurs voire correcteurs d'erreurs, pour améliorer leurs performances sous les conditions imparfaites du canal de transmission. Le codage de canal a pour but de sécuriser les paquets de données envoyés vers le récepteur.

Selon la théorie [48, 82], si deux mots de code successifs diffèrent de α bits (α est la distance de Hamming), on peut :

- détecter toute erreur portant sur $(\alpha-1)$ bits.
- corriger toute erreur portant sur $(\alpha-1) / 2$ bits.

Ainsi, le champ de correction d'erreur est de moitié de celui de la détection d'erreur. En plus, les codes auto-correcteurs nécessitent souvent des algorithmes de calcul fastidieux pour leur application. Par conséquent, nous nous limitons à une détection d'erreurs mais la plus robuste possible, afin que le récepteur soit averti à la moindre erreur de donnée.

Dans la foule des codes détecteurs d'erreur, la méthode du CRC (cyclic redundancy check) constitue un moyen à la fois élégant et efficace. Il est très souvent utilisé dans les différents équipements de transmission et d'instrumentation.

Dans la détection par code cyclique, l'information redondante ou clé CRC est déterminée par une opération mathématique effectuée dans le corps de Galois [45, 82], qui est une arithmétique modulo 2. Elle est appliquée au bloc de données à transmettre puis transmise avec celui-ci, figure 6-2.

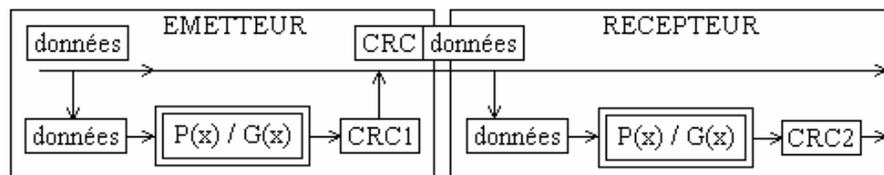


Figure 6-2: Principe de la détection d'erreur par CRC

La méthode de contrôle par CRC, considère le bloc de N bits à transmettre comme un polynôme $P(x)$ de degré $N-1$. Ce polynôme est divisé par un autre, dit polynôme générateur $G(x)$ selon les règles de calcul dans le corps de Galois. Le reste de cette division $R(x)$ constitue le CRC.

Le CRC calculé est transmis à la suite du bloc de données. En réception, le destinataire effectue

la même opération sur le bloc reçu. Le CRC transmis et celui calculé par le récepteur sont comparés, si les valeurs diffèrent une erreur est signalée.

En réalité, la méthode utilisée est un peu différente. En effet, si D est le dividende, d le diviseur et R le reste, la division $(D - R) / d$ donne un reste nul. En arithmétique modulo 2, l'addition et la soustraction sont une même opération, l'opération $(D - R)$ est équivalente à l'opération $(D + R)$. Il leur correspond en algèbre booléenne, l'opération en « ou exclusif ». Dans ces conditions et au récepteur de la figure 6-2, la division par le générateur $G(x)$ de l'ensemble bloc de données et CRC, soit $P(x) + R(x)$, doit rester égale à zéro pour déclarer le bloc reçu exact. L'arithmétique modulo 2 est une arithmétique sans retenue, l'affirmation précédente n'est donc exacte que si le reste est ajouté à une séquence binaire nulle. Pour réaliser cette condition avant d'effectuer la division, on multiplie le polynôme $P(x)$ par x^m où m est le degré du polynôme générateur, ce qui correspond à une translation de m positions. Rappelons que le reste de la division par un diviseur de degré m est de degré $m-1$, il comporte donc m termes. Cette opération a pour effet d'insérer m bits à zéro, pour y ajouter les termes du reste.

Exemple : Si on veut protéger le message « 110111 » par un CRC à l'aide du polynôme générateur $x^2 + x + 1$, on fait correspondre au message le polynôme :

$$x^5 + x^4 + 0x^3 + x^2 + x^1 + x^0.$$

Pour permettre l'addition du CRC au message, on multiplie le polynôme représentatif du message par x^m où m est le degré du polynôme générateur. Ceci revient à ajouter 2 zéros à la donnée. Le dividende devient : $x^7 + x^6 + 0x^5 + x^4 + x^3 + x^2 + 0 + 0$. Lorsque le reste de la division polynomiale est de degré inférieur à celui du diviseur, la division est terminée.

$$\begin{array}{r}
 x^7 + x^6 + 0 + x^4 + x^3 + x^2 + 0 + 0 \quad | \quad x^2 + x + 1 \\
 \underline{x^7 \quad x^6} \\
 x^5 \\
 x^4 \\
 x^3 \\
 x^2 0 0 \\
 x 1 \\
 1 \\
 \text{RESTE} \longrightarrow 1
 \end{array}$$

Le reste comporte deux termes, il est de degré -1 par rapport au polynôme générateur. Le reste ou CRC est donc 11. Le message à transmettre est $P(x) + R(x) = 11011111$. En réception, l'ensemble du message, donnée et CRC, subit la même opération ; si le reste de la division est égal à zéro, on estime que le message n'a pas été affecté par une erreur de transmission.

La division est réalisée physiquement par des opérateurs qui effectue des « ou exclusif » sur des éléments de retard, donc de mémoire [82]. La chaîne des éléments de retard matérialisent au fait, la suite de puissance en x du polynôme générateur. Alors que les opérations en « ou exclusif », appartenant à l'algèbre de Boole, sont l'équivalent des additions-soustractions sans retenue dans le corps de Galois. En d'autres termes, cela repose sur l'identité des tables de vérité d'une porte XOR d'une part, et celles d'une addition ou soustraction modulo 2 de l'autre part.

Le système de génération du CRC de l'exemple précédent, sera donc celui de la figure 6-3.

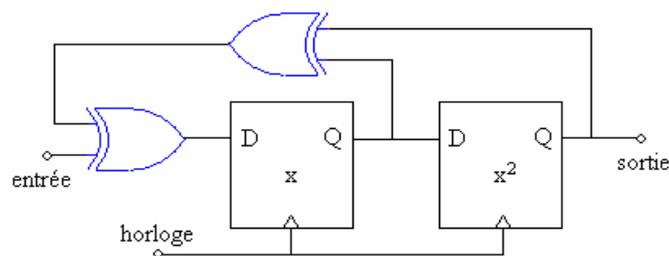


Figure 6-3: Codeur CRC de l'exemple

Le choix d'un polynôme générateur, repose sur une combinaison binaire telle que la probabilité de non détection d'une erreur soit aussi faible que possible, et que le calcul du CRC ne pénalise pas trop la transmission. Le degré du polynôme et d'autant plus élevé que la probabilité d'apparition d'une erreur est grande, c-à-d que la longueur du bloc à protéger est importante.

Les polynômes générateurs utilisés font l'objet de normalisation, et dans notre cas nous retenons l'avis du CCITT n°41 [40, 82] qui recommande l'utilisation du polynôme générateur suivant :

$x^{16} + x^{12} + x^5 + 1$, dont la résistance aux erreurs est excellente. Il détecte :

- toutes les erreurs simples, doubles et en nombre impair.
- toutes les rafales d'erreurs de longueur ≤ 16 bits.
- permet de détecter 99.99 % des erreurs de longueur > 16 bits.

Le codage de canal par CRC sera concrétisé dans la suite, par programmation.

6.1.3. Codage de source

Il existe un certain nombre de codes en ligne dont les principales caractéristiques à prendre en compte lors de leur choix, sont [48] :

- L'occupation spectrale où largeur de bande fréquentielle occupée, est fonction du code utilisé. Par exemple, les spectres du bipolaire ou celui du code Manchester n'ont pas de composantes continues.
- Le nombre de transitions, qui peuvent rendre aisée la restitution d'horloge dans le récepteur.
- L'immunité au bruit, puisque le taux d'erreur binaire est influencé par le code utilisé.

L'occupation spectrale est théoriquement infinie pour tous les types de codes en ligne, mais en pratique, souvent seul importe la prédominance du lobe spectral principal sur les lobes secondaires.

Ainsi, le spectre du simple codage NRZ est plus étroit que ceux des autres codes en ligne (bipolaire, Manchester, Miller), qui sacrifient de la bande pour privilégier les transitions dans le signal émis, en vue de simplifier l'extraction d'horloge. On remarque en effet que dans le code NRZ, les transitions dans le signal émis sont celles de la séquence binaire du message.

Une conséquence de cette propriété est que le signal ne donne pas d'information de rythme (ou d'horloge) lors de la transmission d'une longue séquence de '0' ou de '1'. Dans notre réalisation, ce problème sera résolu par l'utilisation d'une modulation discrète en bande de base, riche en transitions et qui adapte en plus, le signal d'émission à la porteuse HF analogique.

6.1.4. Choix de la modulation discrète

Dans le cadre d'une propagation électromagnétique, il faut assurer une faible occupation spectrale. Pour cela, il s'avère nécessaire d'avoir un spectre limité après modulation en bande de base et avant sa transposition fréquentielle par modulation HF. Un autre point important est que la modulation doit être à enveloppe constante pour maximiser la puissance d'émission [6].

Les transitions du code NRZ génèrent un spectre riche en contenu harmonique qui ne colle pas à une transmission RF. Les modulations discrètes changent les états de la porteuse (sinusoïdale ou impulsionnelle) par des sauts en amplitude, fréquence, phase ou une combinaison des trois. Ces différents sauts sont lissés avant l'opération de modulation. Ceci réduit l'élargissement en bande passante utile.

Il y a principalement trois groupes de modulation discrète. La table 6-1 récapitule les caractéristiques essentielles de ces modulations [75].

ASK	PSK-FSK	MAQ (≡ ASK + PSK)
besoin d'un ampli linéaire pour préserver l'enveloppe.	enveloppe constante (peut être affectée par filtrage).	meilleure efficacité spectrale et de puissance, si symboles à grand nombre d'états.
vulnérable aux fluctuations du canal.	insensible aux fluctuations du canal.	besoin d'un ampli linéaire.
insensible aux dérives de la porteuse.	sensible aux dérives de la porteuse.	sensible aux distorsions du gain et de phase dues au canal.
rarement utilisé.	format le plus courant.	utilisé surtout dans les grands débits.

Table 6-1: Comparaison entre familles de modulation discrète

Le débit de transmission fixé par le cahier des charges à 4800 bit/sec étant relativement faible, la modulation MAQ est écartée. L'ASK, la moins utilisée en pratique, est de même écartée puisqu'elle nécessite en plus, un amplificateur de puissance linéaire. Chose allant à l'encontre du cahier des charges, concernant l'utilisation d'un ampli de puissance à la fois moins encombrant et à haut rendement.

Les modulations que nous allons cibler, seront donc le groupe PSK-FSK à enveloppe constante. Les modulations angulaires discrètes sont nombreuses [24, 115].

Nous rappelons que dans une transmission M-aire [41, 48], un message binaire peut être décomposé en tranches de n éléments binaires successifs. Chaque groupe de ces mêmes éléments, appelé symbole, peut prendre 2^n valeurs. On a donc une relation entre un nombre binaire de n chiffres et un alphabet de $M = 2^n$ états, $\left[\pm 1, \pm \frac{M-3}{M-1}, \pm \frac{M-5}{M-1}, \dots, \pm \frac{1}{M-1} \right]$, qui peuvent

être des niveaux d'amplitude ou un ensemble quelconque de M états discrets autour d'une porteuse. Cet ensemble d'états est appelé constellation. Cependant, les modulations à deux ou plus de bits par symbole (M-aire), requièrent un signal sur bruit plus élevé que les modulations binaires pour un même taux d'erreur de bit (BER) [41, 48].

Dans les systèmes de transmission, les modulations M-aire présentent une bande passante d'autant plus petite que le nombre d'états est plus élevé, mais en contre partie réclament une meilleure linéarité de l'amplificateur de puissance. Ce dernier point constitue un handicap pour notre cahier des charges et par conséquent, nous nous orientons vers les modulations binaires (à 2 états). Une famille entière de modulation catégorisée comme modulation à enveloppe constante et à phase continue, est appelée CPM (continuous phase modulation) et présente de bonnes propriétés spectrales [41, 48]. Elle diminue la consommation en largeur de bande par élimination des discontinuités de phase. Les changements d'état d'une CPM, sont représentés par des changements non abrupts en phase et en fréquence, tandis que l'amplitude de la porteuse demeure constante (modulation angulaire). En définitif, notre choix de modulation discrète se borne uniquement à la famille de modulations CPM 2-aire.

6.1.4.1. Modulation CPM [115] :

Un modulateur CPM typique est le modulateur de fréquence à base d'un VCO, où la déviation de fréquence instantanée est proportionnelle au signal modulant $m(t)$, et la phase est proportionnelle à son intégral.

Soit :
$$e(t) = \text{Re} \{ A_c \exp[j \phi(t)] \} \quad 6.1$$

où, $\phi(t) = \omega_c t + \phi_c + \psi(t)$,

$e(t)$ = signal de modulation angulaire,

A_c = amplitude (non modulée) de la porteuse,

ω_c = fréquence porteuse, et ϕ_c = sa phase initiale,

$\psi(t)$ = phase instantanée, modulée par le signal $m(t)$.
 Pour une CPFSK, ou FSK à phase continue, on a :

$$e(t) = A_c \cos\left[\omega_c t + \phi_c + k \int_{-\infty}^t m(\tau) d\tau\right] \quad 6.2$$

Où, $m(t)$ est le signal modulant et k une constante de proportionnalité.
 Pour un train d'impulsions bipolaires :

$$m(t) = \sum_{n=-\infty}^{+\infty} S_n r(t - nT_b) \quad 6.3$$

Où, $S_n = \pm 1$ pour la polarité des données bipolaires, et $r(t)$ l'impulsion rectangulaire d'un bit d'amplitude $1/(2T_b)$ et de durée T_b . Donc :

$$\psi(t) = k \sum_{n=-\infty}^{+\infty} S_n \int_{-\infty}^t r(\tau - nT_b) d\tau \quad 6.4$$

Pour un intervalle $0 \leq t < T_b$:

$$\psi(t) = k \sum_{n=-\infty}^{+\infty} S_n \frac{t}{2T_b} \quad 6.5$$

Avec ce résultat, l'équation 6.2 devient:

$$e(t) = A_c \sum_{n=-\infty}^{+\infty} \cos\left[\omega_c t + k S_n \frac{t}{2T_b} + \phi_n\right] \quad 6.6$$

Où, $nT_b \leq t < (n+1)T_b$ et ϕ_n est ajustée sur la phase du symbole précédent ϕ_{n-1} .

Pour une CPFSK, le signal modulant étant les données NRZ bipolaires ($S_n = \pm 1$) et en posant :

$$\omega_c + kS_n/2T_b = 2\pi f_n \quad \text{et} \quad f_d = |f_2 - f_1| \quad 6.7$$

avec: f_1 = fréquence pour la logique 0 (cas $S_n = -1$)

f_2 = fréquence pour la logique 1 (cas $S_n = +1$).

$$\text{il vient :} \quad k |S_2 - S_1| / 2T_b = 2\pi f_d \quad \text{soit} \quad k = 2\pi f_d T_b \quad 6.8$$

Le terme $f_d T_b$, est connu sous le nom d'indice de modulation h .

En substituant $k=2\pi h$ dans (6.6), on obtient pour $nT_b \leq t < (n+1)T_b$:

$$e(t) = A_c \sum_{n=-\infty}^{+\infty} \cos\left[2\pi f_c t + \frac{\pi h S_n}{T_b} t + \phi_n\right] \quad 6.9$$

6.1.4.2. Modulation MSK [115] :

Pour MSK, l'indice de modulation est fixé à $h = 0.5$.

L'expression (6.9) devient pour $nT_b \leq t < (n+1)T_b$:

$$e(t) = A_c \sum_{n=-\infty}^{+\infty} \cos\left[2\pi \left(f_c + \frac{S_n}{4T_b}\right) t + \phi_n\right] \quad 6.10$$

Pour que deux signaux soient orthogonaux sur un intervalle T_b , il faut vérifier la condition :

$$\int_0^{T_b} x(t) y^*(t) dt = 0 \quad 6.11$$

Où, $y(t) \neq x(t)$ et $y^*(t)$ est le conjugué complexe de $y(t)$.

Si $x(t)$ et $y(t)$ sont réels, (6.11) devient pour n'importe quel système FSK où $\omega_1=2\pi f_1$ et $\omega_2=2\pi f_2$:

$$\int_0^{T_b} x(t) y(t) dt = \int_0^{T_b} \cos(\omega_1 t) \cos(\omega_2 t) dt = 0 \quad 6.12$$

L'intégration de (6.12) donne :

$$\frac{\sin(\omega_1 - \omega_2) T_b}{2(\omega_1 - \omega_2)} + \frac{\sin(\omega_1 + \omega_2) T_b}{2(\omega_1 + \omega_2)} = 0 \quad 6.13$$

Comme $f_d = 1/(2T_b) = (\omega_1 - \omega_2) / 2\pi$, il vient : $(\omega_1 - \omega_2) = \pi / T_b$ 6.14

La fréquence porteuse est : $\omega_c = (\omega_1 + \omega_2) / 2$ 6.15

En substituant (6.14) et (6.15) dans (6.13), on obtient : $\sin(2\omega_c T_b) = 0$ 6.16

Qui a comme solution, avec K un entier : $\omega_c = \frac{K\pi}{2T_b}$ 6.17

Ainsi, si l'équation (6.17) est satisfaite, f_1 et f_2 seront orthogonales. En fait, l'équation (6.14) représente le « minimum » de séparation entre f_1 et f_2 (ou déplacement minimal de fréquence), d'où le nom de « Minimum Shift Keying ou MSK ».

Les valeurs des fréquences de saut, f_1 et f_2 , peuvent être déduites en fonction de K.

De (6.15) et (6.17), on tire : $f_1 + f_2 = K / (2T_b)$ 6.18

alors que (6.14) donne : $f_1 - f_2 = 1 / (2T_b)$ 6.19

D'où, on tire :

$$\begin{cases} f_1 = \frac{K+1}{4T_b} \\ f_2 = \frac{K-1}{4T_b} \end{cases} \quad 6.20$$

Ainsi, si on prend $f_2 = 1/T_b$, on obtient $K = 5$ et $f_1 = 3f_2 / 2$,
 et, si on prend $f_1 = 1/T_b$, on obtient $K = 3$ et $f_2 = f_1 / 2$.

La figure 6-4 illustre un ensemble de quatre ondes MSK répondant au premier cas.

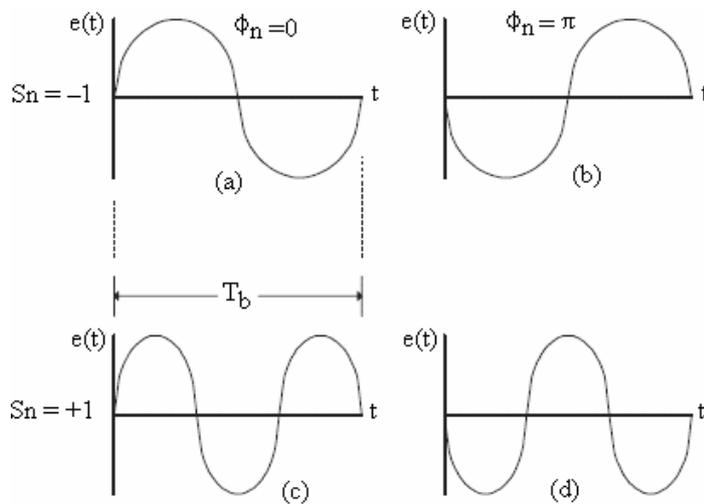


Figure 6-4 : Signaux MSK typiques (cas pour $f_2 = 1/T_b$) avec : (a) logique 0 et $\phi_n=0$
 (b) logique 0 et $\phi_n=\pi$, (c) logique 1 et $\phi_n=0$, (d) logique 1 et $\phi_n=\pi$.

Pour le second cas, la firme MX.com, inc [25], en montre l'exemple sur son modem MSK, le MX469 à 1200/2400/4800 bps. Selon cette firme, ce circuit présente une meilleure résistance au bruit avec une bande passante plus étroite que celle des modems FSK.

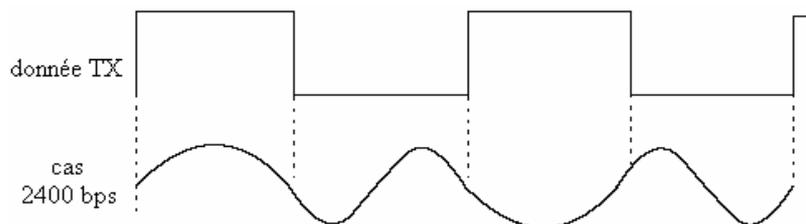


Figure 6-5: Signal MSK du circuit MX469 (cas $f_1=1/T_b=2400$ bps, et $f_2=1200$ bps)

On remarque sur la figure 6-5, que la phase est bien continue et que les changements de fréquence s'opèrent aux passages par zéro du signal de la porteuse.

Vis-à-vis de ses propriétés, MSK est une modulation binaire où chaque symbole est représenté par un seul bit. Elle nécessite donc un signal sur bruit moins élevé que les modulations M-aires pour un même taux d'erreur de bit (BER).

MSK peut recourir à une amplification non linéaire à meilleur rendement et présentant une bande passante non large (pour un spectre encore plus étroit, la GMSK en est une alternative).

Dans les transmissions à grand débit ou dans les milieux perturbés, la transmission synchrone est préférée à l'asynchrone. Une transmission synchrone de données est caractérisée par la présence de l'horloge dans le signal modulé, qui peut être récupérée en réception. MSK est une modulation synchrone. Toutes ces caractéristiques de MSK, font qu'elle est préférée dans le domaine des communications.

6.1.5. Implémentation du modulateur MSK sur GAL

6.1.5.1. Présentation et description du GAL [42, 43, 45, 114] :

Les PLD (programmable logic devices) sont de simples réseaux logiques programmables. Parmi les nombreux PLD existant sur le marché, on trouve le GAL22v10 fabriqué par Lattice sous la marque déposée de « Generic Array Logic » [114]. Il est depuis, réalisé par d'autres fabricants.

Un GAL [42, 43, 45] est muni d'une macrocellule ou OLMC (Output Logic MacroCell) pouvant indifféremment être programmée en logique séquentielle ou en logique combinatoire, à sortie directe ou inversée. A raison de dix macrocellules pour 22 entrées possibles, le GAL22v10 est capable de résoudre à lui seul, des problèmes logiques déjà conséquents (sa densité atteint l'équivalent de 500 portes). Lattice annonce pour son GAL, la possibilité d'une centaine de cycles d'effacement-écriture avec une rétention de données excédant 20 années [114].

Le GAL22v10 est issu de la technologie E²CMOS (CMOS Electrically Erasable) utilisant des grilles flottantes dans ses transistors. Le temps de propagation maximal est de 4 ns (fréquence maximale de 250 MHz). Avec des temps d'effacement inférieurs à 100 ms, le GAL peut être reprogrammé ou reconfiguré rapidement.

Le schéma de la figure 6-6 montre l'organisation de la matrice programmable commune à toutes les macrocellules.

Tous les signaux sont reliés par un opérateur de produit, AND. Le GAL présente un nombre variable de termes de produit par OLMC, reliés par un opérateur OR.

Puisque toutes les entrées et réinjections (feedback) de l'OLMC, aussi bien les non-inversées que les inversées, sont reliées aux lignes de la matrice AND, on peut réaliser dans un GAL des fonctions booléennes de forme canonique disjonctive.

En plus, chaque OLMC possède un terme produit supplémentaire dédié au contrôle de validation de sortie (OE : Output Enable).

La polarité de la sortie de chaque OLMC peut être individuellement programmée en direct ou inversée, en mode registre ou mode combinatoire. Ceci permet à chaque sortie d'être configurée individuellement comme active haute ou active basse. Le GAL possède un terme produit pour un Reset Asynchrone (AR) et un terme produit pour un Preset Synchrone (SP).

Ces deux termes de produit sont communs à toutes les OLMCs en mode registre.

Le GAL présente un auto-reset à la mise sous tension. Chaque macrocellule a deux modes fonctionnels : registre et E/S combinatoire. Les modes et la polarité de sortie d'une macrocellule sont fixés par deux fusibles (bits S_0 et S_1), figure 6-6 et table 6-2.

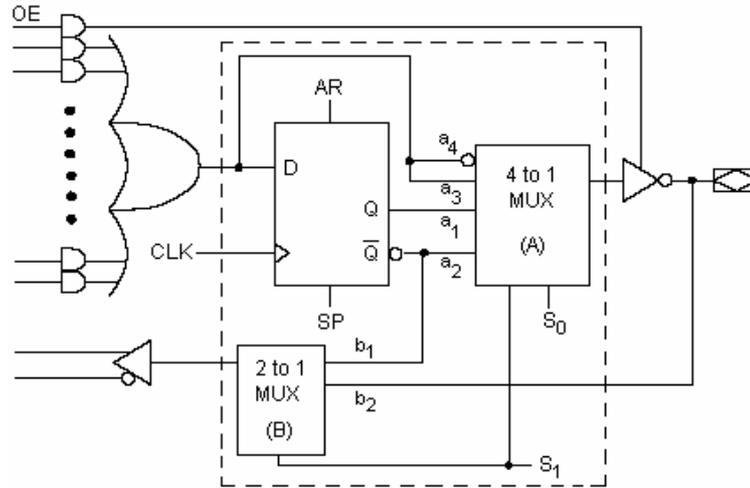


Figure 6-6: OLMC du GAL22V10

S ₁	S ₀	MUX	
		A	B
0	0	a ₁	B ₁
0	1	a ₂	
1	0	a ₃	B ₂
1	1	a ₄	

Table 6-2: Entrées de sélection des deux multiplexeurs

6.1.5.2. Programmation du GAL :

La programmation du GAL ne prend que quelques secondes. L'affectation des broches du GAL est différente lors de son utilisation normale et pendant sa programmation. Les fusibles dans un GAL sont numérotés selon la norme JEDEC, dans laquelle, les fusibles à programmer sont indiqués par L suivi de l'adresse ou numéro de fusible. A la fin du fichier JEDEC, une somme de vérification à quatre chiffres contrôle son CRC.

Un compilateur GAL permet de traduire un fichier source écrit dans un langage relativement évolué, en un fichier JEDEC qui sera exploité par un programmeur de PLD. Un bon compilateur de GAL décharge l'utilisateur de la partie la plus fastidieuse de la programmation du GAL : la mise au point du plan des fusibles (élaboration du fichier JEDEC).

Parmi les compilateurs, nous avons utilisé PROLOGIC de la firme INLAB [98] (offert par Texas Instruments et disponible sur CD d'accompagnement avec son manuel). Ce logiciel réunit deux outils distincts : un compilateur et un simulateur logique. L'emploi de ce dernier demeure parfaitement facultatif. PROLOGIC attribue aux broches utilisées ainsi qu'aux signaux, des noms logiques qu'il faudra utiliser pour préparer un fichier source à partir d'équations booléennes, de tables de vérité, ou de diagrammes d'état. Après compilation, et en plus du fichier JEDEC, un fichier contenant une documentation technique détaillée (plan des fusibles, équations simplifiées), est fourni.

6.1.5.3. Application au modulateur MSK :

Nous nous inspirons des diagrammes du modem MX469, établis en figure 6-5. La fréquence des bits d'entrée étant fixée à 4800 bps, nous choisissons donc les déplacements en fréquence : $f_0 = 4800$ Hz et $f_1 = 2400$ Hz, telle que le bit 0 lui sera attribué une période complète de f_0 et le bit 1 une demie période de f_1 . D'après les expressions (6.17) et (6.20), $K=3$ et la fréquence porteuse est de $F_c = 3600$ Hz ($f_0 = F_c + 1200$ Hz et $f_1 = F_c - 1200$ Hz).

Le signal modulé sera donc construit par imbrication de périodes de f_0 et de demie périodes de f_1 . La fréquence f_1 sera générée de façon synchrone à partir de f_0 , à la sortie Q_1 d'une bascule D montée en diviseur de fréquence. Chaque intervalle de bit T_b est représenté, soit par un état d'une période f_0 ou son inverse $\overline{f_0}$, soit par un état d'une demie période f_1 ou son inverse $\overline{f_1}$. Ces quatre états possibles du signal modulé S_m , seront disponibles à la sortie d'un multiplexeur à quatre entrées. La sélection de ces états se fait selon la table de vérité 6-3. Les différents états sont désignés par S_i , et les sélecteurs I_j sont déterminés ultérieurement.

état	I_1	I_0	$S_m(\text{MSK})$
S_0	0	0	f_0
S_1	0	1	$\overline{f_0}$
S_2	1	0	f_1
S_3	1	1	$\overline{f_1}$

Table 6-3: Table de vérité du multiplexeur de sortie MSK

Afin d'assurer la continuité de phase du signal modulé S_m , aux changements de bit d'entrée e , il faut savoir coller une période de f_0 (ou $\overline{f_0}$) à la suite d'une demie période de f_1 (ou $\overline{f_1}$) et inversement : le nouvel état est toujours l'inverse de l'état précédent. Dans ce but, les sélecteurs $I_1 I_0$ seront tamponnés aux sorties $Q_2 Q_3$ de deux bascules D.

Le diagramme d'état de la figure 6-7, montre comment lier adroitement les différents états des signaux f_0 et f_1 en fonction de l'entrée binaire e , l'état actuel de Q_1 , et les états passés $Q_2 Q_3$ des sélecteurs $I_1 I_0$. Ces derniers sont désignés par les états encadrés S_i ($i = 0$ à 3), table 6-3. On peut rester sur un état S_i , comme on peut passer à un état adjacent, selon les valeurs de e , Q_1 , et de l'état S_i lui-même. Dans le fichier source réalisé, est programmé le diagramme d'état donnant $I_1 I_0$, la table de vérité délivrant le signal S_m ou MSK, et enfin le diviseur de fréquence générant f_1 . La table de vérité 6-4 aide à mieux comprendre le diagramme d'état de la figure 6-7, qui détermine essentiellement les sélecteurs ou entrée $I_1 I_0$ du multiplexeur MSK.

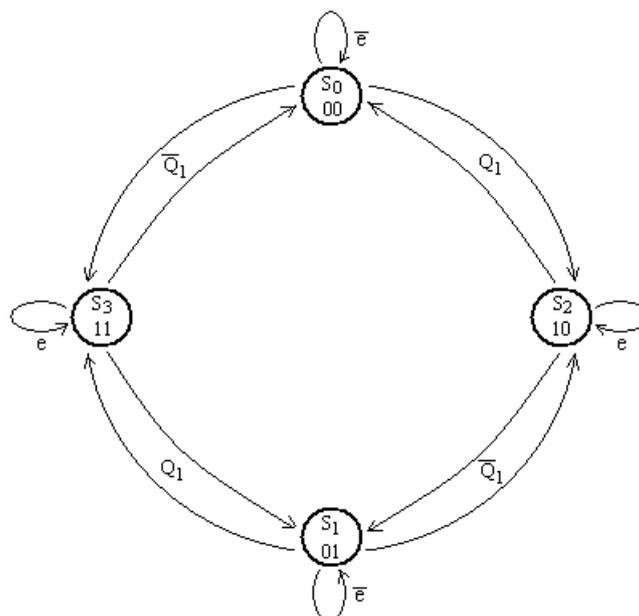


Figure 6-7: Diagramme d'état des sélecteurs de sauts de fréquence

E	Q_1	Q_2	Q_3	I_1	I_0	S_m
---	-------	-------	-------	-------	-------	-------

(entrée)	(actuel)	(I ₁ passé)	(I ₀ passé)	(actuel)	(actuel)	(MSK)
0	x	0	0	0	0	f ₀
0	x	0	1	0	1	f ₀
0	0	1	0	0	1	f ₀
0	1	1	0	0	0	f ₀
0	0	1	1	0	0	f ₀
0	1	1	1	0	1	f ₀
1	0	0	0	1	1	f ₁
1	1	0	0	1	0	f ₁
1	0	0	1	1	0	f ₁
1	1	0	1	1	1	f ₁
1	x	1	0	1	0	f ₁
1	x	1	1	1	1	f ₁

Table 6-4: Table de vérité du modulateur MSK

On trouve comme solution : $I_1 = e$ et $I_0 = Q_3 \oplus [(Q_2 \oplus e) \oplus \overline{Q_1}]$.

6.1.5.4. Application au générateur pseudo-aléatoire :

Parmi les applications du générateur pseudo-aléatoire, on cite la génération de bruit blanc discret, qui ne contient pas toutes les fréquences mais seulement des fréquences à intervalles réguliers. C'est un pseudo bruit périodique. L'utilisation du générateur pseudo-aléatoire est mentionnée dans le cahier des charges comme moyen d'évaluation des caractéristiques de l'appareil de transmission, une fois réalisé.

Nous rappelons que la théorie de génération de données (binaires) aléatoires, rentre dans le cadre des opérations arithmétiques faites dans le corps de Galois GF(2), voir paragraphe 6-1. Une méthode très répandue consiste à utiliser des polynômes premiers, modulo 2. Un polynôme premier est irréductible et ne peut être factorisé, des algorithmes existent pour tester la primitivité des polynômes dans GF(2). Il y a aussi des tables qui contiennent un polynôme premier pour chaque degré jusqu'au degré 168 [45, 105]. L'ordre n du polynôme définit le nombre de séquences différentes qui sont générées avant que celles-ci se répètent. Ce nombre est égal à $2^n - 1$. On peut concevoir un générateur pseudo aléatoire par programmation ou par réalisation.

Pour réaliser un circuit qui opère suivant un polynôme donné, il suffit d'avoir un registre à décalage disposant d'autant de pas que l'ordre du polynôme désiré, et un nombre de portes ou-exclusif égal au nombre d'additions dans le polynôme moins un. Un décalage correspond à une multiplication (ou division) par deux, et un ou-exclusif à une addition modulo 2. Pour fonctionner correctement après démarrage ou un reset, il suffit d'initialiser à « 1 » une seule bascule du registre (généralement la première), une valeur initiale égale à zéro de tout le registre se perpétuant indéfiniment. Ayant consommé cinq OLMC du GAL dans l'application du modulateur MSK, il nous reste les cinq autres pour réaliser notre générateur pseudo-aléatoire. Le polynôme premier doit donc être de degré 5, et présente alors une longueur de 31 séquences. Le polynôme premier retenu est $x^5 + x^2 + 1$, avec une période de séquence de 6.46 ms sous une horloge de 4800Hz. La figure 6-8 illustre un tel générateur, conçu dans le GAL par simple programmation de son polynôme sous forme d'équation combinatoire. La sortie s_g a été prise au lieu de e_1 pour éviter de gaspiller une bascule (une sortie combinatoire dans une OLMC du GAL, gaspille la bascule correspondante).

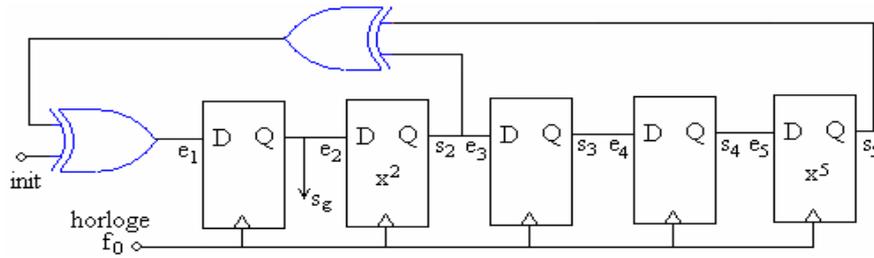


Figure 6-8: Générateur pseudo-aléatoire implémenté dans le GAL22V10

6.1.6. Filtre de transmission

Un filtre de transmission est un filtre en bande de base, c'est-à-dire filtre le signal modulant MSK. Les signaux NRZ issus du GAL (sorties MSK et générateur pseudo-aléatoire) obligent à filtrer leurs transitions abruptes avant de les appliquer au modulateur HF.

Les signaux sont à bande illimitée si l'impulsion élémentaire est limitée à la durée T_s (ou T_b dans le cas du symbole binaire). Cette propriété découle du fait qu'un signal ne peut être à la fois à bande limitée et à durée limitée. Or, dans les applications pratiques, la bande de fréquence est une ressource rare qu'il faut utiliser de la manière la plus efficace possible. En utilisant un signal à bande limitée, les impulsions élémentaires associées aux symboles composant ce signal sont de durée infinie.

En général, pour transmettre à la cadence $1/T_s$ des symboles M-aires sans qu'il y ait une perte en performances par rapport à la transmission à bande illimitée (impulsions rectangulaires de durée T_s), il faut déterminer des impulsions élémentaires avec passages par zéro espacées de T_s secondes. Cette propriété est précisément le critère de Nyquist [48] pour une transmission sans interférences entre symboles (IES). Elle assure que chaque échantillon du signal reçu provient d'un seul symbole. La bande minimale nécessaire pour transmettre sans IES à une cadence $1/T_s$ est de $1/2T_s$, que l'on appelle la bande de Nyquist.

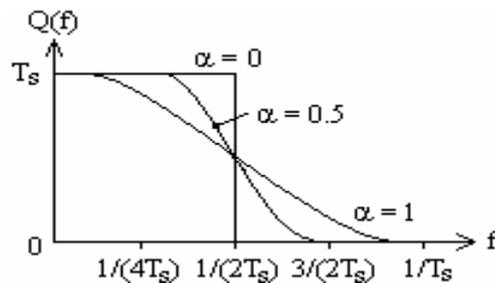


Figure 6-9: Fonction de transfert du filtrage de Nyquist

La figure 6-9 montre la fonction de transfert d'un filtre passe-bas idéal indiqué par $\alpha = 0$. La réponse impulsionnelle de ce filtre est un sinus cardinal passant par zéro en tout point de forme nT_s sauf pour $n=0$. Ce filtre n'étant pas réalisable, il existe d'autres filtres dont la bande est inférieure à $1/T_s$ et vérifiant le critère de Nyquist pour transmettre des symboles à la cadence $1/T_s$. Par exemple, le filtre en cosinus surélevé [48] est paramétré par α ; pour $\alpha=0$, on obtient le filtre passe-bas idéal de bande $1/2T_s$ et pour $\alpha=1$, on a un filtre occupant une bande $1/T_s$.

Le facteur α qui définit la bande supplémentaire par rapport à la bande de Nyquist est appelé facteur d'arrondi (ou facteur de roll-off). Cependant, il faut signaler qu'il existe une infinité de fonctions de transfert qui s'annulent pour $|f| > 1/T_s$ et vérifient le critère de Nyquist pour annuler l'IES.

Dans notre cas, nous avons opté pour un filtre passe-bas actif du 2^{ème} ordre, figure 6-10.

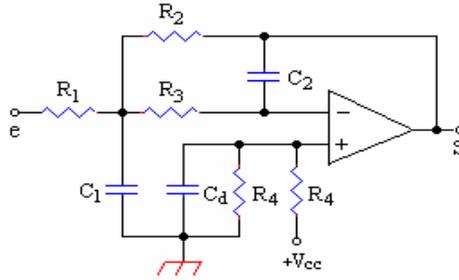


Figure 6-10: Filtre passe-bas de transmission

La fonction de transfert d'un tel filtre est :

$$G(f) = \frac{G_0}{1 + j \frac{1}{Q} \frac{f}{f_n} - \left(\frac{f}{f_n}\right)^2} \quad 6.21$$

où,

$$\begin{cases} G_0 = -R_2/R_1 \\ Q = \sqrt{\frac{C_1}{C_2}} \frac{R_1 \sqrt{R_2 R_3}}{R_1 R_2 + R_2 R_3 + R_3 R_1} = \frac{1}{2\varepsilon} \\ f_n = \frac{1}{2\pi \sqrt{R_2 R_3 C_1 C_2}} \end{cases} \quad 6.22$$

Comme la modulation MSK présente un indice de modulation $h=0.5$, et que le rythme des bits est à $F_b = 4800$ bps, la bande passante est estimée à : $B = 2(h + 1).F_b = 14.4$ kHz (la bande de Nyquist étant de 2.4 kHz).

Le coefficient Q dans l'expression (6.21) du gain complexe du filtre, n'est autre que $1/2\varepsilon$ où ε est le coefficient d'amortissement du filtre. En fixant ce dernier à $\varepsilon = 1/\sqrt{2}$, on obtient $Q = 1/\sqrt{2}$.

En fixant de même, le gain statique $G_0 = -1$, la fréquence de coupure à $f_n = 15$ kHz, et la polarisation de l'ampli-op à 12V ; le calcul des éléments du circuit donne comme valeurs normalisées : $R_1 = R_2 = 12.7$ K Ω , $R_3 = 9.3$ K Ω , $R_4 = 33$ K Ω , $C_1 = C_2 = 1$ nF, $C_d = 22$ nF.

La réponse impulsionnelle du filtre de transmission, a été calculée par transformation inverse de

Laplace, et elle s'écrit :

$$g(t) = \frac{\omega_n \exp(-\varepsilon \omega_n t) \sin(\omega_n \sqrt{1 - \varepsilon^2} t)}{\sqrt{1 - \varepsilon^2}} \quad 6.23$$

En posant : $\omega_0 = \frac{\omega_n}{\sqrt{1 - \varepsilon^2}}$, la réponse impulsionnelle réduite g / ω_0 a été tracée avec les valeurs

numériques des différentes grandeurs établies précédemment, et se trouve en figure 6-11. La réponse fréquentielle en amplitude est représentée en figure 6-12. Un programme réalisé en matlab, nous a permis de concrétiser ces tracés.

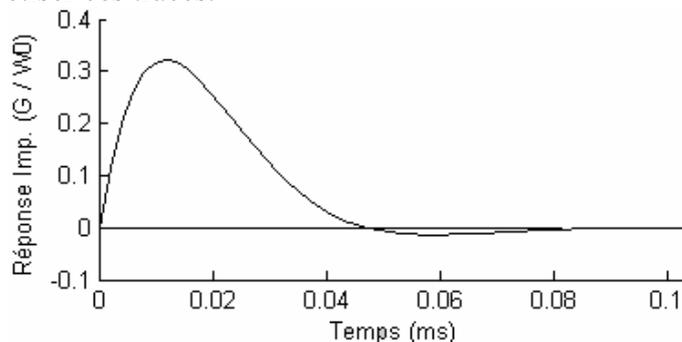


Figure 6.11: Réponse impulsionnelle du filtre de transmission

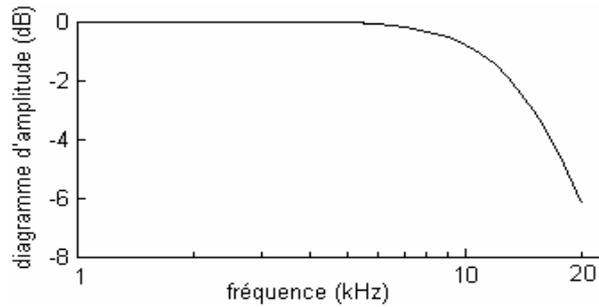


Figure 6.12: Réponse fréquentielle en amplitude du filtre de transmission

En respectant la bande passante du signal MSK, par filtrage passe-bas à 15kHz, on dépasse la bande (théorique) de Nyquist de 6 fois et la bande (pratique) minimale d'environ 3 fois, figure 6-9. Cette largeur de bande nous a permis d'avoir une réponse impulsionnelle du filtre, s'étendant sur une même durée que présente un état minimal du signal MSK d'entrée, soit $1/2F_b=0.1\text{ms}$. Ceci, a pour effet d'annuler toute distorsion intersymbole, comme c'est confirmé par la pratique.

6.1.7. Modulation HF à PLL

Une modulation FM est la plus propice pour succéder à la modulation MSK filtrée. Elles sont toutes les deux, des modulations angulaires qui jouissent d'une bonne immunité contre le bruit tout en tolérant une amplification non linéaire en classe C [73]. L'inconvénient d'étalement du spectre, dû à ce type de modulation, nous pousse à utiliser une modulation FM à bande étroite. Il est naturel de se tourner vers une PLL, lorsque celle-ci est asservie à une fréquence porteuse stable et précise. Une modulation de fréquence par le biais du VCO seul de la PLL, peut souffrir d'instabilité à cause des éléments constituant le VCO et dont les valeurs sont sujettes à des dérives, notamment en fonction de la température et du temps. Mais l'utilisation de toutes les ressources de la PLL, offre l'avantage d'une excellente stabilité.

La PLL [6, 36] se compose de trois éléments, l'oscillateur contrôlé en tension VCO, le comparateur de phase, et le filtre de boucle, figure 6-13.

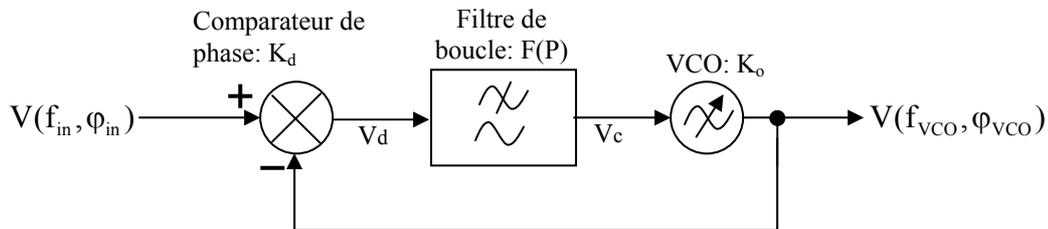


Figure 6-13: Schéma fonctionnel d'une PLL

Le VCO : est l'élément principal de la boucle puisque l'asservissement va porter sur ses paramètres, fréquence et phase. La fréquence du VCO varie autour d'une fréquence centrale f_0 sous l'action d'une commande $V_C(t)$:

$$2\pi \cdot f_{vco}(t) = 2\pi \cdot f_0 + K_O V_C(t) \quad 6.24$$

Avec K_O coefficient de proportionnalité à $V_C(t)$, appelé sensibilité du VCO.

En considérant des grandeurs variables, et comme la fréquence est la dérivée de la phase instantanée, on peut écrire :

$$\frac{d\phi_{vco}(t)}{dt} = K_O V_C(t) \Rightarrow \phi_{vco}(P) = \frac{K_O V_C(P)}{P} \quad 6.25$$

Le comparateur de phase : encore appelé détecteur de phase, effectue la différence entre les phases des signaux des deux entrées et délivre un signal de sortie V_D proportionnel à un facteur K_D , appelé gain ou sensibilité du comparateur de phase :

$$V_D(P) = K_D [\varphi_{in}(P) - \varphi_{vco}(P)] \quad 6.26$$

Le filtre de boucle : élément important de la boucle, il permet par le choix de ses paramètres de réaliser un système stable, rapide, et précis. La tension de sortie du filtre de boucle est donnée par la relation :

$$V_C(P) = F(P) V_D(P) \quad 6.27$$

En plus de la stabilisation du VCO par la boucle à verrouillage de phase, il faut assurer de même sa modulation en fréquence. Selon le schéma fonctionnel de la figure 6-14, il suffit d'ajouter la tension de modulation $m(t)$ au signal de sortie du filtre de boucle.

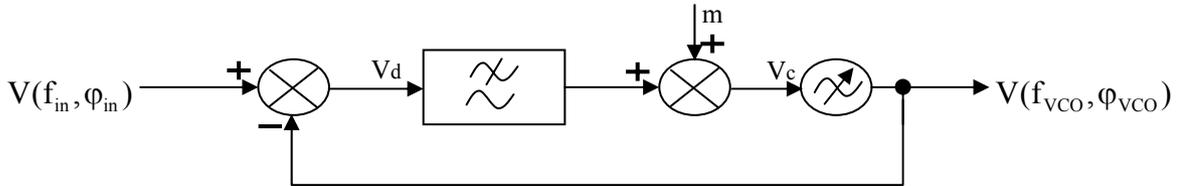


Figure 6-14: VCO modulé en fréquence par m et stabilisé par PLL

Les trois principales équations régissant le fonctionnement du système de la figure 6-14, sont les expressions 6.25, 6.26, et 6.27. Cette dernière expression devient dans ce cas :

$$V_C(P) = m(P) + F(P) V_D(P) \quad 6.28$$

En résolvant ce système de trois équations, on peut écrire:

$$\varphi_{vco}(P) = \frac{K_o}{P + K_o K_D F(P)} \cdot m(P) + \frac{K_o K_D F(P)}{P + K_o K_D F(P)} \cdot \varphi_{in}(P) \quad 6.29$$

Dans le 1^{er} terme, le facteur de $m(P)$ est dû à la modulation et dans le 2^{ème} terme, le facteur de $\varphi_{in}(P)$ est représentatif de la stabilisation de la fréquence centrale du VCO (fréquence porteuse).

- cas de boucle sans modulation, $m(P)=0$:

$$\text{il reste de (6.29): } \varphi_{vco}(P) = \frac{K_o K_D F(P)}{P + K_o K_D F(P)} \cdot \varphi_{in}(P) = T_F(P) \cdot \varphi_{in}(P) \quad 6.30$$

La transmittance en boucle fermée de la PLL vaut donc:

$$T_F(P) = \frac{\varphi_{vco}}{\varphi_{in}} = \frac{f_{vco}}{f_{in}} = \frac{K_o K_D F(P)}{P + K_o K_D F(P)} \quad 6.31$$

La fonction de transfert de l'erreur de phase est donnée par la relation :

$$\frac{\varphi_{in} - \varphi_{vco}}{\varphi_{in}} = \frac{\varepsilon(P)}{\varphi_{in}} = 1 - T_F(P) = \frac{P}{P + K_o K_D F(P)} \quad 6.32$$

La transmittance en boucle fermée peut s'exprimer en fonction de la transmittance en boucle ouverte par :

$$T_F(P) = \frac{T_o(P)}{1 + T_o(P)} \quad 6.33$$

$$\text{D'où: } T_o(P) = \frac{\varphi_{vco}}{\varepsilon(P)} = \frac{T_F(P)}{1 - T_F(P)} = \frac{K_o K_D F(P)}{P} \quad 6.34$$

- Cas de boucle avec modulation, $m(P) \neq 0$:

L'équation 6.29 peut être exprimée en fonction de $T_F(P)$:

$$\varphi_{vco}(P) = \frac{K_o}{P} [1 - T_F(P)] \cdot m(P) + T_F(P) \cdot \varphi_{in}(P) = \frac{K_o \varepsilon(P)}{P} \cdot m(P) + T_F(P) \cdot \varphi_{in}(P) \quad 6.35$$

En FM, $V(f_{in}, \varphi_{in})$ joue le rôle de la porteuse avec f_{in} et φ_{in} des constantes, donc stable, issue d'un oscillateur à quartz par exemple.

Ainsi pour une absence de variation de la phase d'entrée, $\varphi_{in}(P) = 0$, (6.35) devient:

$$\varphi_{VCO}(P) = \frac{K_O \varepsilon(P)}{P} \cdot m(P) \quad 6.36$$

La fréquence étant la dérivée de la phase, on obtient:

$$2\pi \cdot f_{VCO}(P) = K_O \varepsilon(P) \cdot m(P) \quad 6.37$$

On obtient ainsi, une modulation FM où la fréquence $f_{VCO}(P)$ varie autour de sa fréquence centrale $f_0 = f_{in}$, au rythme du message $m(P)$. Mais pour cela, il faut que la PLL soit assez lente pour transmettre l'information $m(P)$ avant l'annulation de l'erreur $\varepsilon(P)$. Dans l'équation 6.32, l'erreur est fonction de $1-T_F(P)$ où $T_F(P)$ est la transmittance en boucle fermée de la PLL assurant un filtrage passe-bas. En conséquence, le terme $1-T_F(P)$ représente un filtrage passe-haut. Ceci signifie que le signal $m(t)$ ne doit pas contenir des composantes à très basses fréquences, ce domaine étant réservé à l'asservissement de fréquence et de phase du VCO. La fréquence de coupure basse de la PLL, doit donc être inférieure à la fréquence la plus basse du signal à transmettre.

La fréquence centrale de la PLL dépend de celle d'entrée, $f_0 = f_{in}$. Si la fréquence d'entrée f_{in} est issue d'un quartz, la fréquence de sortie f_0 conserve la même précision que celle d'entrée.

Dans notre cas, en choisissant une fréquence porteuse de 2 MHz, située au milieu de la gamme [1, 3 MHz] fixée par le cahier des charges, le modulateur FM sera à base d'un circuit courant, la PLL CD4046. Ce circuit est une PLL semi-numérique qui contient deux comparateurs de phase, un VCO, et un suiveur. Son VCO délivre un signal carré modulable en fréquence, qui atteint une fréquence maximale d'environ 1.4 MHz.

On sait qu'un signal carré ne possède que les harmoniques impaires. De ce fait, en considérant la troisième harmonique égale à la porteuse 2 MHz, la fréquence centrale de la PLL doit être fixée à $f_0 = 2 / 3 = 0.66$ MHz.

Si V_{DD} désigne la tension d'alimentation du circuit, ses différentes sensibilités sont :

- $K_D = V_{DD} / \pi$ pour comparateur de phase 1, du type porte XOR,
- $K_D = V_{DD} / 4\pi$ pour comparateur de phase 2, du type circuit séquentiel à trois états [5, 6, 36],
- $K_O = 2\pi \Delta f_{VCO} / (V_{DD} - 2v)$, Où $\Delta f_{VCO} = f_{max} - f_{min}$ est la plage fréquentielle linéaire du VCO.

La figure 6-15 montre la PLL montée en modulateur FM, et utilisant le comparateur de phase 2.

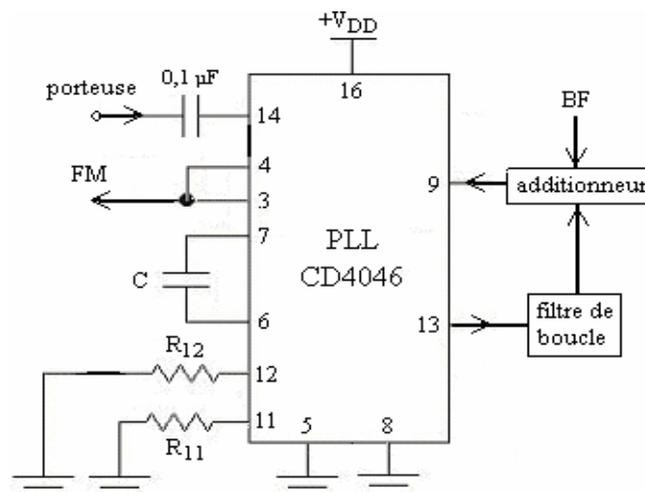


Figure 6-15: Le circuit CD4046 monté en modulateur FM

La résistance R_{12} fixe la fréquence minimale du VCO : Si R_{12} diminue, f_{min} augmente ($f_{min} = 0$ si R_{12} enlevée). La résistance R_{11} fixe la fréquence centrale du VCO sous $V_{DD} / 2$ appliquée au pin 9, et en fonction de la valeur du condensateur C .

Le circuit PLL délivre un signal carré qui subit, par la suite, une intégration par effet de la capacité parasite d'entrée de l'amplificateur de puissance. Faute de quoi, le nouveau signal présente un deuxième harmonique non nul. Ainsi, au lieu de fixer la fréquence centrale de la PLL à $f_0 = 0.66$ MHz comme prévu, nous sommes amenés à la fixer aux environs de 1 MHz pour sélectionner son deuxième harmonique sur 2 MHz comme porteuse.

Avec les composants satellites du circuit de la figure 6-15, fixés à : $R_{11} = 10$ k Ω , $R_{12} = 22$ k Ω , $C = 100$ pF, et sous une alimentation $V_{DD}=12$ V, nous avons relevé les mesures suivantes en fonction de la tension continue appliquée à l'entrée du VCO (pin 9) :

$$\begin{aligned} f_{\min} &= 0.66 \text{ MHz} && \text{à environ 1 volt,} \\ f_0 &= 1 \text{ MHz} && \text{à environ 6.2 volts,} \\ f_{\max} &= 1.3 \text{ MHz} && \text{à environ 11 volts.} \end{aligned}$$

Les sensibilités de la PLL deviennent :

$$K_D = V_{DD} / 4\pi \approx 1 \text{ V/rd} \text{ pour le performant comparateur de phase n}^\circ 2,$$

$$K_O = 2\pi (f_{\max} - f_{\min}) / (11 - 1\text{V}) = 4.10^5 \text{ rd / Vs.}$$

$$\text{Ce qui donne comme produit de sensibilité : } K_V = K_O K_D = 4.10^5 \text{ s}^{-1}.$$

6.1.7.1. Filtre de boucle :

Il est préconisé d'utiliser une PLL du troisième ordre [36], dans le cas d'utilisation d'un comparateur de phase à trois états, et aussi pour pouvoir optimiser facilement à la fois, la bande passante, le coefficient d'amortissement, et le gain du système.

Le filtre d'une boucle PLL étant généralement responsable des performances de toute la boucle d'asservissement, il sera dans ce cas, du deuxième ordre selon l'expression 6.31. Il doit aussi être intégrateur pour annuler les erreurs statiques en fréquence de la boucle.

Le filtre finalement retenu et répondant aux exigences précédentes, est donnée en figure 6-16.

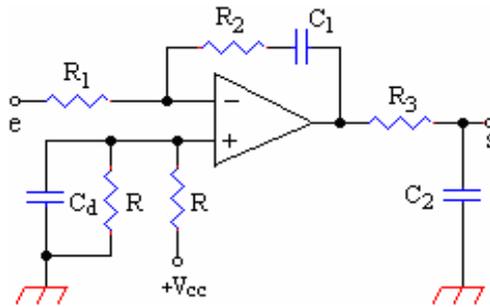


Figure 6.16: Filtre du deuxième ordre de la boucle PLL

Sa fonction de transfert est de la forme :

$$F(P) = -\frac{1 + \tau_2 P}{\tau_1 P (1 + \tau_3 P)} \quad 6.38$$

où, $\tau_1 = R_1 C_1$, $\tau_2 = R_2 C_1$, et $\tau_3 = R_3 C_2$.

La présence du terme $1 + \tau_2 P$ assure une correction par avance de phase.

En posant $K_v = K_O K_D$, les transmittances en boucle fermée 6.31 et ouverte 6.34, s'écrivent respectivement :

$$T_F(P) = \frac{K_v \tau_2}{\tau_1 \tau_3} \frac{P + \tau_2^{-1}}{P^3 + \tau_3^{-1} P^2 + K_v \tau_2 \tau_1^{-1} \tau_3^{-1} P + K_v \tau_1^{-1} \tau_3^{-1}} \quad 6.39$$

$$T_O(P) = \frac{K_v \tau_2}{\tau_1 \tau_3} \frac{P + \tau_2^{-1}}{P^2 (P + \tau_3^{-1})} \quad 6.40$$

La figure 6-17 donne les diagrammes de Bode de $T_O(P)$ pour une meilleure marge de stabilité.

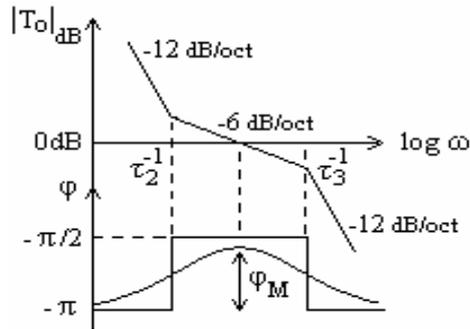


Figure 6-17: Diagrammes de Bode de $T_O(j\omega)$

De l'équation caractéristique de $T_F(P)$, on détermine la stabilité par le critère de Routh: $\tau_2 > \tau_3$. L'équation caractéristique se reporte à une équation normalisée du troisième ordre, qui s'écrit [36]:

$$F(P) = P^3 + \omega_n(1 + 2\varepsilon)P^2 + \omega_n^2(1 + 2\varepsilon)P + \omega_n^3 \quad 6.41$$

tel que, ε est le coefficient d'amortissement et ω_n la pulsation naturelle (ici, pulsation de coupure de la PLL). Par identité, on trouve :

$$\tau_1 = \frac{K_V(1 + 2\varepsilon)}{\omega_n^2}, \quad \tau_2 = \frac{1 + 2\varepsilon}{\omega_n} \quad \text{et} \quad \tau_3 = \frac{1}{\omega_n(1 + 2\varepsilon)} \quad 6.42$$

Le choix optimal des constantes de temps, s'obtient lorsque la marge de phase est $\varphi_M = 45^\circ$, correspondant à $\varepsilon = 1/\sqrt{2}$. D'autre part, la fréquence de coupure de la PLL devant être inférieure à la fréquence minimale du signal modulant, $\omega_n < \omega_{\min}$ (MSK), nous la fixons à $\omega_n = 2\pi \cdot 500\text{Hz}$. On obtient donc les constantes de temps : $\tau_1 = 0.1 \text{ s}$, $\tau_2 = 0.805 \text{ ms}$, et $\tau_3 = 0.138 \text{ ms}$.

Soit les valeurs des composants : $C_1 = 0.33 \mu\text{F}$, $C_2 = 22 \text{ nF}$, $R_1 = 300 \text{ k}\Omega$, $R_2 = 2.2 \text{ k}\Omega$, $R_3 = 6.2 \text{ k}\Omega$, et les deux $R = 33 \text{ k}\Omega$ découplées par $1 \mu\text{F}$.

La valeur prise pour le coefficient d'amortissement assure, pour le système approximé à celui du deuxième ordre [44], un temps de réponse de la PLL de $t_r = 3 / \omega_n = 1 \text{ ms}$.

Le système obtenu est aussi stable et astatique (erreur statique nulle pour des sauts de fréquence ou de phase à son entrée, à cause de la double intégration à l'origine de $T_O(P)$).

6.1.7.2. Additionneur du signal modulant :

Le circuit de la figure 6.18 est un additionneur des deux signaux de filtres passe-bas. A savoir, le filtre de transmission à l'entrée e_1 et le filtre de boucle à l'entrée e_2 .

$$\text{Le signal de sortie s'écrit donc : } s = -R_3(e_1/R_1 + e_2/R_2) \quad 6.43$$

Le signe moins de l'expression 6.43 compense ceux issus des deux filtres inverseurs d'entrée e_1 et e_2 .

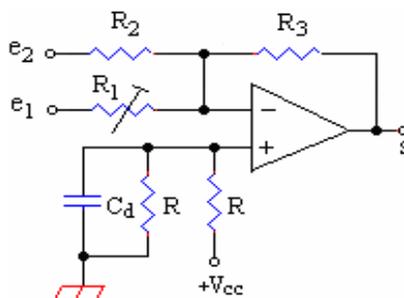


Figure 6.18: Circuit additionneur des signaux de filtres

Les résistances de pondération R_1 , R_2 , et R_3 , se calculent de la façon suivante :

L'ampli de puissance faisant suite à la PLL, doit être sélectif pour s'accommoder avec une FM à bande étroite. Dans le calcul optimal des circuits accordés d'alimentation de l'implant, mené au chapitre 3, la bande passante du système était estimée à environ 76 kHz.

On arrondit donc la bande FM transmise à la valeur $B_{FM} = 100$ kHz.

Sachant que le signal modulant n'est autre que le signal MSK filtré, la bande fréquentielle occupée par ce dernier s'étend à : $B_{MSK} = 14.4$ kHz, voir paragraphe 6-1-6.

Ainsi d'après la bande de Carson, l'indice de modulation de la FM est : $\delta_{FM} = (0.5 B_{FM} / B_{MSK}) - 1$ soit $\delta_{FM} = 2.5$.

L'excursion en fréquence devient : $\Delta F = \delta_{FM} F_{max} = \delta_{FM} B_{MSK} = 36$ kHz. Ce qui donne comme fréquences limites de modulation : $f_{max} = f_{porteuse} + \Delta F = 2.036$ MHz, et $f_{min} = f_{porteuse} - \Delta F = 1.964$ MHz.

Pour cela, la tension crête à crête du signal MSK ne doit pas dépasser :

$$V_{c\text{-}\grave{a}\text{-}c} = 2\pi \Delta F / K_O = 2\pi 36.10^3 / 4.10^5 = 0.56 \text{ volt.}$$

Le signal MSK issu du GAL étant de type TTL, et comme son filtre présente un gain statique de $G_0 = -1$, on obtient donc l'expression de la tension crête à crête du signal MSK à la sortie de l'additionneur : $e_1 R_3 / R_1 = V_{c\text{-}\grave{a}\text{-}c}$.

D'autre part, la tension e_1 mesurée vaut 4V crête à crête avec composante continue de 6V. Soit : $R_3 / R_1 = 0.56 / 4 = 0.14$. En fixant le gain de l'additionneur vis-à-vis du filtre de boucle, à -1 , il vient : $R_2 = R_3$. Comme la sortie du filtre de boucle est capacitive, nous choisissons pour R_2 une valeur assez élevée, soit $R_2 = 68 \text{ k}\Omega = R_3$ et la valeur normalisée de $R_1 = 470 \text{ k}\Omega$, qu'on prend ajustable afin d'obtenir un indice de modulation réglable.

6.1.8. Amplification HF

L'ampli de puissance doit s'adapter avec l'antenne d'émission d'une part, et le modulateur FM à PLL d'autre part. Nous commençons par l'étude du dispositif de rayonnement électromagnétique avant d'arriver à l'ampli en question.

6.1.8.1. Organe de rayonnement HF

Le cahier des charges mentionne la nature inductive de la transmission sans fil, comme discuté au chapitre 3. Le couplage d'induction HF doit être assuré sans noyau et avec peu de spires, notamment du côté récepteur, afin d'être moins encombrant à l'implantation.

Un tel système de faible puissance est schématisé en figure 6-19, lui-même déduit de la figure 3-7. L'énergie HF de la porteuse est amplifiée avant de l'appliquer à la bobine accordée L_1 de transmission. Cette bobine étant couplée à un secondaire L_2 , également accordé.

Le secondaire est suivi d'un circuit limiteur composé d'un redresseur rapide D, d'une capacité de filtrage C_0 à minimum de pertes, et d'une zener interdisant toute augmentation de la tension continue obtenue. Cet effet de limitation a pour rôle d'écarter l'amplitude du signal FM reçu et de la débarrasser d'éventuels parasites superposés. Le rapport signal sur bruit est ainsi amélioré, tout en aménageant une tension V_0 redressée, filtrée, et stabilisée, qui sera utilisée à une fin d'alimentation du récepteur implanté à faible consommation.

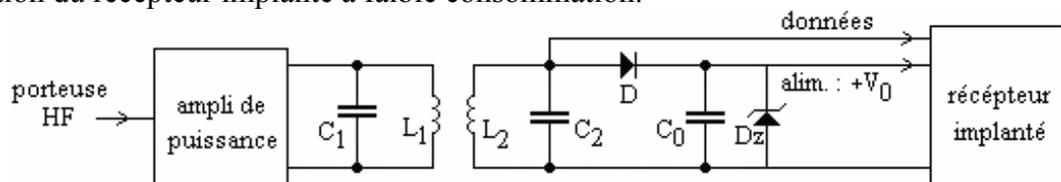


Figure 6-19: Transmission par induction HF avec limiteur avant réception

Ainsi, un même système de rayonnement HF à base de circuits accordés et couplés, est utilisé pour transmettre données et énergie d'alimentation par champ. Le choix et calcul optimisé d'un tel système a été présenté au paragraphe 3-4-4 du chapitre 3, et donne les valeurs suivantes :

- rayon de la bobine primaire $a=1.8$ cm, $n_1=7$ spires, L_1 d'environ $4.4 \mu\text{H}$ accordée par $C_1=1$ nF,
- rayon de la bobine secondaire $b=1.25$ cm, $n_2=5$ spires, L_2 d'environ $1.8 \mu\text{H}$ accordée par $C_2=3.3$ nF,
- distance de séparation entre bobines : $d = 1$ cm.

6.1.8.2. Amplificateur de puissance

Un ampli polarisé en classe C, est tout à fait désigné pour répondre aux exigences du cahier des charges, à savoir, une simplicité associée à un rendement élevé. En plus, les modulations FM et MSK utilisées, tolèrent une amplification non linéaire de ce type. La figure 6-20 montre le circuit que nous avons réalisé.

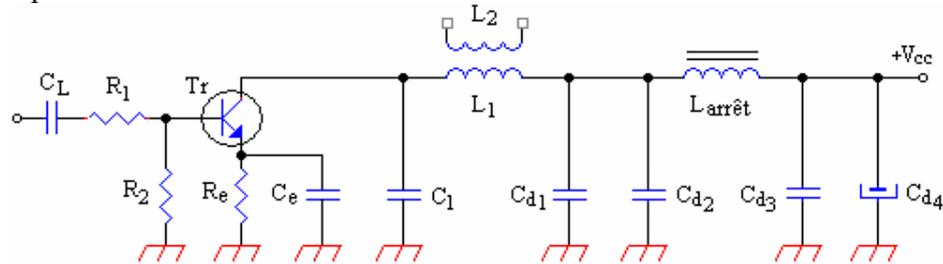


Figure 6-20: Amplificateur de puissance RF

La classe C suppose que le transistor est polarisé négativement à son entrée, de manière à ce que son angle d'ouverture soit faible et que la conduction soit assurée sur moins d'une demie période d'entrée. Dans notre cas, nous nous sommes inspirés de la polarisation automatique d'un FET, dans laquelle la chute de tension à la source est exploitée. De même, la chute de tension continue aux bornes d'une faible résistance d'émetteur R_e , causée par le passage du courant collecteur moyen, nous donne une tension de polarisation V_{BE} négative: $V_{BE} = 0 - V_E$. La composante continue positive du signal de la PLL à l'entrée, est éliminée par la capacité de liaison C_L .

Le circuit $L_1 C_1$ au collecteur, est accordé sur le deuxième harmonique de la porteuse 1 MHz de la PLL. Il régénère ainsi à la résonance, une nouvelle porteuse obtenue par multiplication de fréquence, soit 2 MHz. La modulation y est conservée.

La tension d'alimentation de l'ampli est isolée en dynamique par une self d'arrêt, associée à des condensateurs de découplage en amont et en aval. Ceci, pour éviter que les oscillations HF perturbent la tension $+V_{CC}$.

Le condensateur C_L , choisi égal à $0.1 \mu F$, est à faible impédance à la fréquence d'entrée de 1MHz et ne modifie en rien la forme du signal. Or, la capacité d'entrée C_{be} du transistor intègre le signal sous la constante de temps $\tau = (R_1 // R_2) C_{be}$. Le signal présent à la base ainsi que le courant collecteur résultant, sont illustrés à la figure 6-21.

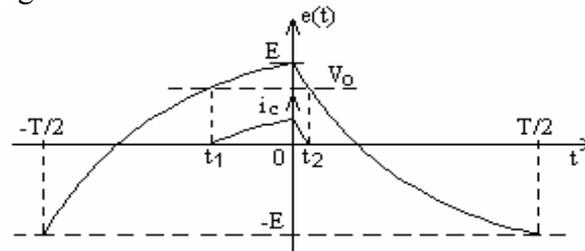


Figure 6-21: Allures de la tension de base $e(t)$ et du courant collecteur $i_c(t)$

Le signal $e(t)$ est composé d'arcs exponentiels entre $\pm E$, avec la relation entre la période T et la constante de temps τ : $k\tau = T/2$ tel que $3 < k \leq 5$.

Les équations du signal $e(t)$ sur une période $-T/2$ à $T/2$, sont :

$$\begin{aligned}
 e(t) &= E(1 - 2e^{\frac{-t - T/2}{\tau}}) & \text{pour } -T/2 \leq t \leq 0 & \text{ avec } e(0) = E(1 - 2e^{-k}) \approx E \\
 e(t) &= E(2e^{-\frac{t}{\tau}} - 1) & \text{pour } 0 \leq t \leq T/2 & \text{ avec } e(T/2) = E(2e^{-k} - 1) \approx -E
 \end{aligned} \quad 6.44$$

Le courant collecteur ne circule que si la tension de base $e(t)$ dépasse celle de l'émetteur V_0 , c'est à dire entre les instants d'ouverture t_1 et t_2 de chaque période.

soit, $i_c = 0$ pour $e < V_0$
 $i_c = S(e - V_0)$ pour $e \geq V_0$, avec $S = \beta / h_{11}$, transconductance du transistor.

Au seuil de conduction, on a l'égalité : $e(t_1) = e(t_2) = V_0$, soit :

$$t_1 / \tau = 0.7 - k - \ln\left(1 - \frac{V_0}{E}\right) < 0 \quad 6.45$$

$$t_2 / \tau = 0.7 - \ln\left(1 + \frac{V_0}{E}\right) > 0 \quad 6.46$$

Le courant collecteur est donc formé d'une suite d'impulsions « intégrées », de valeur moyenne :

$$I_0 = \frac{1}{T} \int_{t_1}^{t_2} i_c(t) dt = \frac{S}{T} \left[2E \tau e^{-T/2\tau} + V_0(t_1 - t_2) - E(t_1 + t_2) \right] \quad 6.47$$

et d'harmoniques :

$$I_n = \frac{2}{T} \int_{t_1}^{t_2} i_c(t) e^{jn\omega t} dt = \frac{2S}{n\pi(jn\omega\tau - 1)} \left[(E - V_0) e^{jn\omega \frac{t_1}{2}} \sin\left(n\omega \frac{t_1}{2}\right) + (E + V_0) e^{jn\omega \frac{t_2}{2}} \sin\left(n\omega \frac{t_2}{2}\right) - E n\omega\tau e^{-T/2\tau} \right] \quad 6.48$$

Dans notre cas, seul l'harmonique 2 est sélectionné soit $n = 2$, et le courant collecteur s'écrit selon la série de Fourier :

$$i_c(t) = I_0 + \operatorname{Re}(I_2) \cos(2\omega t) + \operatorname{Im}(I_2) \sin(2\omega t) \quad 6.49$$

avec $\omega = 2\pi / T$, la fréquence angulaire d'entrée correspondant à 1 MHz, figure 6-22.

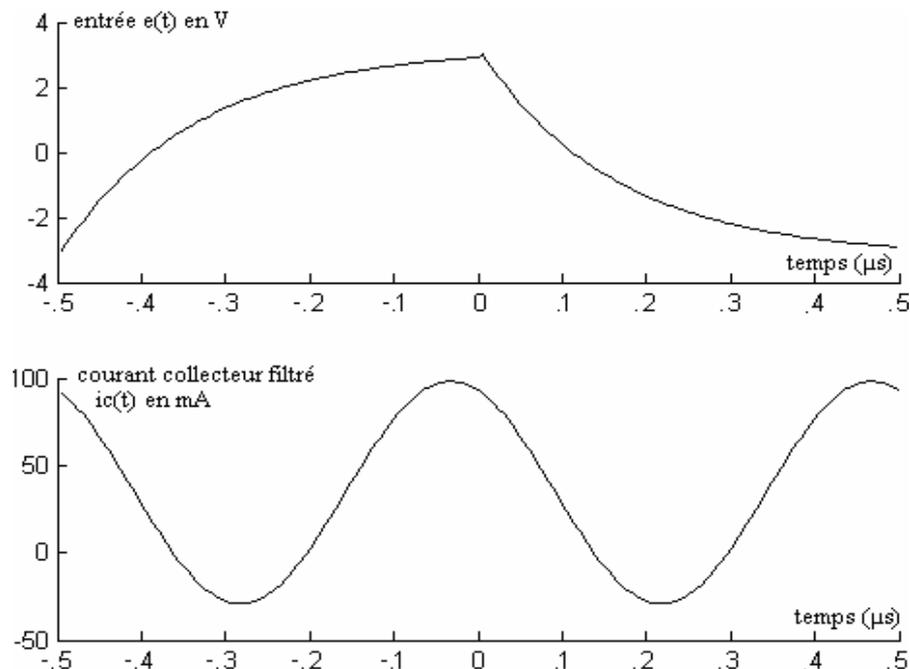


Figure 6-22: Conversion d'une période ($-T/2$ à $T/2$) de la tension exponentielle d'entrée en deux périodes de courant sinusoïdal en sortie

Le signal modulé du VCO est à 10 V crête, il devient à ± 5 V après C_L . La résistance h_{11} d'entrée du transistor étant de l'ordre de 1 k Ω , on choisit les résistances R_1 et R_2 du diviseur de tension du même ordre de grandeur pour adaptation, soit $R_1 = 1$ k Ω et $R_2 = 1.8$ k Ω .

Les tensions de crête du signal intégré de base, se trouvent alors à $E = \pm 3.2$ V.

La tension d'émetteur est donc à $V_0 = E - 0.6 = 2.6$ V.

On fixe une faible valeur de résistance à l'émetteur, $R_e = 75\Omega$, qu'on découple par un condensateur de 0.1 μ F. Le courant moyen du collecteur devient égal à : $I_0 = V_0 / R_e = 34.6$ mA.

Un programme réalisé en Matlab, nous permet de retrouver la valeur de I_0 et ceci, en essayant plusieurs valeurs du rapport $k = T / 2\tau$. La valeur trouvée est égale à : $k = 3.1$, soit $\tau = 161$ ns avec $T = 1$ μ s. On a de même :

- la pente du transistor, $S = 40 I_0 = 1.36 \text{ A/V}$.
- les seuils temporels, $t_1 = -117 \text{ ns}$ et $t_2 = 17 \text{ ns}$.
- les angles d'ouverture du transistor, $\theta_1 = \omega t_1 = -42^\circ$ et $\theta_2 = \omega t_2 = 6^\circ$.

On déduit aussi :

- le courant crête du collecteur, $I_2 = (58.3 - j 25.3) \text{ mA}$, soit $|I_2| = 63.6 \text{ mA}$.
- la puissance à l'entrée du transistor est de l'ordre de $P_e = (E - V_0)^2 / 8h_{11} = 0.8 \text{ mW}$.
- la puissance utile au primaire étant limitée à $P_p = 0.44 \text{ W}$, le gain en puissance de l'ampli devient : $G_p = P_p / P_e = 553.6$ soit 27.4 dB .
- le rendement de l'ampli, vis-à-vis de son alimentation $V_{cc} = 12 \text{ V}$, est donc :

$$\eta_{\text{ampli}} = \frac{P_p}{V_{cc} \left[I_0 + \frac{|I_2|}{\sqrt{2}} \right]} = 46.4 \% .$$

Ce rendement est relatif au second harmonique et non au fondamental de I_C .

Les circuits accordés au primaire et secondaire, transforment la porteuse FM impulsionnelle en porteuse FM sinusoïdale de tension crête à crête de : $V_{c_{\text{pic à pic}}} = 4 P_p / |I_2| = 27.6 \text{ V}$.

Notre transmetteur final d'implant se compose donc d'un circuit externe associé à la carte DSP, figure 6-23. Il renferme plusieurs circuits.

Un premier étage assure une bufférisation des données à l'aide d'un simple transistor 2N2222. Un condensateur accélérateur de 75 pF shunte la résistance de base afin d'améliorer les temps de commutation du transistor. Ceci est une exigence de l'étage suivant, constitué par le circuit GAL 22v10, pour qu'il reçoive convenablement des données TTL. Dans ce circuit, sont implémentés un modulateur MSK et un générateur de bits pseudo-aléatoires (GPA) basé sur un polynôme générateur de degré 5. Ce générateur réclame à la mise sous tension, une initialisation à 1 de son entrée au pin 4. Un circuit RC dérivateur connecté à la ligne d'alimentation +5V, assure cette fonction automatiquement. Un bouton poussoir est aussi prévu pour l'assurer manuellement.

Le filtre de transmission fait suite au modulateur MSK. Un condensateur de liaison d'au moins $1 \mu\text{F}$ est indispensable, afin de ne pas altérer les deux composantes continues appliquées chacune à l'une de ses armatures : 2.5 V du côté GAL et 6 V du côté filtre.

Quant au filtre, il s'agit d'un passe-bas actif du deuxième ordre. Le signal MSK ainsi filtré est additionné au signal d'erreur d'une PLL, filtré lui aussi par le filtre de boucle. La PLL constitue la base d'un modulateur FM, assurant par régulation une grande stabilisation de la fréquence porteuse, chose recommandée dans un modulateur FM. Un timer du DSP, piloté par un quartz, fournit cette fréquence porteuse de 1 MHz .

L'ampli est équipé d'un seul transistor polarisé en classe C. Son circuit au collecteur, est accordé sur la deuxième harmonique de la porteuse 1 MHz . Il régénère ainsi à la résonance, une nouvelle porteuse obtenue par multiplication de fréquence, soit 2 MHz .

En pratique, le transistor 2N2222 utilisé comme transistor de puissance, a présenté d'excellentes performances où nous avons pu avoir une tension crête à crête au collecteur jusqu'à 30 V . Mais cela n'a pas persisté car il chauffait trop et tomba souvent en panne. La cause est liée à la puissance moyenne continue qui a été jusqu'ici ignorée. La tension du collecteur au repos est de $V_{ce0} = V_{cc} - V_0 = 9.45 \text{ V}$, contribuant à une puissance totale dissipée au collecteur de $P_{\text{tot}} = V_{ce0} I_0 + P_p = 0.76 \text{ W}$, qui dépasse la puissance maximale de dissipation du 2N2222. Ne voulant pas équiper le transistor d'un dissipateur de crainte d'encombrer davantage la réalisation, nous l'avons échangé contre un BD135 de puissance de dissipation totale de 8 W et une fréquence de transition de 190 MHz . Seulement, nous n'avons pas touché au calcul déjà fait des éléments l'entourant, et qui sont prévus pour une puissance émise de 0.44 W . La seule modification apportée se ramène au renforcement de la capacité d'intégration C_{BE} par un condensateur en parallèle, de valeur : $C = \tau / (R_1 // R_2) = 0.16 \mu\text{s} / 0.68 \text{ k}\Omega = 270 \text{ pF}$. Le filtre d'entrée du

transistor prend alors la structure d'un pont de Wien. Le montage pratique du transmetteur de données est en figure 6-23.

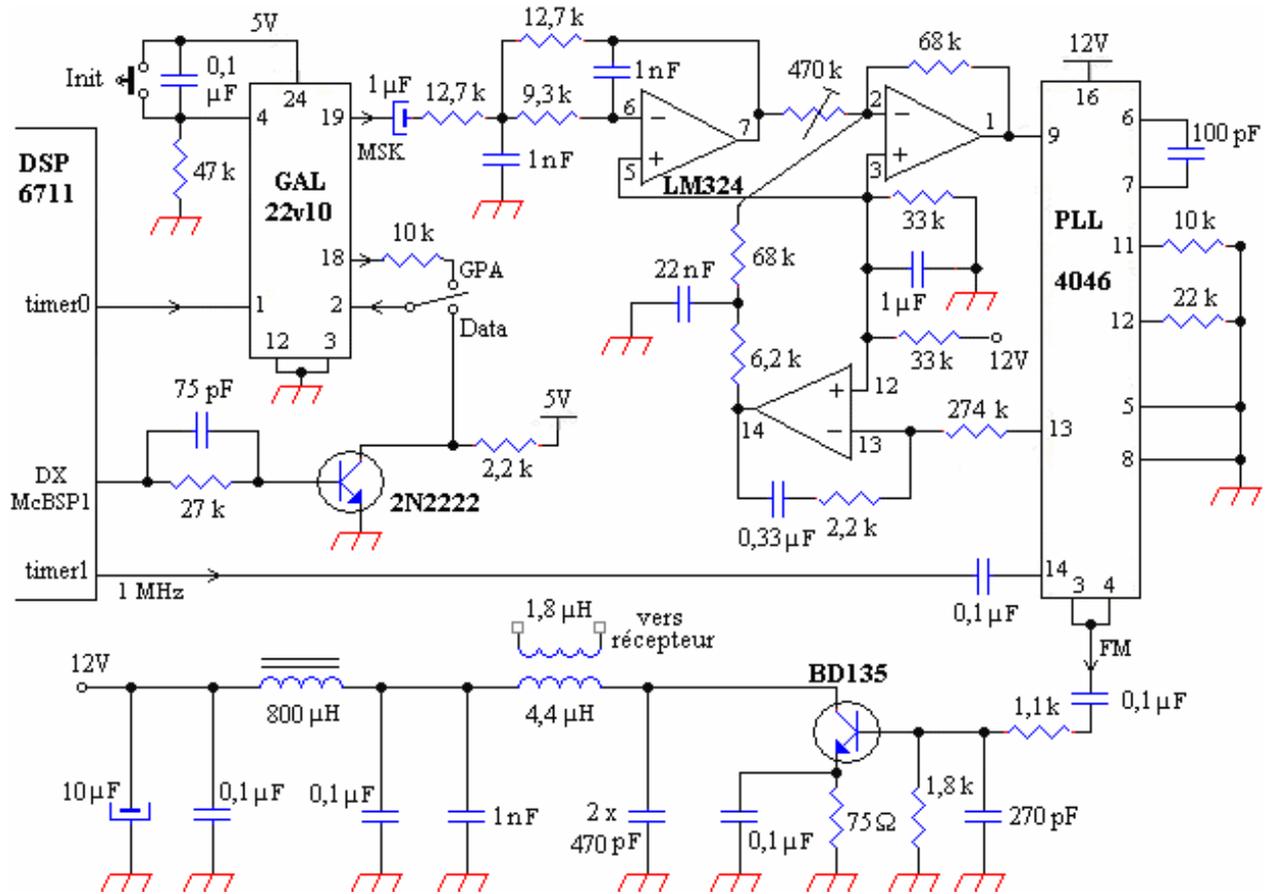


Figure 6-23: Transmetteur MSK-FM de paramètres de stimulation cochléaire

6.2. Conception et réalisation du récepteur d'implant

6.2.1. Aspect général

Le modèle de réception sera le modèle réciproque du modèle de transmission établi en figure 6-1, d'où découle le schéma synoptique de la figure 6-24.

Après l'antenne de réception, vient l'amplificateur HF qui doit présenter un faible facteur de bruit. Dans notre cas, il est précédé du circuit limiteur attaché à la modulation FM.

Le démodulateur HF transpose en bande de base, l'information contenue dans la porteuse modulée. Le décodeur de source permet de convertir la forme analogique de l'information obtenue en signal numérique, et ceci grâce à un seuil de décision. Dans notre cas, il assure aussi la démodulation MSK. En sortie, deux voies au moins sont générées; une de données et l'autre propre à leur cadence. Le décodeur de canal détecte les éventuelles erreurs imprégnant le message utile.

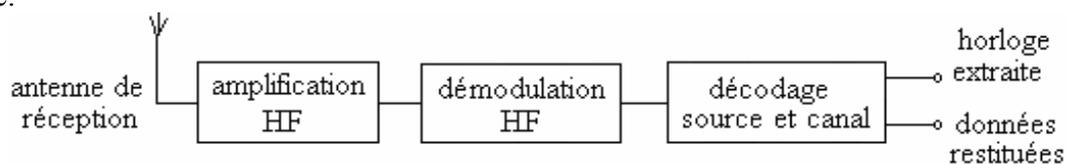


Figure 6-24: Schéma général de réception

6.2.2. Circuit limiteur

Vu la tolérance des composants d'accord utilisés dans le couplage du primaire et secondaire,

un signal optimal est récupéré sur une porteuse de 1.9 MHz. Le recours aux condensateurs ajustables n'apporte aucun bénéfice, vu qu'ils n'atteignent pas le nanoFarad.

Le circuit du secondaire, jouant le rôle d'antenne réceptrice de la FM sur 1.9 MHz, est connecté à un circuit de limitation d'amplitude comme le montre la figure 6-25.

Le circuit est muni de deux cellules d'écrtage assurant deux niveaux symétriques par rapport à zéro. Le niveau capté au secondaire étant assez élevé, une diode rapide, la 1N4148, suivie d'une diode Zener, la BZX460 de 5.6 V, composent chaque cellule. L'amplitude du signal est donc limitée symétriquement à ± 5.6 V et se voit ainsi débarrassée des parasites HF superposés. En ajoutant des condensateurs de filtrage, tels que $C = 0.1 \mu\text{F}$ et $C' = 6.8 \mu\text{F}$ (condensateur-goutte à 6.3 V), les niveaux de tension récoltés aux bornes de chaque Zener deviennent par là, filtrés et stabilisés pouvant alimenter sous ± 5.6 V les circuits à faible puissance du récepteur.

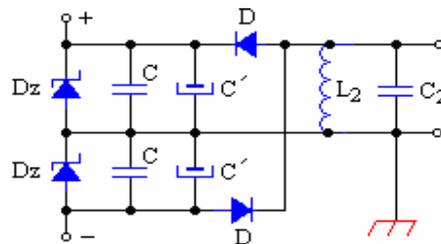


Figure 6-25: Circuit de limitation et d'alimentation à faible puissance

6.2.3. Amplificateur-isolateur HF

Un étage simple et compact est utilisé pour assurer d'une part, l'amplification HF du signal capté et d'autre part, une séparation nécessaire pour isoler le circuit accordé du démodulateur FM de celui du secondaire. En effet, il a été constaté comme dans le cas des étages FI d'un poste de radiodiffusion, qu'il y a interaction mutuelle entre les trois circuits accordés du primaire, du secondaire, et du circuit de démodulation FM. Le présent étage vise donc à réduire cet effet en introduisant une certaine isolation. Comme illustré dans la figure 6-26, l'élément actif de l'étage n'est autre qu'un transistor MOSFET double grille à canal N diffusé, le MFE130 de Motorola, qui peut monter jusqu'à 200 MHz. Il est logé dans un boîtier métallique du type TO72 à quatre broches.

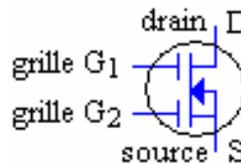


Figure 6-26: MOSFET double grille à canal N appauvri

La caractéristique de transconductance la plus utilisée dans le cas des MOSFETs est celle de saturation, à forme quadratique : $I_D = K (V_{GS} - V_T)^2$, avec K , paramètre de construction et V_T , tension seuil de fermeture du canal.

Sous une alimentation de 12 V, des mesures ont été menées sur ce transistor pour déterminer ses constantes. Le courant du drain s'annule lorsque $V_{G1S} \leq -1$ V ou $V_{G2S} \leq -1.3$ V, soit :

$$V_{T1} = -1 \text{ V} \quad \text{et} \quad V_{T2} = -1.3 \text{ V}.$$

Pour $V_{G1S} = 0$, on mesure un courant-drain de $I_D = K_1 V_{T1}^2 = 2.2$ mA, soit $K_1 = 2.2$ mA/V².

Alors que pour $V_{G2S} = 0$, on mesure $I_D = K_2 V_{T2}^2 = 3.6$ mA, soit $K_2 = 2.1$ mA/V².

Les deux caractéristiques de transconductance deviennent alors :

$$\begin{cases} I_D \text{ (mA)} = 2.2 (V_{G1S} + 1)^2 \\ I_D \text{ (mA)} = 2.1 (V_{G2S} + 1.3)^2 \end{cases} \quad 6.50$$

Leurs tracés sont donnés en figure 6-27.

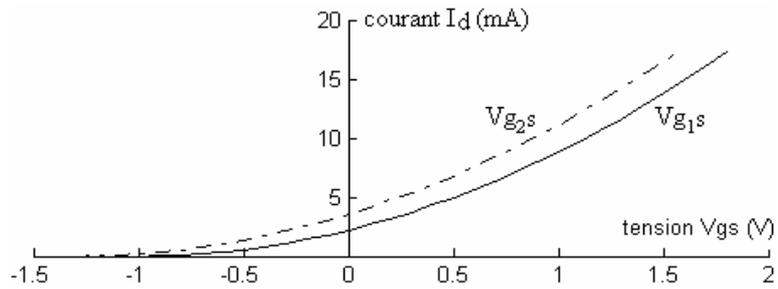


Figure 6-27: Caractéristiques de transconductance, I_D en fonction de V_{G1S} et V_{G2S}

Le MOSFET double grille peut être considéré comme un double transistor pouvant réaliser simplement et efficacement le célèbre montage cascode. C'est un montage stable dans lequel, une source commune est suivie d'une grille commune. Dans notre cas, la grille G_1 sera l'entrée pour le premier étage à source commune, qui à son tour, communique le signal préamplifié à l'étage à grille G_2 commune. Le signal amplifié est récupéré au niveau du drain, figure 6-28.

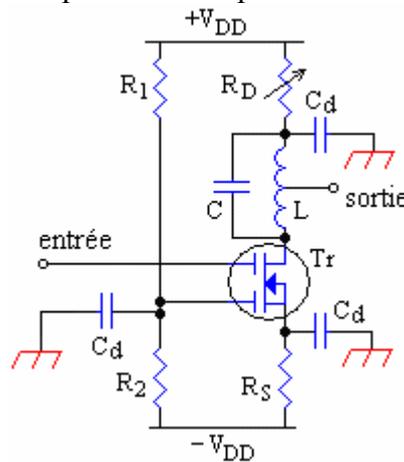


Figure 6-28: Ampli cascode à base d'un MOSFET double grille

La deuxième grille sert à fixer la polarisation du transistor et de là, fixe sa pente et son amplification. Sous une tension d'alimentation de ± 5 V, la tension continue de la grille G_2 est portée à 0 V grâce aux deux résistances de pont $R_1 = R_2 = 470$ k Ω , munies d'une capacité de découplage à la masse de 22 nF.

Selon la figure 6-27, une bonne linéarité est obtenue sous un fort courant de drain. En fixant le courant de repos à $I_D = 14$ mA sous $V_{G2S} = -V_{T2} = 1.3$ V, nous obtenons une tension de source égale à $V_S = -1.3$ V, fixant à son tour la résistance de source à $R_S = (5 - 1.3) / 0.014 = 264$ Ω . En pratique, cette résistance est prise égale à 300 Ω , réduisant ainsi le courant de repos à $I_D = 12.7$ mA. La résistance du drain, quant à elle, est prise du même ordre de grandeur que R_S , soit $R_D = 220$ Ω ajustable.

En régime dynamique, la résistance au drain R_D est découplée par une capacité de 22 nF tandis que celle à la source R_S , est fortement découplée par une 10 μ F. La charge du drain n'est alors constituée que par le circuit accordé LC en parallèle avec la résistance d'entrée R du circuit démodulateur qui suit, figure 6-29.

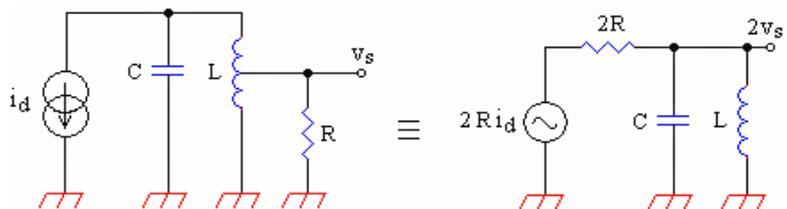


Figure 6-29: Charge du drain avec son circuit de Thévenin équivalent
 La prise médiane de la bobine L, permet de ramener une résistance double aux bornes de celle-ci.
 Le circuit de Norton obtenu est ensuite remplacé par son équivalent de Thévenin.
 La transmittance v_s / v_e d'un tel circuit, où $v_e = R i_d$, est :

$$\frac{v_s}{v_e} = \frac{1}{1 + jQ \left(\frac{\omega}{\omega_r} - \frac{\omega_r}{\omega} \right)} \quad 6.51$$

avec, $\omega_r = (LC)^{-1/2}$ est la fréquence angulaire de résonance,

$Q = 2R / L\omega_r$ est le coefficient de qualité du circuit chargé par $2R$.

Le choix de ces deux paramètres ainsi que la résistance d'amortissement R, sera discuté dans le paragraphe suivant.

Le gain en tension du montage est : $|G_V| = g_m 2R$, avec la pente du transistor au point de repos P_0 fixé par la tension continue V_{G2S} :

$$g_m = \left. \frac{\partial I_D}{\partial V_{G2S}} \right|_{P_0} = 2K_2 (V_{G2S} - V_{T2}) = 2 \times 2.1 \times 2.6 = 10.9 \text{ mA/V} \quad 6.52$$

6.2.4. Discriminateur d'amplitude

Le discriminateur d'amplitude à flanc figure parmi les plus simples schémas de démodulation FM. Comme préconisé par le cahier des charges, il s'agit du circuit le moins encombrant faisant appel à une transformation du signal FM en AM par le biais d'un classique circuit LC. La courbe en cloche d'un circuit résonant LC présente deux flancs autour de sa fréquence de résonance f_r , figure 6-30.

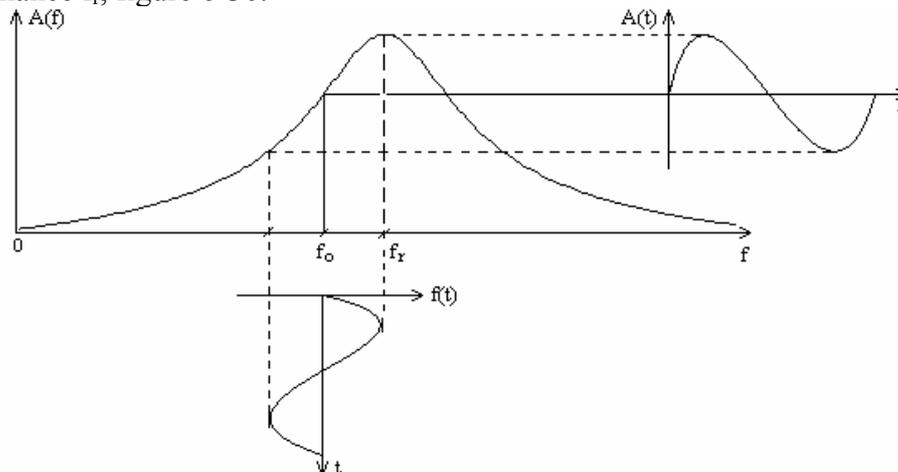


Figure 6-30: Flanc d'une courbe en cloche utilisé dans la transformation FM-AM

Un seul flanc suffit pour assurer la dite transformation. Cependant, le circuit doit résonner à une fréquence f_r voisine de la porteuse f_0 . Pour cela, f_r sera au minimum décalée de la valeur d'une excursion de la porteuse f_0 de l'onde FM. L'excursion en fréquence estimée dans le paragraphe 6-1, est de $\Delta F = 36 \text{ kHz}$, soit $f_r = f_0 + \Delta F = 1.9 + 0.036 = 1.936 \text{ MHz}$. En fabricant une self de l'ordre de $L = 1 \mu\text{H}$, il vient : $C = 6.8 \text{ nF}$. La bande minimale du circuit LC est donc, $B = 2 \Delta F$. Ce qui donne comme coefficient de qualité : $Q \geq f_r / B = 1.936 / 0.072 = 27$.

Le problème de non linéarité soulevé par ce type de discriminateur, même allégé par un faible coefficient de qualité, n'est point un handicap. Ceci est dû à ce que la démodulation porte sur la transformation d'un saut de fréquence (celui de la MSK) en un saut d'amplitude.

La résistance de charge minimale correspondant à $Q = 27$, peut être déterminée par :

$R \geq Q L \omega_r / 2 = 164 \Omega$, qui représente la résistance d'amortissement du détecteur d'enveloppe. La résistance effectivement utilisée dans la cellule de détection, s'élève au double à cause de la self à point milieu utilisée, figure 6-31, d'où $2R \geq 328 \Omega$.

La résistance de charge R de la figure 6-29, représente la résistance d'amortissement du détecteur d'enveloppe de la figure 6-31. Tandis que la résistance, effectivement utilisée dans la cellule de détection, s'élève au double, soit $2R$.

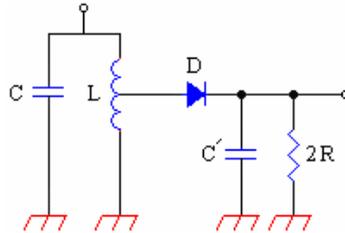


Figure 6-31: Discriminateur d'amplitude à flanc

La constante de temps du détecteur d'enveloppe doit vérifier l'inégalité suivante :

$T_{HF} \ll 2RC' < T_{BF}$ où, $T_{HF} = 1 / 2\text{MHz}$ et $T_{BF} = 1 / B_{MSK}$ désignant respectivement, la période de la porteuse et celle limitant la bande du signal modulant, soit : $0.5\mu\text{s} \ll 2RC' < 70\mu\text{s}$. En fixant $R = 4.7 \text{ k}\Omega$, il vient : $53 \text{ pF} \ll C' < 7.5 \text{ nF}$, qu'on prend égale à $C' = 2 \text{ nF}$. Le gain en tension de l'amplificateur s'élève donc à $|G_V| = g_m 2R = 10.9 \times 9.4 = 102$, soit 40 dB.

Une liaison capacitive entre le détecteur d'enveloppe et le circuit résonnant qui le précède, figure 6-28, est responsable d'un effet nuisible (effet clamping) qui apparaît lors de la présence en série d'un condensateur avec une diode. Nous avons eu l'idée d'ajuster la tension continue du drain sur zéro volt, ou mieux, sur 2.5 V de façon à pouvoir utiliser un Trigger de Schmitt à circuit logique et de se dispenser ainsi de la liaison capacitive.

6.2.5. Démodulateur MSK

Ce circuit peut être assigné à la fonction du décodeur de source de la figure 6-24. La démodulation consiste donc à convertir les signaux sinusoïdaux de la MSK en signaux NRZ digitaux. Une démodulation non cohérente se fait par un discriminateur de fréquence suivi d'un circuit de décision à seuil. Cette dernière technique simplifie beaucoup la réalisation, mais les performances diminuent de 6.6 dB par rapport à une modulation cohérente [48].

Dans le cadre d'une démodulation non cohérente, on constate sur la figure 6-5 que le signal MSK présente pour la logique 0, un passage par zéro supplémentaire que la logique 1. En d'autres termes, un démodulateur MSK doit pouvoir discriminer jusqu'à un seul passage par zéro pour pouvoir démoduler le signal.

Dans notre cas, le démodulateur MSK traite les signaux démodulés en FM pour en extraire, données et horloge, selon le procédé de la figure 6-32.

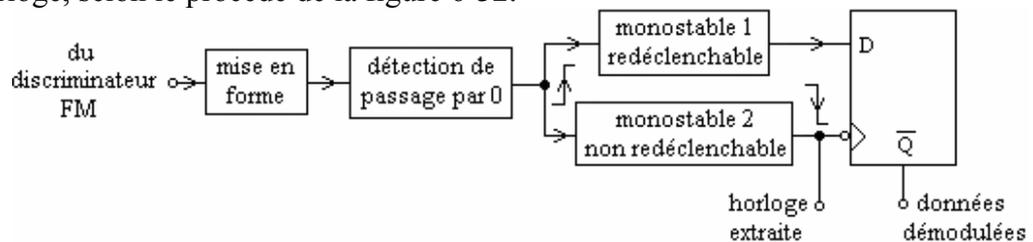


Figure 6-32: Principe de notre démodulation MSK non cohérente

Il y a cinq blocs constituant le démodulateur MSK. La sortie du dernier bloc est celle des données récupérées. Les différents blocs seront discutés dans l'ordre de la progression du signal.

6.2.5.1. Circuit de mise en forme :

Le trigger de Schmitt de la figure 6-33, reproduit la forme digitale du signal MSK tout en filtrant les parasites éventuels placés dans sa fenêtre d'hystérésis. Les seuils d'hystérésis jouent le rôle de seuils de décision dans la mise en forme NRZ du signal d'entrée.

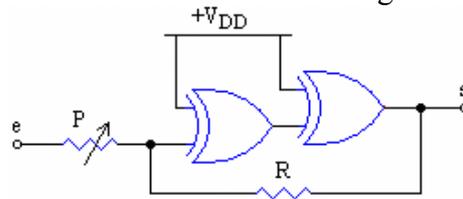


Figure 6-33: Trigger de Schmitt à portes logiques

Les portes XOR de la figure 6-33 assurent chacune, la simple fonction d'inversion logique. Elles sont de type CMOS. Les seuils d'hystérésis de ce type de trigger sont fonction des résistances P et R. Ces seuils se trouvent symétriquement, de part et d'autre, de la mi-tension d'alimentation V_{DD} :

$$e_{1,2} = \left(1 \pm \frac{P}{R}\right) \frac{V_{DD}}{2} \quad 6.53$$

La sortie du démodulateur FM doit donc présenter une composante continue de $V_{DD}/2 = 2.5$ V. En fixant la fenêtre d'hystérésis à $\Delta e = PV_{DD}/R = 0.5$ V, on tire, $R = 10$ P.

En choisissant $R = 100$ k Ω , la résistance P sera un ajustable de 10 k Ω pour pouvoir régler en pratique, la fenêtre d'hystérésis et ses deux seuils de décision, de 0 à ± 0.25 V autour de 2.5 V.

6.2.5.2. Détecteur de passage par zéro :

L'indicateur de passage par zéro génère une impulsion à chaque changement de polarité du signal d'entrée. Lorsque le signal mis en forme par le trigger passe par zéro, la sortie du détecteur bascule à l'état haut durant un bref instant ajusté par la cellule RC de retard, figure 6-34.

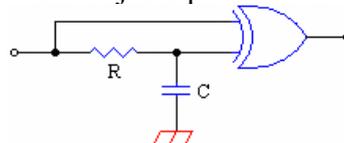


Figure 6-34: Indicateur de passage par zéro

La porte XOR, en technologie CMOS, dispose à ses deux entrées du signal d'entrée NRZ et de son intégration par la cellule RC. Ainsi, si le signal d'entrée transite à l'état haut, son signal intégré arrive à $V_{DD}/2$ au bout de 0.7 RC. La constante de temps d'intégration RC doit être négligeable devant la durée d'un bit $T_b = 208$ μ s, soit $RC < T_b/10$. Pour $RC = 12$ μ s et R fixée à 10 k Ω , on tire $C = 1.2$ nF.

Les impulsions délivrées par le détecteur de passage par zéro, sont donc de durée 8.4 μ s et arrivent à une cadence combinée de T_b et $T_b/2$ selon le rythme de la modulation.

6.2.5.3. Temporisateurs de prétraitement :

La démodulation en fréquence doit différencier entre les deux sauts fréquentiels du MSK, basée sur la fréquence du signal de passage par zéro. Cette fonction est confiée à deux monostables, qui répondent à chaque front montant d'impulsion délivrée par le détecteur de passage par zéro, figure 6-35.

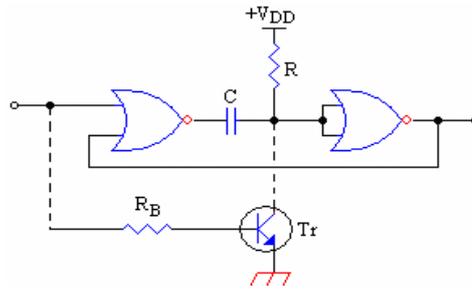


Figure 6-35: Configuration des deux temporisateurs en redéclenchable et non redéclenchable. Toujours en technologie CMOS, l'un des monostables est pris redéclenchable (tout le circuit de la figure 6-35, y compris le circuit en pointillé) tandis que l'autre est non redéclenchable (sans le circuit en pointillé de la figure 6-35).

Lorsque la temporisation du monostable expire, sa sortie retourne à l'état bas. Si une nouvelle impulsion d'entrée arrive plus tôt, avant expiration du délai de temporisation, elle ne sera pas prise en compte dans le cas 'non redéclenchable', mais elle le sera dans le cas 'redéclenchable' grâce au circuit additionnel à transistor. Ce dernier a pour rôle de décharger la capacité C du monostable afin de réinitialiser la temporisation tout en maintenant la sortie du monostable à l'état haut.

Les signaux de démodulation sont montrés sous des conditions parfaites, sur la figure 6-36.

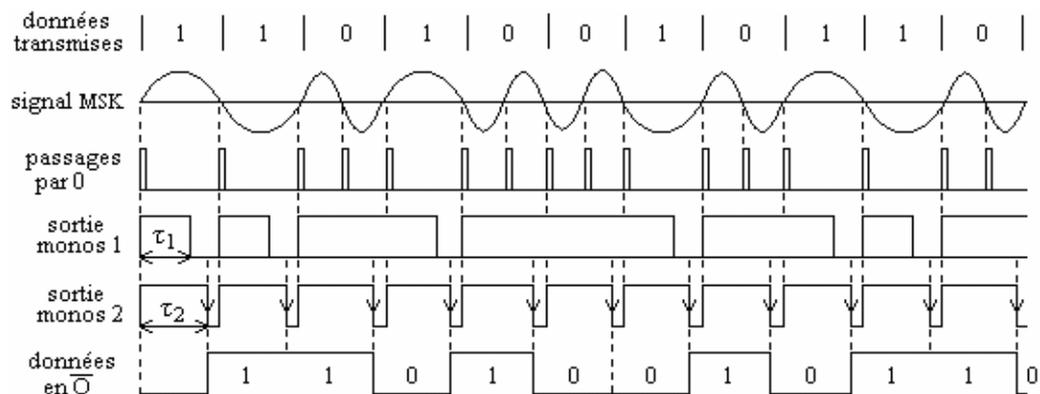


Figure 6-36: Diagrammes temporels des signaux de démodulation MSK

Dans cette figure, l'exemple de la donnée transmise 11010010110, illustre les signaux générés au niveau des différents blocs. Le signal MSK, objet de la démodulation, correspond à la donnée transmise. Comme on le voit sur ce signal et durant un temps de bit, un bit de donnée avec une valeur de 1 possède un seul passage par zéro, un bit de donnée avec une valeur de 0 possède deux passages par zéro.

Le monostable 1 (redéclenchable) fournit à sa sortie des données codées, tel qu'un « un logique » lui correspond une impulsion de durée $\tau_1 = 0.7 R_1 C_1$ telle que : $T_b / 2 < \tau_1 < 3T_b / 4$, alors qu'un « zéro logique » lui est associé par recouvrement, une impulsion de durée d'un bit T_b . Les éléments R_1 et C_1 de ce monostable sont déduits à partir de l'inégalité de temporisation :

$$149 \mu\text{s} < R_1 C_1 < 223 \mu\text{s}, \text{ soit } R_1 = 18 \text{ k}\Omega \text{ et } C_1 = 10 \text{ nF}.$$

Le monostable 2 (non redéclenchable) génère une horloge de référence à la vitesse de transmission de données (baud rate). La durée d'impulsion $\tau_2 = 0.7 R_2 C_2$ est fixée entre :

$$3T_b / 4 < \tau_2 < T_b. \text{ Les éléments } R_2 \text{ et } C_2 \text{ du monostable 2 sont déduits à partir de cette inégalité : } 223 \mu\text{s} < R_2 C_2 < 298 \mu\text{s}, \text{ soit } R_2 = 26 \text{ k}\Omega \text{ et } C_2 = 10 \text{ nF}.$$

6.2.5.4. Régénérateur de données :

Une bascule D utilise les signaux en sorties des monostables pour reconstruire l'information digitale. Les pseudo-données issues du monostable 1 sont échantillonnées sur chaque front descendant de l'horloge émanant du monostable 2. La sortie \bar{Q} de la bascule délivre alors les données synchronisées sur l'horloge restituée, figure 6-36. Cependant, la bascule peut être omise en programmant convenablement le port série de réception de la carte de traitement des données qui suit. De même, la fonction de décodage du canal se fera par programmation en tant que premier traitement accompli par cette carte.

Il est à noter que ce type de démodulation est non cohérent du moment que le récepteur n'est pas synchronisé avec le transmetteur, ni en phase ni en temps (retard τ_1 d'apparition des données en figure 6-36). La figure 6-37 donne le schéma global du récepteur.

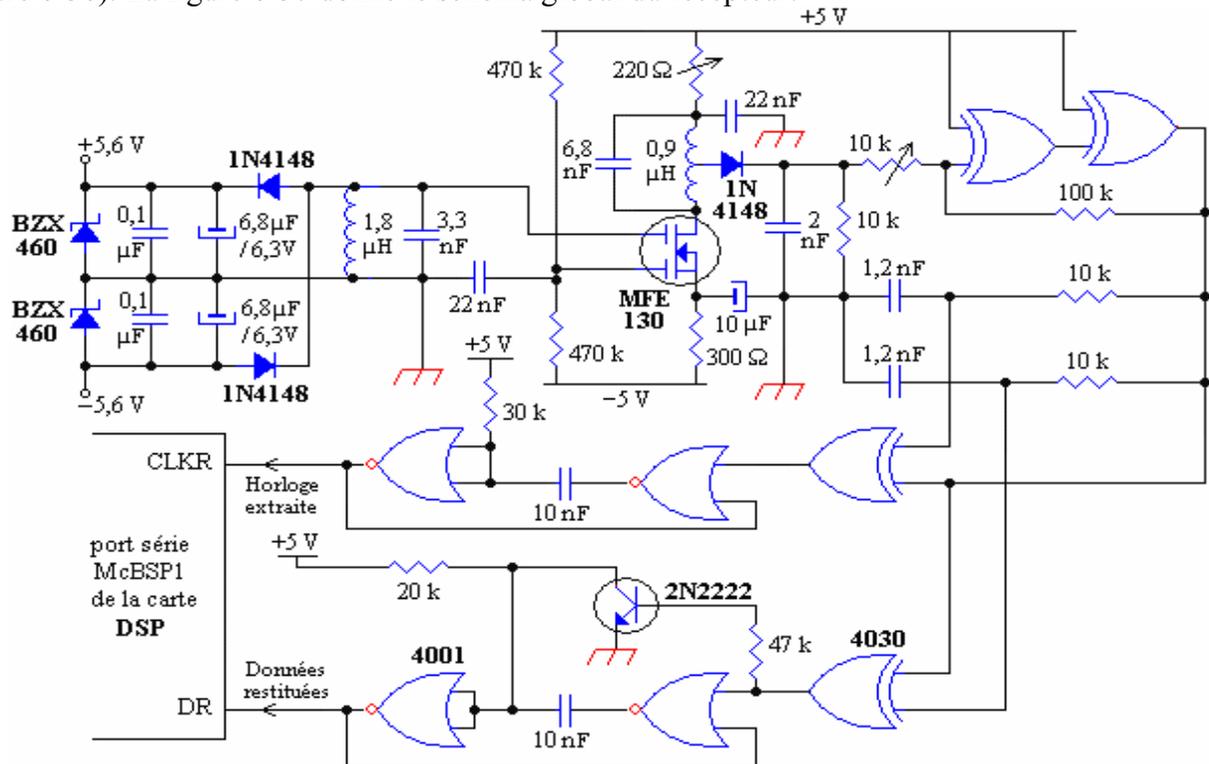


Figure 6-37: Récepteur FM-MSK de paramètres de stimulation cochléaire

6.3. Logiciel de la réalisation

Ce paragraphe décrit la gestion software en bande de base de l'ensemble du transmetteur et récepteur de données de notre implant cochléaire. Les différentes tâches sont réparties sur les périphériques du DSP pendant leurs configurations, alors que les connexions entre la carte à DSP et les cartes de transmission ou de réception, sont illustrées par les figures 6-23 et 6-37.

6.3.1. Développement du projet d'implant sous CCS

Le logiciel de l'application est constitué de six fichiers sources, quatre écrits en langage C : aquisition.c, analyse.c, trans_rcpt.c, rtdx.c, et deux fichiers écrit en VisualBasic : writertdx.bas et readrtdx.bas. Les contenus de ces fichiers sont discutés ci-après.

Une fois les fichiers sources de notre application sur la transmission-réception de données étant préparés, la construction du projet gérant l'application sous l'environnement CCS (Code Composer Studio), suit les étapes de développement suivantes :

Au menu *Project \ New*, notre projet est nommé « implant.pjt » en choisissant le type exécutable (et non librairie) avec la cible *TMS320C67XX*.

6.3.1.1. Fichier de configuration :

Le fichier de configuration nommé *implant.cdb*, renferme diverses configurations d'objets saisies graphiquement. Pour cela :

Au menu *File \ New \ DSP/BIOS Configuration*, la plate-forme *dsk6711.cdb* propre à notre carte DSP, est sélectionnée. L'outil DSP/BIOS de configuration graphique s'affiche et certaines catégories sont à programmer :

a) Application de la catégorie Instrumentation :

Au pop-menu du gestionnaire *LOG-Event Log Manager*, on clique sur *Insert LOG* afin d'insérer un nouvel objet *LOG0* qu'on renomme « Trace ». Cet objet est un buffer d'affichage de messages ou données avec un impact minimal sur l'application en temps réel.

b) Application de la catégorie Scheduling :

Dans le module d'interruptions hardware «HWI-SR Manager» : Voie «HWI_RESERVED1» : est programmée la fonction *_RTDX_Poll* pour la communication RTDX (Real Time Data eXchange). Dans notre cas, l'application RTDX sert à capturer en temps réel, les sorties d'algorithmes de notre vocodeur à formants et les afficher sur des fenêtres conçues sous Visual Basic afin d'en vérifier l'implémentation.

Dans Voie «HWI_INT8 (EDMA_Controller)» : est programmée la routine *_hwiEdma*.

Dans Voie «HWI_INT15 (McBSP1_Receive)» : est programmée la routine *_HWIRecept*.

Dans le module «SWI- Manager» :

Objet «swiProcess» : est programmée la routine *_swiEdma* à priorité=1.

Objet «SWIobj» : est programmée la routine *_SWIenvoi* à priorité=2.

c) Application de la catégorie Input/Output :

Dans le module «RTDX-Settings»: est programmé *entreeRTDX* comme canal d'entrée des données de configuration de l'implant cochléaire, vers le DSP. Le canal *sortieRTDX* est programmé en sortie du DSP pour visualiser en temps réel, les paramètres de la stimulation cochléaire.

d) Application de la catégorie CSL :

Dans le module «Extern Declaration»: est programmé, *paquet*, comme un tableau de trois entiers à 32 bits non signés (représentant les paramètres de stimulation cochléaire, concaténés avec leurs CRC).

Dans le module «EDMA»: Voie «Configuration Manager»: sont programmées les configurations de transfert de données suivantes :

- *CfgReceive* : destinée au canal EDMA n°13, table 5-1, elle assure un transfert à haute priorité de 160 données à 16 bits reçues au port de réception du McBSP0, vers la mémoire par décrémentation d'adresse. Après transfert, une réinitialisation automatique s'effectue pour le prochain transfert tout en enchaînant un autre transfert via le canal EDMA n°8.
- *CfgChainage* : destinée au canal EDMA n°8, elle assure un transfert en bloc à basse priorité de 160 données à 32 bits, d'une zone mémoire à une autre par décrémentation d'adresses. Après transfert, une réinitialisation automatique s'effectue pour le prochain transfert tout en envoyant une interruption hardware au CPU pour l'informer de la fin du transfert.
- *CfgSynch* : destinée au canal EDMA n°14, elle assure un transfert initial à basse priorité de la première donnée à 32 bits de "*paquet*", vers le port de transmission du McBSP1. Ce transfert est répétée 24 fois et constitue une salve de synchronisation à l'entrée du port de réception, par manque d'un vrai signal de synchronisation des trames à l'entrée FSR de ce même port. Après ce transfert, le canal s'initialise avec la configuration *CfgEnvoi*.

- *CfgEnvoi* : Elle assure un transfert à haute priorité via le canal EDMA n°14, des trois données à 32 bits de "paquet" par incrémentation d'adresses, vers le port de transmission du McBSP1. Après transfert, une réinitialisation automatique s'effectue pour le prochain transfert de "paquet".

Dans le module «MCBSP»: Voie «Configuration Manager»: est programmée la configuration *mcbSPCfg0*, qui est destinée au port série McBSP0. Ce port s'occupe de la réception des données à 16 bits de la parole, échantillonnée à 8 kHz par le codec.

La configuration *mcbSPCfg1* est destinée au port série McBSP1, qui s'occupe de la transmission et la réception des paquets de paramètres de la stimulation cochléaire.

Dans le module «TIMER»: Voie «Configuration Manager»: est programmée la configuration *timerCfg0*, qui est destinée au timer0. Ce temporisateur assure un débit des données transmises au rythme de 4800 bps.

La configuration *timerCfg1* est destinée au timer1, qui doit générer une porteuse de 1 MHz.

En quittant l'outil de configuration graphique, nous sauvegardons nos configurations sous le nom de fichier, *implant.cdb*. Ceci génère automatiquement un fichier de liaison ou linkage, *implantcfg.cmd*, et deux fichiers C : un fichier de source, *implantcfg.c.c*, et un fichier d'entête (ou header), *implantcfg.h*. Le fichier source *implantcfg.c.c* contient, la déclaration des objets et structures de configuration établie graphiquement. Le fichier d'entête *implantcfg.h* est à insérer manuellement dans les fichiers source en C de notre application par : `#include "implantcfg.h"`.

Sous l'option *Add Files to Project* du menu déroulant de *Project*, on ajoute successivement au projet les fichiers sources C de notre application, le fichier de configuration *implant.cdb*, et le fichier de linkage *implantcfg.cmd*.

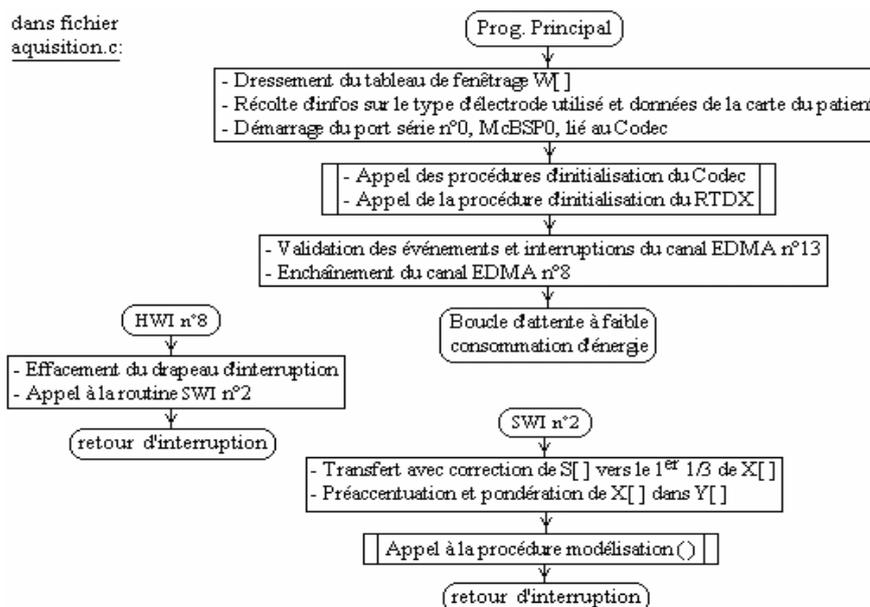
Sous l'option *Scan All Dependencies* du menu déroulant de *Project*, on inclut au projet entre autres, les fichiers d'entête mentionnés dans les fichiers de source C.

La construction du projet, passe par le menu *Project\Rebuild All* qui compile, assemble, et lie tous les fichiers du projet en un seul fichier exécutable *implant.out*.

6.3.1.2. Fichiers-source :

a) *Aquisition.c* :

Ce fichier renferme le programme principal, une fonction de paquets, et deux routines d'interruptions, l'une hardware et l'autre software, voir organigramme de la figure 6-38.



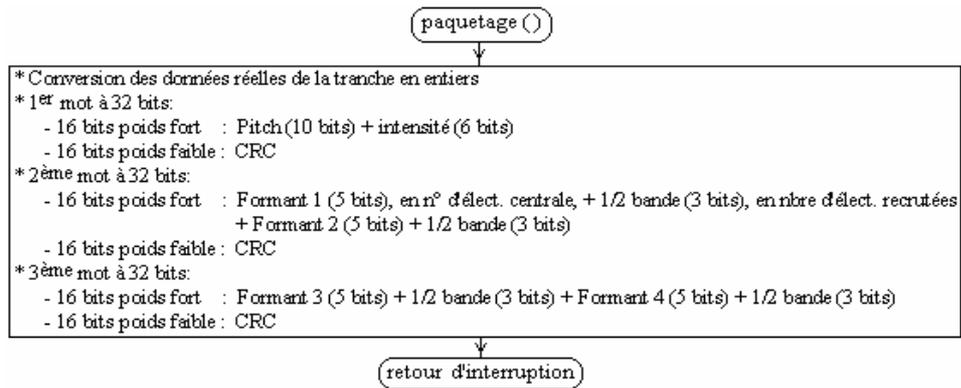


Figure 6-38: Organigrammes du programme principal et ses routines d'interruptions

Le programme principal commence par l'initialisation des différentes constantes, drapeaux (ou indicateurs), et variables globales. Il remplit le tableau W[], avec 480 valeurs réelles propres au fenêtrage de type Hamming.

Il collecte ensuite les informations relatives au type d'électrode sélectionné, en plus des données propres au mal-entendant à partir de sa carte de patient, voir plus loin le dressement d'une telle carte. Ces différentes informations lui proviennent d'une fenêtre de sélection en Visual Basic. Le programme démarre alors le canal de réception du port série McBSP0. Ce dernier devient donc prêt à recevoir les échantillons convertis en provenance du Codec. Le programme appelle successivement les fonctions d'initialisation du Codec (six fonctions imbriquées), et celle d'initialisation d'un canal RTDX en sortie (au nom d'objet 'voieRTDX', voir fichier de configuration). Cette dernière fonction figure dans le fichier RTDX.c.

Après validation des événements et interruptions propres au canal EDMA n°13 ainsi que la validation d'enchaînement du canal EDMA n°8, Le programme rentre dans une boucle d'attente de type IDLE à faible consommation d'énergie. Le programme n'en ressort que pour répondre aux différents appels d'interruptions.

Dans notre cas, les numéros attribués aux différentes interruptions évoquent leurs priorités. De ce fait, la routine hardware HWI n°8 est prioritaire devant HWI n°15, et la routine software SWI n°1 est prioritaire devant SWI n°2. Une façon d'abaisser la priorité d'une routine d'interruption hardware lorsque celle-ci entame une phase longue de son travail, est élucidée dans le cas de la routine HWI n°8. Cette routine hardware assure à un niveau de priorité élevé, un travail court mais urgent (effacement du drapeau d'interruption). Arrivant à la phase longue du travail, elle cède la main à la routine software SWI n°2 de priorité beaucoup plus faible.

Une fois le programme principal est en attente dans la boucle idle, la saisie des échantillons de la parole se fait alors automatiquement selon un plan prescrit et exécuté par les deux canaux EDMA précédemment validés, voir fichier de configuration. Le plan suivi, se déroule de la manière suivante, figure 6-39 :

- Les échantillons de la parole, convertis par le Codec à la fréquence de 8 kHz et véhiculés via le port série McBSP0, sont rangés par le canal EDMA n°13 dans le tableau S[]. Ce tableau présente une taille égale à 160 échantillons. Le canal EDMA n°13 est auto-initialisé, de façon à ce qu'un rafraîchissement continu du tableau s'accomplisse après chaque remplissage complet de ce dernier.
- Un tableau X[], de taille trois fois plus grande que celle de S[], est dressé en vue d'accueillir les échantillons de la tranche courante de la parole en plus des deux tranches précédentes. A chaque remplissage du tableau S[], une mise à jour du tableau X[] est entamée pour assurer le recouvrement des échantillons. Elle consiste donc à faire reculer les anciennes tranches d'échantillons vers sa fin en gardant son début pour la récente

tranche. Ceci est effectué par le canal EDMA enchaîné n°8 qui se charge de transférer en bloc les anciennes tranches. Une fois accompli, une interruption hardware (Hwi n°8) suivie d'une interruption software (Swi n°2), se voient destinées au CPU afin d'exécuter le transfert de la tranche récente tout en masquant le bit LSB de chaque échantillon. Ce bit étant réservé au contrôle du Codec.

- Après le transfert corrigé de la tranche récente des échantillons dans X[], une préaccentuation et une pondération de tout le tableau X[] sont réalisées au sein d'un tableau Y[] de même taille. La préaccentuation (pre-emphasis) se fait selon la norme GSM, figure 5-2, alors que la pondération n'est autre que l'opération de fenêtrage de Hamming à l'aide du tableau W[].

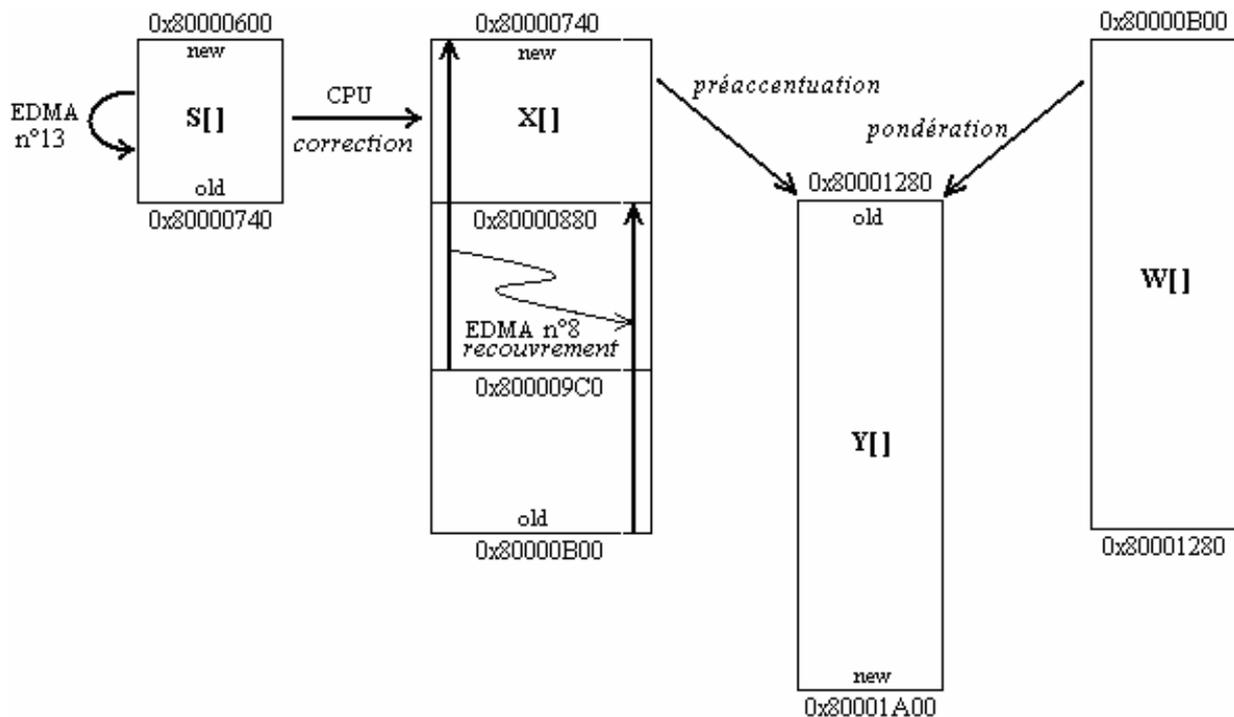


Figure 6-39: Tableaux d'acquisition et de prétraitement des échantillons

Ainsi, à chaque rafraîchissement du tableau X[] par une récente tranche de 160 échantillons de la parole et une fois le tableau Y[] complètement prétraité, le signal est prêt à être paramétré pour en tirer les traits les plus pertinents tels que les coefficients LPC, etc...

Pour cela, le programme appelle la procédure « modélisation() » figurant dans le fichier Analyse.c pour le traitement du signal, voir organigrammes de la figure 5-4.

b) Analyse.c :

Inclut deux procédures, modélisation() et Laguerre(), qui se basent chacune sur un algorithme. modélisation() fait appel à l'algorithme de Schur pour extraire la valeur RMS et les coefficients LPC des segments de la parole, tandis que Laguerre() fait appel à l'algorithme de Laguerre pour extraire les formants et leurs bandes passantes, figures 5-4 et 5-6.

c) Trans_Rcpt.c :

Dans ce fichier, ne figurent que deux routines d'interruptions, l'une hardware relative à la réception sur port série McBSP1, et l'autre software relative au démarrage des périphériques

internes tels que : les deux timers, les canaux de transmission et réception du port série McBSP1, et le canal EDMA n°14, figure 6-40.

L'attribution des tâches aux périphériques internes, responsables de la gestion des deux cartes externes, s'effectue au niveau du fichier de configuration de façon à assurer :

- Une cadence des données transmises, rythmée à $f_0 = 4800$ Hz par le Timer0.
- Une génération de la mi-porteuse FM, aux alentours de 1 MHz, par le Timer1.
- Une transmission et réception de données via les canaux correspondants du port McBSP1.

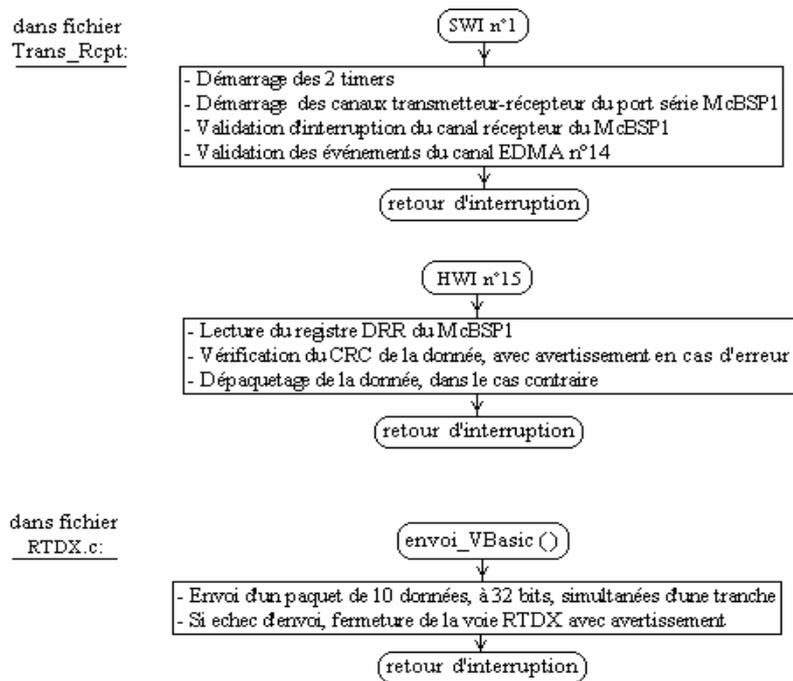


Figure 6-40: Organigrammes propres à deux fichiers source

Il faut rappeler qu'un port série du DSP exige, pour son fonctionnement comme transmetteur ou récepteur, deux signaux de synchronisation : horloge CLK(X/R) et trame FS(X/R). Ainsi notre choix des signaux pour le port McBSP1, est indiqué par le schéma de la figure 6-41. Les liaisons indiquées en couleur bleue, sont établies par programmation lors de la configuration du McBSP1. La liaison en couleur rouge est assurée extérieurement par câblage.

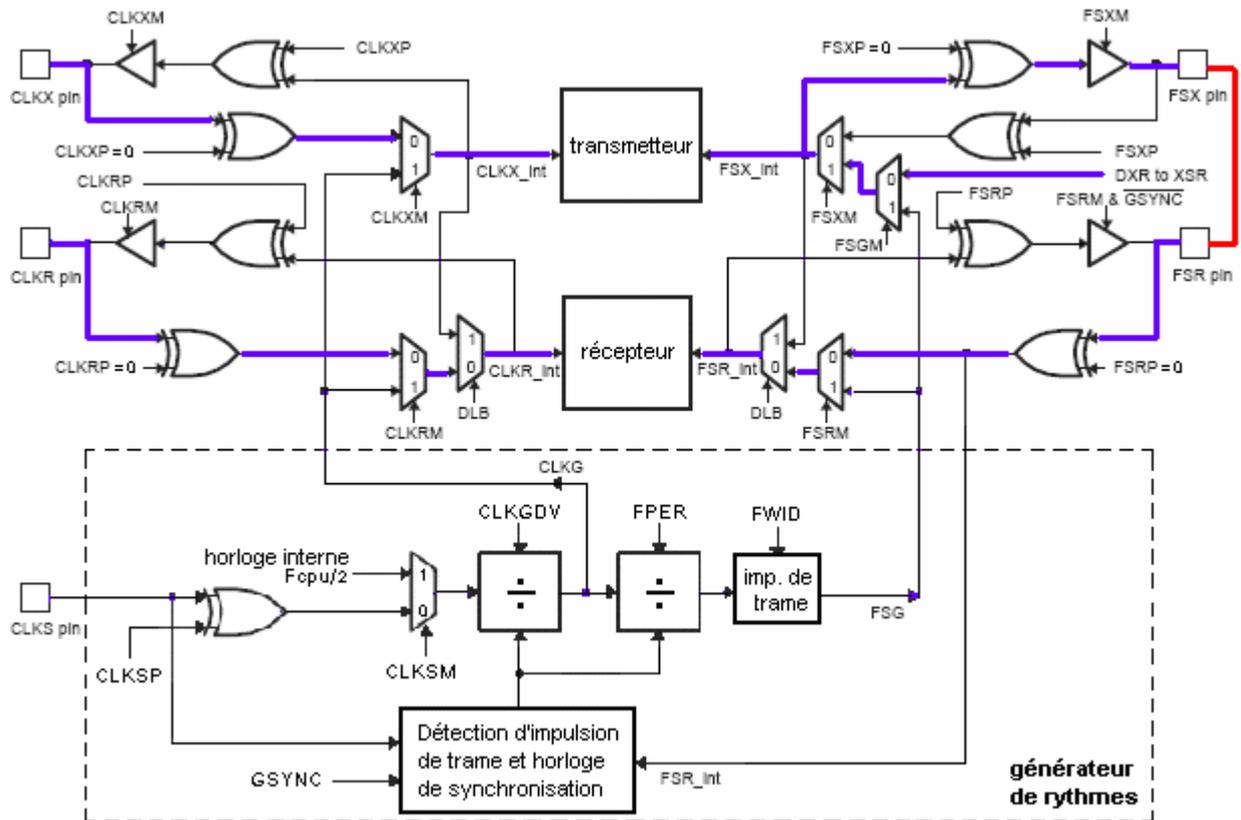


Figure 6-41: Choix des signaux de synchronisation du port McBSP1

Nous rappelons que les paramètres extraits de la parole et transmis vers l'implant sont : le pitch, l'intensité, les couples formant-bande. Ces derniers couples ont été avantageusement remplacés par un autre jeu de couples : numéro d'électrode centrale du formant et nombre d'électrodes recrutées par demi-bande. Ceci a pour conséquence, de maintenir un débit de 4800 bps avec une paquetsation adéquate des paramètres, qui sera discuté dans le paragraphe suivant.

Dans ce contexte, la formule de Hartmann [38] exprimée en 1-1, permet de déduire le lieu de résonance en mm sur la membrane basilaire (ou lieu d'accord des fibres nerveuses associées), en fonction de la fréquence acoustique (ou de stimulation électrique). De cette même formule, on peut déduire l'étendue de la membrane basilaire relative à une gamme fréquentielle donnée. Ainsi, connaissant la distance d de séparation entre électrodes d'un implant cochléaire commercial, on peut convertir une demi-bande fréquentielle D_B d'un formant en un nombre d'électrodes pouvant couvrir cette même plage : $N_{elec} = D_B / d$.

Paquetsation des paramètres :

D'après la théorie des télécommunications, la transmission d'une source d'information sur un canal doit mettre en œuvre un codage de source (compressif) et un codage de canal (protecteur) efficaces. Dans notre cas, le codage de source est le codage autorégressif (LPC), alors que le codage de canal adopté est un CRC à 16 bits selon la norme CCITT n°41, voir la fin du paragraphe 6-1-2, [40, 82].

La transmission de la parole paquetsation opère à 4800 bps avec un codage des paramètres toutes les 20 ms. Un total de 96 bits sont donc disponibles à chaque trame pour encoder les paramètres du signal. Ces bits sont alloués aux différents paramètres, comme indiqué en table 6-5.

Paramètre	Nombre de bits	Valeur limite atteinte
Pitch	10	$F_{0 \max} = 1024 \text{ Hz}$
Intensité	6	64 pas

4 électrodes centrales (Formants)	4 x 5	32 électrodes / Formant
4 groupes d'électrodes (demi-Bandes)	4 x 3	16 électrodes / Bande
3 x CRC	3 x 16	50% redondance
Total bits / 20 ms	96	4800 bps

Table 6-5: Allocation des bits pour notre vocodeur LPC

La répartition des différents paramètres paquetsés sur des mots à 32 bits est illustrée par la figure 6-42.

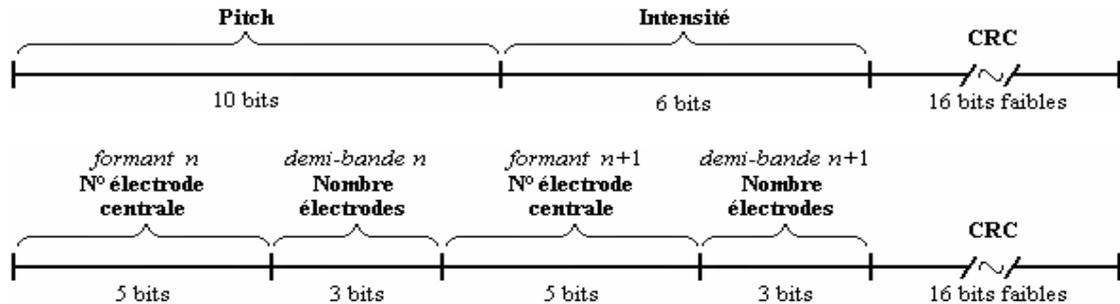


Figure 6-42: Paquetsésation des paramètres d'une tranche de 20 ms de parole

En fin d'organigramme de la procédure modélisation(), figure 5-4, deux appels de procédures sont effectués :

- l'un à la procédure paquetage(), pour la paquetsésation des paramètres extraits.
- l'autre à la procédure envoi_VBasic(), pour l'envoi des paramètres vers les fenêtres d'affichage développées sous Visual Basic.

Les tranches de la parole paquetsésées et protégées, sont amenées automatiquement au port série de transmission du McBSP1, via le canal EDMA n°14. Une fois dans la carte externe de transmission, les paquets de données sont modulés en MSK au rythme de l'horloge 4800 bps, puis modulent en FM une porteuse impulsionnelle de 1MHz, pour être amplifiés ensuite tout en changeant de porteuse contre une sinusoïdale de 2 MHz, voir paragraphe précédent.

Captés au niveau de la carte de réception, le signal subit des traitements inverses, à savoir, une amplification HF suivie d'une démodulation FM, une mise en forme du signal détecté, et une démodulation MSK pour en extraire données et horloge.

De retour à la carte DSP, le canal de réception du port série McBSP1 informe le DSP par interruption HWI n°15, de la présence de données à 32 bits dans son registre DRR. Le DSP lit les données reçues, vérifie leurs CRC, puis les dépaquetsés. Les paramètres ainsi récupérés sont démultiplés et orientés afin de contrôler les courants biphasiques de stimulation au niveau des électrodes appropriées.

Un autre chemin pour les paramètres, est aussi prévu : c'est celui de leur visualisation par le biais de capture HF. Dans notre cas, une capture en temps réel par liaison RTDX est considérée.

Dix données dépaquetsésées sont ainsi envoyées aux fenêtres d'affichage, conçues sous Visual Basic. Il s'agit du signal de pitch et d'intensité représentés par des diagrammes temporels, ainsi que du lieu de stimulation des électrodes dans la cochlée, sous forme d'électrodogrammes.

d) Fichiers en V.Basic :

Le visual basic est un système de programmation, qui se compose d'un environnement de développement élaboré, avec une multitude d'objets, d'assistants de contrôles, et une vaste collection d'instructions de programmation. Notre logiciel de démonstration développé, est appelé Sim-CI (Cochlear Implant SIMulator). Il s'agit d'une interface conviviale qui présente deux fenêtres graphiques, relatives à la saisie de configuration de l'implant et visualisation de ses paramètres de stimulation cochléaire. Deux fichiers sources sont utilisés :

Le fichier writertdx.bas utilise un canal RTDX d'écriture de données, nommé «entreeRTDX». Il assure l'envoi des données suivantes vers le DSP, à travers une boucle de scrutation, figure 6-43 :

- phase de fonctionnement : normal ou test.
- nature de la source audio : enregistrement ou par microphone.
- type d'électrode employé : Nucleus 22 ou 24, Clarion, ou Med-El.
- mode de stimulation : bipolaire ou monopolaire.
- Carte du patient : seuils d'audition MCL et THR, et décalage éventuel d'insertion de la multi-électrodes.

Le fichier readrtdx.bas utilise un canal RTDX de lecture de données, nommé «sortie RTDX». Il assure, à travers une boucle de scrutation, la visualisation graphique des paramètres de stimulation cochléaire envoyés par le DSP, figure 6-44. Il s'agit de deux chronogrammes propres aux pitch et intensité, et d'un électrodogramme propre aux champs d'électrodes activées selon les formants (fréquences et bandes) disponibles. Un petit oscilloscope est prévu à gauche des fenêtres de visualisation précédentes. Il permet de montrer la forme instantanée de l'impulsion biphasique utilisée dans la stimulation de l'implant. L'impulsion est modulée en temps réel, en intensité et en fréquence (du pitch), selon les données en cours.

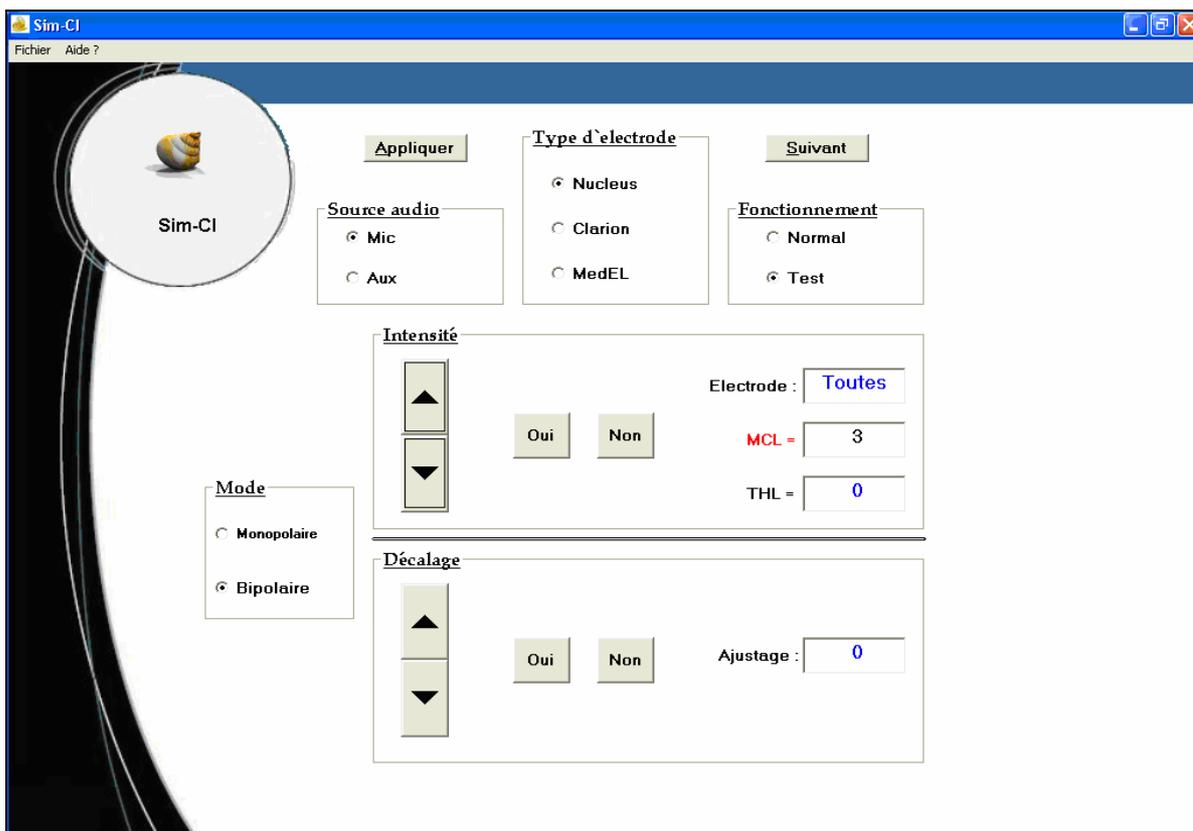


Figure 6-43: Fenêtre de configuration

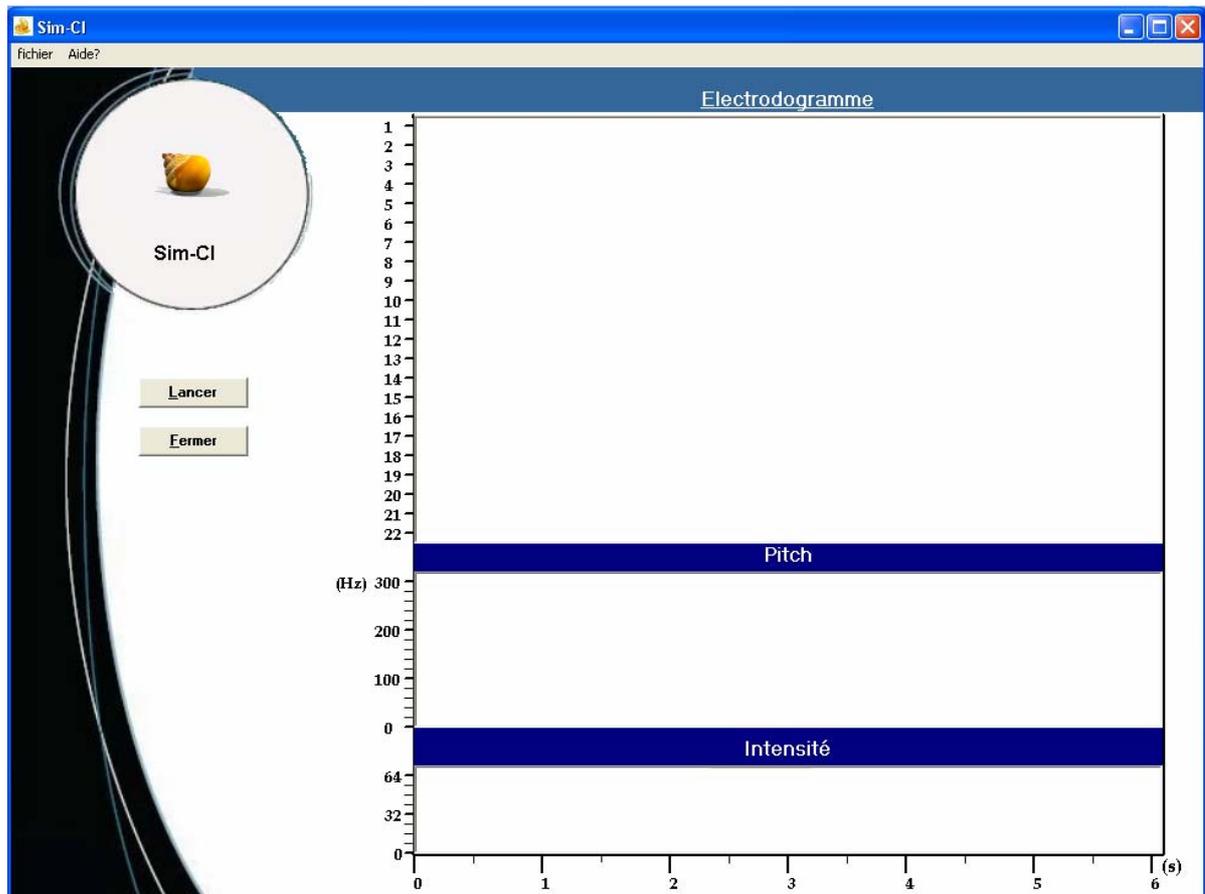


Figure 6-44: Fenêtre d'acquisition

6.3.1.3. Fichiers de mesure :

a) Tracé du spectrogramme à formants :

Afin de valider notre modèle du vocodeur à formants, qui est le noyau de notre stratégie à extraction de paramètres de la parole, on représente graphiquement le spectre à formants pour le comparer à celui d'un sonagramme correspondant, comme mentionné au chapitre 4. Le type de graphisme utilisé est une constellation. Des diagrammes temporels du pitch et intensité, seront également tracés. On réalise pour cela un nouveau projet avec l'IDE CCS, sous le nom « parole.pjt ». Parmi les fichiers à ajouter au projet, figure le fichier source parole.c. Ce fichier est identique au fichier analyse.c, sauf qu'il traite les échantillons de la parole stockés dans un fichier audio, et non en provenance directe du codec.

b) Routines du BER et diagramme en œil :

Il s'agit d'un autre programme de mesure. C'est le programme BER_DiagOeil.c, muni de ses fichiers d'entête oeil.h et gauss32.h, et les fichiers associés gauss32.c, filtre2ord.c, et pseudo.dat. A partir de ces fichiers sources, est conçu le projet BER_DiagOeil.pjt.

Le programme BER_DiagOeil.c, permet la mesure du taux d'erreur de bit ou BER (Bit Error Rate) ainsi que l'affichage d'un diagramme en œil. Le diagramme en œil est relatif à la sortie du filtre de transmission et est dessiné à partir de sa réponse impulsionnelle générée à la cadence des données binaires aléatoires.

Dans le fichier BER_DiagOeil.c, le programme principal invite à faire entrer la durée de test 'temps', puis poste une interruption software. La routine soft TransmissionSWI() utilise avantageusement un tableau de taille 'TABsize'=7714 de données aléatoires à 32 bits, sauvegardées dans le fichier « gauss32.c ». Le GPA matérialisé dans le GAL ne sert donc pas grand-chose devant la capacité de ce tableau.

La routine software commence par calculer la durée nécessaire à la transmission du tableau: $T_{table} = 32 \text{ TABsize} / \text{Clock}$ au rythme de l'horloge 'Clock'=4800 bps. De même, est évalué le nombre de fois à utiliser ce tableau pour pouvoir couvrir la durée de test indiquée :

$n = \text{temps} / (T_{table} + T_{delai})$. Dans cette évaluation, est pris en compte le temps de transit que met un mot à parcourir le trajet en transmission puis en réception à travers le port série, soit : $T_{delai} = 2 \times 32 / \text{Clock} = 13.3 \text{ ms}$.

Ainsi après transmission de chaque donnée, une attente jusqu'à sa réception est effectuée par comptage sur une durée de 13.3 ms avant d'entamer la transmission de la prochaine donnée.

Ce type de synchronisation entre transmission et réception est nécessaire dans l'envoi des données aléatoires et non dans l'envoi des données utiles munies de CRC. L'inconvénient de la synchronisation par attente, est une perte de temps qui se traduit par une diminution du nombre total de bits de test, évalué théoriquement à : $\text{temps} \times \text{Clock}$.

Cependant, un autre temps mort a été constaté en pratique et qui est mesuré à ' T_{cal} ' = 31.57 sec lors d'une seule transmission du tableau « gauss32.c ». Le nombre total de bits à transmettre est déterminé par : $\text{bitMax} = (\text{temps} - n \text{ Tcal}) \text{ Clock}$.

La routine passe ensuite au démarrage des périphériques liés à la transmission-réception, à savoir, les deux timers (horloge et mi-porteuse), et les deux canaux du port série McBSP1 (transmission et réception). Ces périphériques sont configurés graphiquement dans CCS.

Après validation des interruptions du canal de réception de McBSP1, la routine procède à la transmission du tableau « gauss32.c » au nombre de n fois.

Durant ces envois, la routine est régulièrement interrompue par la réception de données au port McBSP1. Chaque fois dans la routine hardware de réception ReceptionHWI(), le DSP compare la donnée reçue avec celle transmise et dénombre les bits erronés éventuels.

A la fin de transmission, le programme affiche le nombre total de bits erronés 'bitCnt' rencontrés durant le temps de test imposé, ainsi que le taux d'erreur de bit correspondant :

$\text{BER} = \text{bitCnt} / \text{bitMax}$.

Une fois terminée la mesure du BER, la routine TransmissionSWI inhibe l'interruption du canal de réception du port McBSP1 puis se sert de données aléatoires pour mesurer les performances de la liaison par le biais du diagramme en œil. Dans ce contexte, est exploité le fichier pseudo.dat, regroupant des données binaires pseudo-aléatoires. Un point de sondage est associé à ce fichier pour rafraîchir le graphe du diagramme en œil, en mode animation.

Le diagramme en œil montre toutes les trajectoires possibles des ondes en bande de base, des données du transmetteur après filtrage, en synchronisation avec la période des symboles.

C'est utile car cela montre rapidement l'efficacité du circuit de transmission conçu et constitue ainsi un puissant outil dans le développement des systèmes de transmission sans fil.

La représentation graphique du diagramme en œil est prise en charge entièrement par l'environnement du CCS. Pour cela, on constate d'abord que la réponse impulsionnelle du filtre de transmission, en figure 6-11, s'étend sur une même durée que celle d'un état minimal du signal MSK d'entrée, soit $T_b = T_H/2 = 0.1 \text{ ms}$ pour une fréquence d'horloge $1/T_H = 4800 \text{ Hz}$. Dans notre cas, la réponse impulsionnelle est échantillonnée en 48 valeurs dans le tableau « TableSinusAmorti[] » du fichier filtre2ord.c.

Une ligne à retard, « transdonnee.delai », constituée d'un paquet de longueur 48 bits, est utilisée pour retenir les 48 échantillons de la réponse impulsionnelle. Ainsi, chaque bit reçu consécutivement par les sous programmes LectureDonnee() et PointConstellation(), est traduit en un point normalisé (± 0.5) de la constellation. La réponse impulsionnelle du filtre de transmission est pondérée par ce point puis superposée aux réponses précédentes éventuelles, mémorisées dans la ligne à retard. La ligne à retard est gérée par le sous programme BufferFiltre(). Les bits aléatoires d'entrée servent d'indices pour une représentation en constellation des symboles binaires, figure 6-45. Il s'agit d'une pseudo constellation, du

moment que la différenciation entre les bits ne se fait qu'au moyen de leurs amplitudes et non par le temps de leur transfert.

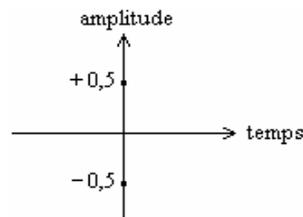


Figure 6-45: Points de constellation des symboles binaires

6.3.2. Espaces de travail des projets de mesure :

6.3.2.1. Espace de travail du projet *Parole.pjt* :

Le programme mis au point est lié en mode lecture, à un fichier extérieur de donnée. On l'associe par conséquent, à un point de sondage (*Probe point*) et un point d'arrêt (*Break point*). Les points de sondage sont utilisés pour :

- Transférer les données d'entrée de ou vers un fichier sur PC, à partir ou vers un buffer de la carte DSP.
- Rafraîchir une fenêtre, tel un graphe, avec les données.

Généralement, on associe un point d'arrêt au point de sondage afin de mettre à jour toutes les fenêtres ouvertes lorsque le point de sondage est atteint. L'activation de ce dernier, commence d'abord par son insertion en marge d'une ligne d'instruction choisie du programme. A partir du menu *File \ File I/O*, on sélectionne le fichier d'entrée et/ou de sortie dans la boîte de dialogue affichée. Dans notre cas, la lecture d'un fichier audio se fait selon la procédure classique suivie par un programme ANSI-C, à savoir, la déclaration de variable de type FILE avec utilisation des mots-clés propres à la librairie *stdio.h* (comme *fopen*, *getc*, et *fclose*). La représentation graphique des résultats d'analyse est prise en charge entièrement par l'environnement du CCS, qui peut afficher une variété de graphes. On cite par exemple, les courbes de réponse temporelle ou fréquentielle, constellation de points, diagramme en œil, et représentation d'image. Le choix du type de graphisme à employer, se fait dans le menu *View\Graph*. Dans notre cas, les types choisis sont, la représentation temporelle (*Time\Frequency*) pour les signaux RMS et Pitch de la parole, et la représentation en constellation pour les formants et leurs bandes. Le rafraîchissement des fenêtres graphiques ainsi obtenues, s'effectue simplement en cliquant sur *Clear Display* depuis leurs menus déroulants. Le point de sondage en conjonction avec un point d'arrêt, est utilisé pour arrêter temporairement le DSP afin d'assurer simultanément, la mise à jour des graphes et le transfert des données du fichier audio. L'utilisation de la commande d'animation permet de reprendre automatiquement l'exécution après que le point d'arrêt soit atteint.

Note: L'application ne peut pas s'exécuter en temps réel si les points de sondage sont utilisés. A ce stade de développement, on teste l'algorithme. L'analyse du comportement en temps réel se fait par d'autres moyens du CCS, comme RTDX.

6.3.2.2. Espace de travail du projet *BER_DiagOeil.pjt* :

Une fois le programme de mesure, mis au point et débogué au moyen d'aides puissantes du debugger de l'environnement CCS, il est plus pratique de lui associer un espace de travail qu'on sauvegarde sous le nom de fichier « *BER_DiagOeil.wks* ». Cet espace de travail sauvegarde la configuration du diagramme en œil et les positions dans le programme, des points associés d'arrêt-sondage. Ainsi, lors de l'exécution sous CCS, on charge l'espace de travail puis le programme exécutable dans le DSP.

Avec l'option *Go Main* du menu déroulant de *Debug*, on se positionne au début du programme principal `main()` de notre fichier source « BER_DiagOeil.c ». Une première exécution jusqu'au point d'arrêt, s'amorce en cliquant sur l'icône *Run*. Cette exécution a pour objectif la transmission-réception de données aléatoires en vue de mesurer le BER de la transmission. Après ceci, et à l'arrêt, une deuxième exécution est à lancer via le bouton *Animate* en sélectionnant le diagramme en œil. Ceci permet de tracer le diagramme en œil à la sortie du filtre de transmission, en réponse aux données aléatoires du fichier « pseudo.dat » associées au point de sondage. Le rafraîchissement du graphe s'effectue simplement en cliquant sur *Refresh* depuis son menu déroulant. Le signal observé à la sortie du filtre de transmission est donné en figure 6-46.

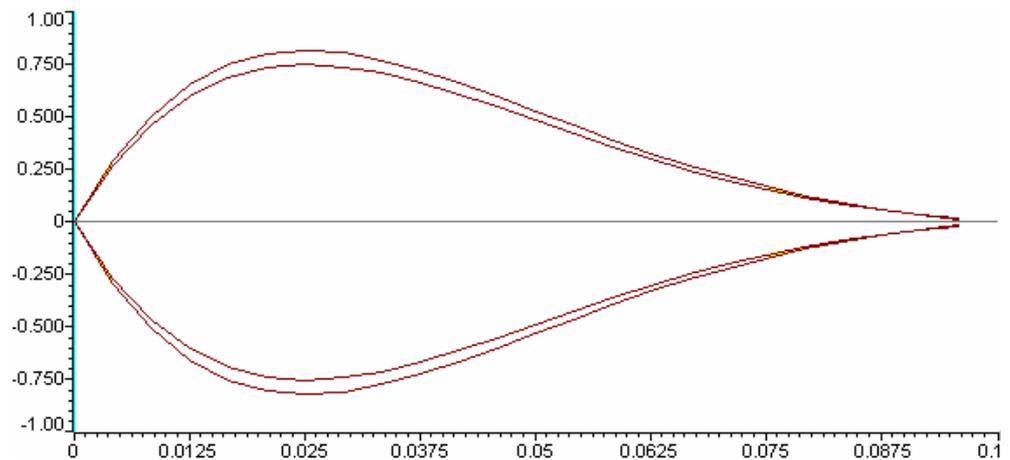


Figure 6-46: Diagramme en œil à la sortie du filtre de transmission

Dans ce diagramme théorique, sans bruit du canal de liaison, l'œil est bien ouverte. Ceci est dû à ce que la réponse impulsionnelle du filtre de transmission s'étend sur une même durée que présente un état minimal du signal MSK d'entrée, $T_S=0.1\text{ms}$. Ce qui a pour effet d'annuler toute distorsion intersymbole (IES), du moment que la réponse impulsionnelle du filtre à un symbole d'entrée (état MSK) s'annule dès qu'arrive le prochain symbole (autre état MSK), ne permettant ainsi, aucun reste de signal superposable à la nouvelle réponse impulsionnelle.

Les seuils de décision, correspondant à la fenêtre d'hystérésis du Trigger de Schmitt de mise en forme à la réception, peuvent être positionnés idéalement aux amplitudes relatives $\pm 0,5$ du diagramme en œil. L'animation du diagramme est arrêtée en cliquant sur l'icône *Halt*, et une deuxième exécution peut être reprise après avoir activé *Debug \ Restart* puis *Go Main*. Il faut noter cependant, qu'une exécution du programme en entier ignorant le point d'arrêt, peut être amorcée en cliquant dès le début sur l'icône *Animate*.

6.4. Schéma fonctionnel final de l'implant cochléaire d'essai

6.4.1. Synoptique de transmission des paramètres d'implant

Notre transmetteur final répond au schéma de principe de la figure 6-47 et au schéma électrique de la figure 6-23. L'acquisition de la parole est assurée par un microphone électrodynamique. Au niveau de la carte à DSP 320C6711, la parole est échantillonnée à la fréquence de 8 kHz par le codec TLC320AD535. Les échantillons numériques sont en complément à 2 dans un format à 16 bits.

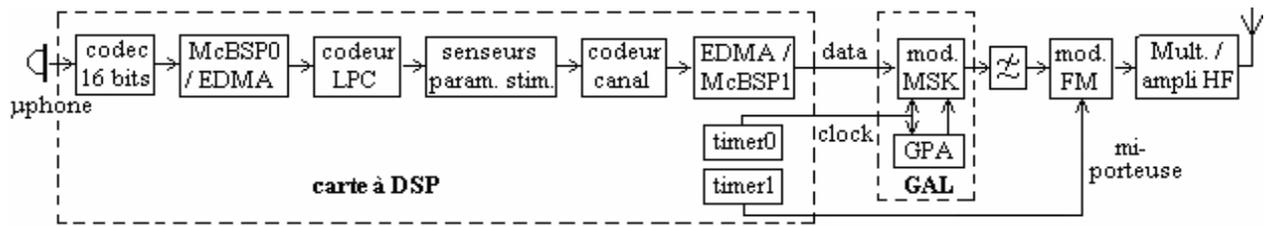


Figure 6-47: Schéma synoptique du circuit de transmission

Le port série McBSP0 interne au DSP, véhicule ces échantillons vers le cache-mémoire du DSP sous le contrôle d'un canal EDMA. Le DSP assure par logiciel les fonctions de traitement tels que le codage LPC, l'extraction des paramètres pertinents de la parole par senseurs, puis protection de ceux-ci par codage du canal. Après paquetage des paramètres avec leurs CRC, un canal EDMA se charge alors de les transmettre périodiquement au circuit extérieur de modulation, via le port série McBSP1.

Le DSP assure de même, à travers ses deux timers internes, le pilotage du circuit extérieur par un signal d'horloge rythmant le débit binaire à 4800 bps, ainsi qu'une mi-porteuse de 1 MHz pour la modulation analogique FM. Viennent ensuite les fonctions de modulation et d'amplification HF accomplies par la carte de transmission.

6.4.2. Synoptique de réception des paramètres d'implant

Le récepteur réciproque au modèle de transmission établi en figure 6-47, se présente sous le schéma synoptique de la figure 6-48.

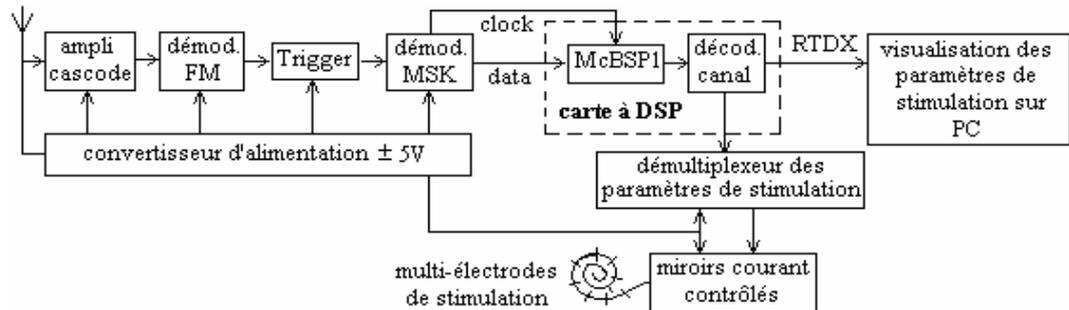


Figure 6-48: Schéma synoptique du circuit récepteur

On y trouve un convertisseur d'alimentation, prévu pour convertir l'énergie HF de la porteuse en source d'alimentation bitension afin de réaliser des courants biphasiques de stimulation. Ces tensions régulées à $\pm 5V$, servent aussi à alimenter les circuits à faible puissance du récepteur. Le circuit d'alimentation est conçu de façon à assurer un autre rôle, celui de circuit limiteur d'ondes FM. Débarrassant par-là, les parasites HF éventuellement superposés sur son enveloppe.

La voie des données est constituée d'une multitude d'étages en cascade. Le premier étage est un ampli HF de type cascode qui, vu sa grande impédance d'entrée, assure une isolation entre les circuits accordés de réception et de démodulation, annulant ainsi leur interaction mutuelle. Un démodulateur FM suivi d'un trigger, permettent de retrouver le signal MSK modulant. Une démodulation MSK non cohérente est utilisée, afin de restituer les données et extraire l'horloge. Le démodulateur MSK peut aussi être vu comme un décodeur de source. Les données et l'horloge sont mises à la disposition du DSP à travers son port série McBSP1.

La première opération à effectuer par le DSP, est la procédure au décodage du canal en vérifiant le CRC accompagnant chaque donnée. Une fois les paramètres dépaquetés, ils sont envoyés par voie RTDX au PC pour leur visualisation en temps réel. Dans certains implants commerciaux, une capture HF des données est utilisée comme moyen de leur visualisation [38].

Au niveau de l'implant cochléaire, le chemin des données est tout autre. Les paramètres de stimulation sont démultiplexés pour être orientés chacun vers sa destination :

- Le pitch pilote la fréquence des impulsions biphasiques de courant.
- La valeur RMS module l'intensité du courant.
- Le courant de stimulation ainsi contrôlé, est reproduit par des miroirs de courant afin d'alimenter les électrodes concernées en mode BP+n. Pour chaque couple formant-bande, l'intervalle d'électrode correspondant est envoyé :
[numéro d'électrode inférieure comme électrode d'arrivée du courant,
et numéro d'électrode supérieure comme électrode de retour].



Figure 6-49: Circuits d'essai de traitement et de transmission de l'implant cochléaire

Conclusion :

Les supports hardware et software de notre réalisation sont décrits dans ce chapitre. Les différents étages constituant les cartes de transmission et de réception, sont présentés et calculés tour à tour, tout en justifiant leurs choix. De même, sont présentés les différents fichiers sources sur lesquels repose le logiciel d'application de notre implant cochléaire d'essai. Des fichiers propres aux mesures et évaluation de la qualité de la transmission-réception des paramètres extraits de la parole, sont aussi discutés.

7.1. Comparaison des traitements LPC-Sonagramme de la parole**7.1.1. Présentation des fichiers audio d'essai**

Deux fichiers enregistrés de la parole sont analysés de deux manières différentes puis les résultats d'analyse sont comparés entre eux.

Il s'agit des fichiers audio, suivants :

a- fichier Parentes.aud : renfermant un seul mot ; « parenthèse ».

b- fichier Lapin.aud : renfermant une seule phrase ; « as tu vu ce fameux lapin ».

Ces fichiers ont été enregistrés avec le logiciel développé Son.pas, du chapitre 4. Ce sont donc des fichiers de données monophoniques, échantillonnées à 11 kHz sous un format à 8 bits.

Comme on peut le constater, les fichiers audio sont choisis à complexité croissante, utilisant différents phonèmes : voyelles, consonnes occlusives, fricatives, etc...

7.1.2. Description et comparaison des spectrogrammes et signaux obtenus

Les deux fichiers sources audio, issus d'un même locuteur, sont utilisés afin de comparer les spectres émanant des deux outils d'analyse employés, à savoir, le sonagramme à large bande (par exécution du logiciel Son.pas sur PC) et le traitement LPC (par exécution du projet Parole.pjt sur carte DSP).

Dans le but de simplifier la phase de comparaison visuelle, trois spectres du même fichier sont réunis sur chaque figure :

- Un spectre (a) est celui généré par le traitement LPC.
- Un spectre (b) est celui généré par le sonagramme à large bande.
- Une superposition (c) des deux spectres précédents.

Le spectre fréquentiel de la parole doit faire apparaître deux informations essentielles : les formants et leurs bandes. Or, l'environnement CCS ne présente que deux possibilités de représentation temporelle; par lignes simples ou par barres. Ces possibilités ne peuvent répondre aux exigences d'un spectre fait de lignes quasi-verticales, réparties sur des abscisses adjacentes en forme de bandes continues ou discontinues, comme le cas du sonagramme à large bande. Afin de se rapprocher au mieux de ce type de représentation spectrale, une visualisation en constellation de points est adoptée pour les formants et leurs bandes. Ainsi, au bout de chaque centaine d'échantillons traitée par LPC, dix points sont tracés de part et d'autre de la fréquence centrale de chaque formant trouvé. Ces points sont répartis uniformément le long de la bande fréquentielle B, autour du formant F, allant donc de $F - B/2$ à $F + B/2$ sur dix points. Le spectre en entier est répété donc 10 fois pour représenter la largeur de bande de chaque formant, de façon discrète sur 10 points. Le spectre prend par conséquent, l'aspect d'un nuage de points qui converge vers celui de lignes contiguës lorsque la bande du formant est faible. Le spectre ressemble alors à celui d'un spectrogramme à large bande. Un autre point est à signaler, à propos de la précision des graduations portées sur l'axe temporel des deux représentations spectrales: les graduations ne sont qu'approximatives. Ceci est dû aux causes suivantes :

- Suppression de l'entête du fichier audio généré par le logiciel Son.pas. Cette entête devient insignifiante lorsque le traitement LPC est entrepris.
- Incrément temporel différent des deux représentations spectrales. LPC opère sur des paquets d'échantillons de longueur non recouverte égale à 100, alors que le sonagramme opère sur des longueurs de 128 échantillons. L'incrément temporel adopté dans chaque spectre respectif est donc : $dt = 100 / 11 \text{ kHz}$ pour LPC, et $dt = 128 / 11 \text{ kHz}$ pour sonagramme.
- Longueur de fichier tronqué approximativement à 70 ko. Ceci afin que le spectre LPC soit généré convenablement par la carte DSP, tout en n'excédant pas la largeur d'une page, une fois imprimé.

Dans les spectres LPC obtenus, chaque petite graduation sur l'axe fréquentiel, représente un intervalle de 166.6 Hz, alors que celle du temps est de 0.1 sec.

7.1.2.1. Analyse du fichier Parenthèse.aud :

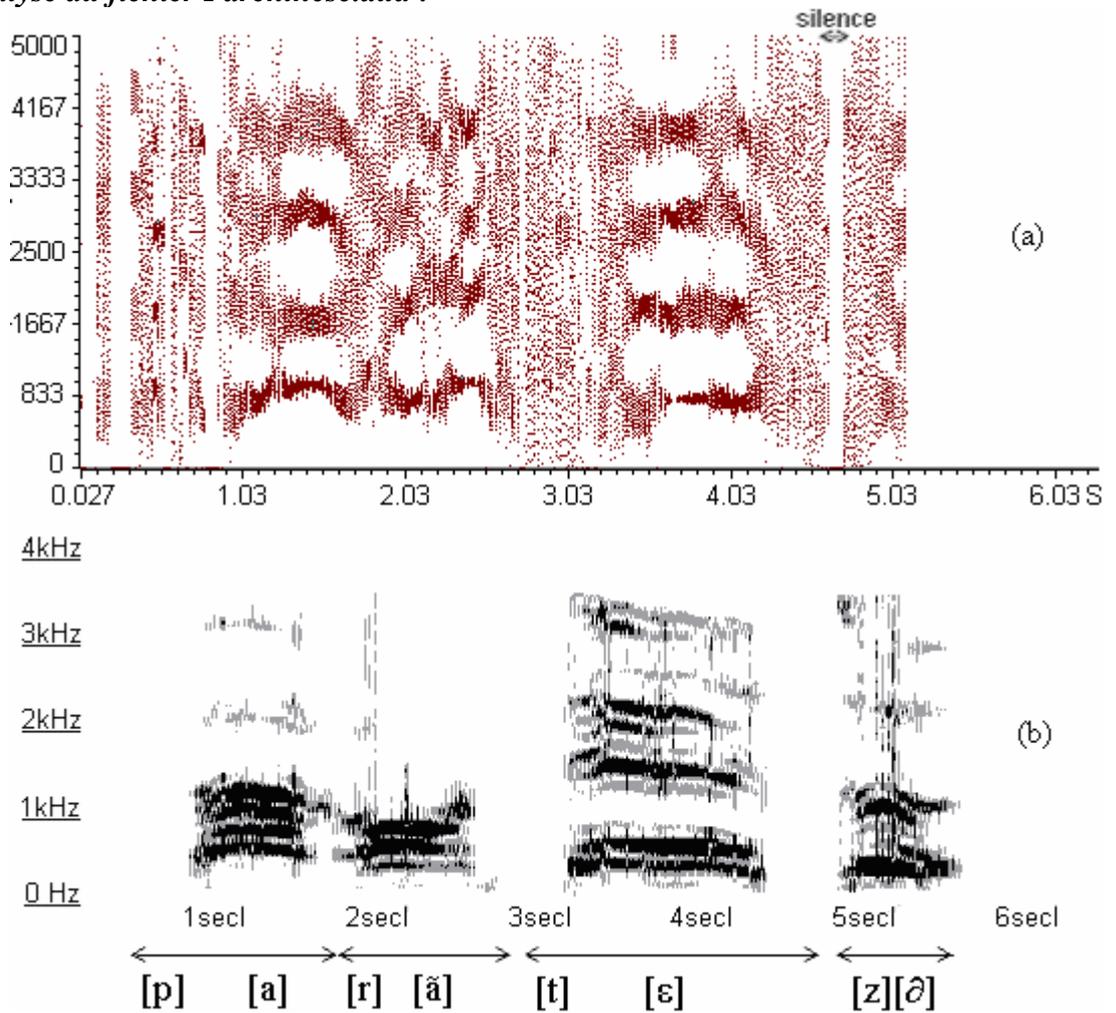


Figure 7-1: Spectre LPC (a) et sonagramme (b), propres au fichier Parenthèse.aud

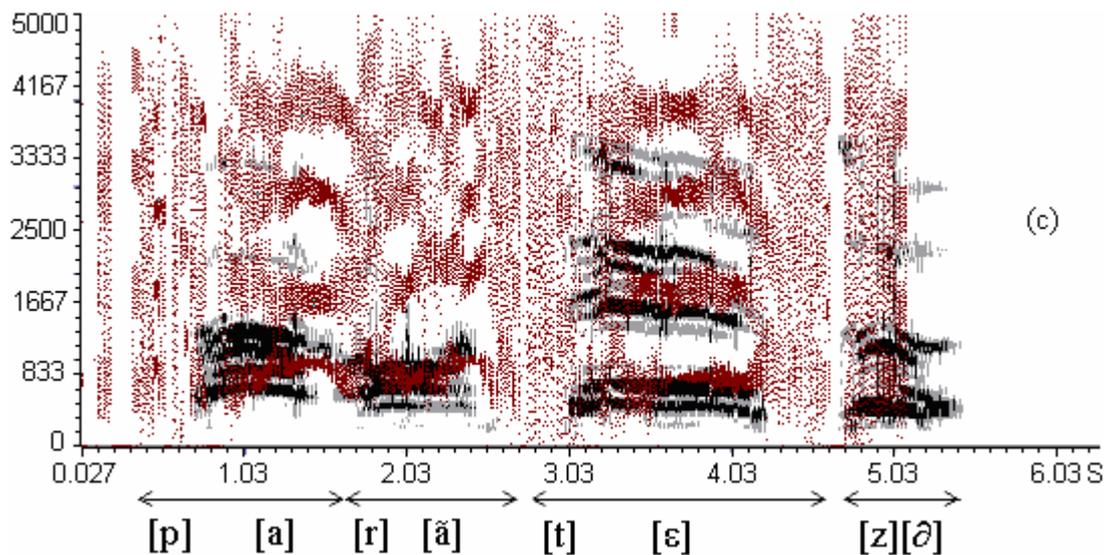


Figure 7-2: Superposition (c) des deux spectres du fichier Parenthèse.aud

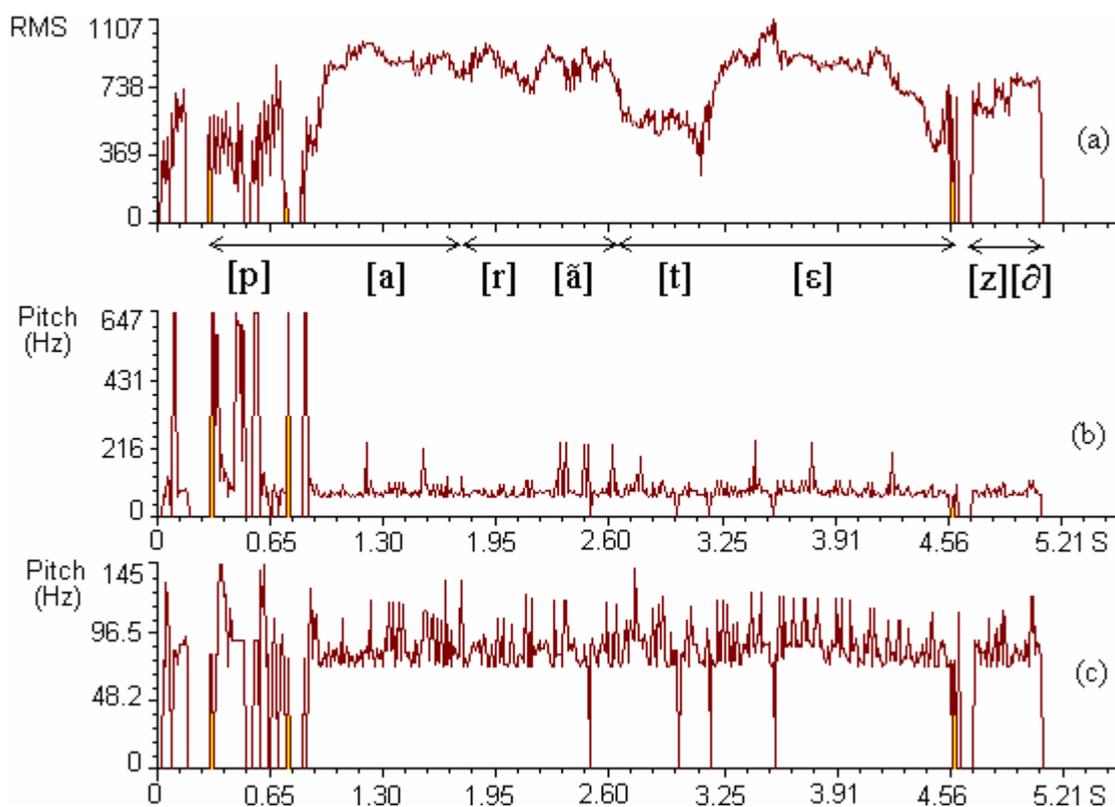


Figure 7-3: Valeur efficace vraie (a), pitch non corrigé (b), et pitch corrigé (c), du fichier Parenthèse.aud

Le fichier renferme le mot « parenthèse » prononcé par un locuteur masculin. Les spectres obtenus à partir de la méthode LPC et celle du sonagramme à large bande, sont donnés en figures 7-1 et 7-2. Les signaux RMS et pitch sont réunis en figure 7-3. La comparaison des spectres, nous recèle les traits de similitude et de complémentarité suivants :

Phonèmes	LPC (en rouge)	Sonagramme (en noir)
[p][a] : occlusive non voisée et voyelle orale.	[p] : aspect diffus le long de F. [a] : quatre formants. mêmes trajets formantiques qu'en sonagramme (forme d'arc, dû au changement de tonalité pour passer au phonème suivant).	[p] : non détecté. [a] : F ₁ englobe celui du LPC. F ₂ et F ₃ , faibles aux débuts des intervalles du LPC. F ₄ absent.
[r][ã] : liquide et voyelle nasale.	[r] : zone diffuse le long de F. [ã] : quatre formants. avec trajets formantiques.	[r] : même zone qu'en LPC. [ã] : même F ₁ qu'en LPC. F ₂ , F ₃ , et F ₄ absents.
[t][ε] : occlusive non voisée et voyelle orale.	[t] : zone diffuse le long de F. [ε] : quatre formants. trajets formantiques à la fin des formants supérieurs, puis diffusion à faible intensité de la voyelle.	[t] : non détectée. [ε] : mêmes F ₁ , F ₂ , et F ₃ qu'en LPC, avec F ₂ plus étalée. diffusion absente à la fin. F ₄ absent.

[z][∂]: fricative voisée et voyelle orale.	[z] : zone diffuse le long de F, précédée d'un silence. [∂] : débuts de quatre formants, (car fichier tronqué).	[z] : zone étroite superposée aux formants de la voyelle. [∂] : quatre formants aux bas des intervalles du LPC.
--	--	--

Table 7-1: Comparaison spectrale de la phrase du fichier Parenthèse.aud

Les consonnes sont représentées par des zones diffuses le long de l'axe fréquentiel du pseudo-spectrogramme LPC. Alors que sur le sonagramme, elles sont :

- soit non détectées comme [p] et [t],
- soit représentées par une zone étroite de formants comme pour [z], ou zone diffuse comme pour [r].

Toutes les voyelles sont représentées par quatre formants dans le spectrogramme LPC, qui coïncident avec ceux détectés par le sonagramme. Les formants ne sont pas tous détectés par le sonagramme: exemple, F₄ absent pour [a], F₁ seul pour [ã], etc....

On constate que les bandes des formants dans le sonagramme sont :

- soit égales à celles du LPC comme dans [ã] et [ε],
- soit sous estimées et n'indiquent que la limite haute des bandes estimées par LPC (comme pour [a]), ou la limite basse (comme pour [∂]),
- soit surestimées et englobent les bandes du LPC comme pour [ε] et [a].

Certaines trajectoires formantiques sont visibles dans les deux spectres. Elles sont dues à des changements de tonalité dans la prononciation de la phrase.

D'une manière générale, Les segments des différents phonèmes rencontrés occupent les mêmes durées d'analyse dans les deux spectres. Cependant, parmi les points forts de l'analyse LPC, on note par exemple, celui des trajectoires formantiques décroissantes, raccordant les voyelles [a] et [ã] des deux premiers phonèmes. Elles sont marquées sur trois formants au moins du spectre LPC, alors qu'il n'est représenté dans le sonagramme qu'une seule trajectoire décroissante pour le premier formant. Le signal RMS du fichier Parenthèse.aud, figure 7-3, présente des maxima en liaison avec les voyelles [a], [ã], et [ε] des phonèmes [p][a], [r][ã], et [t][ε] respectivement. Les voyelles sont les parties les plus énergétiques des phonèmes. On retrouve de même, les valeurs RMS correspondantes aux segments des consonnes occlusives comme [p] et [t], qui ne sont pas détectées dans le sonagramme. Le pitch est détecté en correspondance parfaite avec le signal RMS. Même les segments des occlusives, leur correspondent un pitch. Le pitch non corrigé, figure 7-3, présente deux types d'erreurs :

- pitch trop élevé lié à l'occlusive [p] du premier phonème. La même consonne lui est attribuée d'ailleurs, une grande variation d'énergie ou valeur efficace.
- dédoublement du pitch, atteignant dès-fois les 216 Hz. On constate que cette erreur apparaît en liaison avec les différentes voyelles des phonèmes prononcés.

Une fois corrigé, le pitch de la figure 7-3 accuse une variation moyenne entre 70 Hz et 115 Hz selon l'évolution de la tonalité. La valeur moyenne de pitch du locuteur masculin est donc 90 Hz. Quelques rares pics des occlusives [p] et [t], sont limités à 144 Hz. Valeur pouvant encore être interprétée comme un dédoublement du pitch.

7.1.2.2. Analyse du fichier Lapin.aud :

Ce fichier contient la phrase enregistrée : « as tu vu ce fameux lapin », issue du même locuteur masculin. Ce fichier source est tronqué à 70 ko avant d'être traité par le DSP. Le phonème [p][é] du mot « lapin » se voit donc éliminé de son spectre LPC. Les spectres obtenus, à partir de la méthode LPC et celle du sonagramme à large bande, sont donnés en figures 7-4 et 7-5

alors que les signaux RMS et pitch sont réunis en figure 7-6.

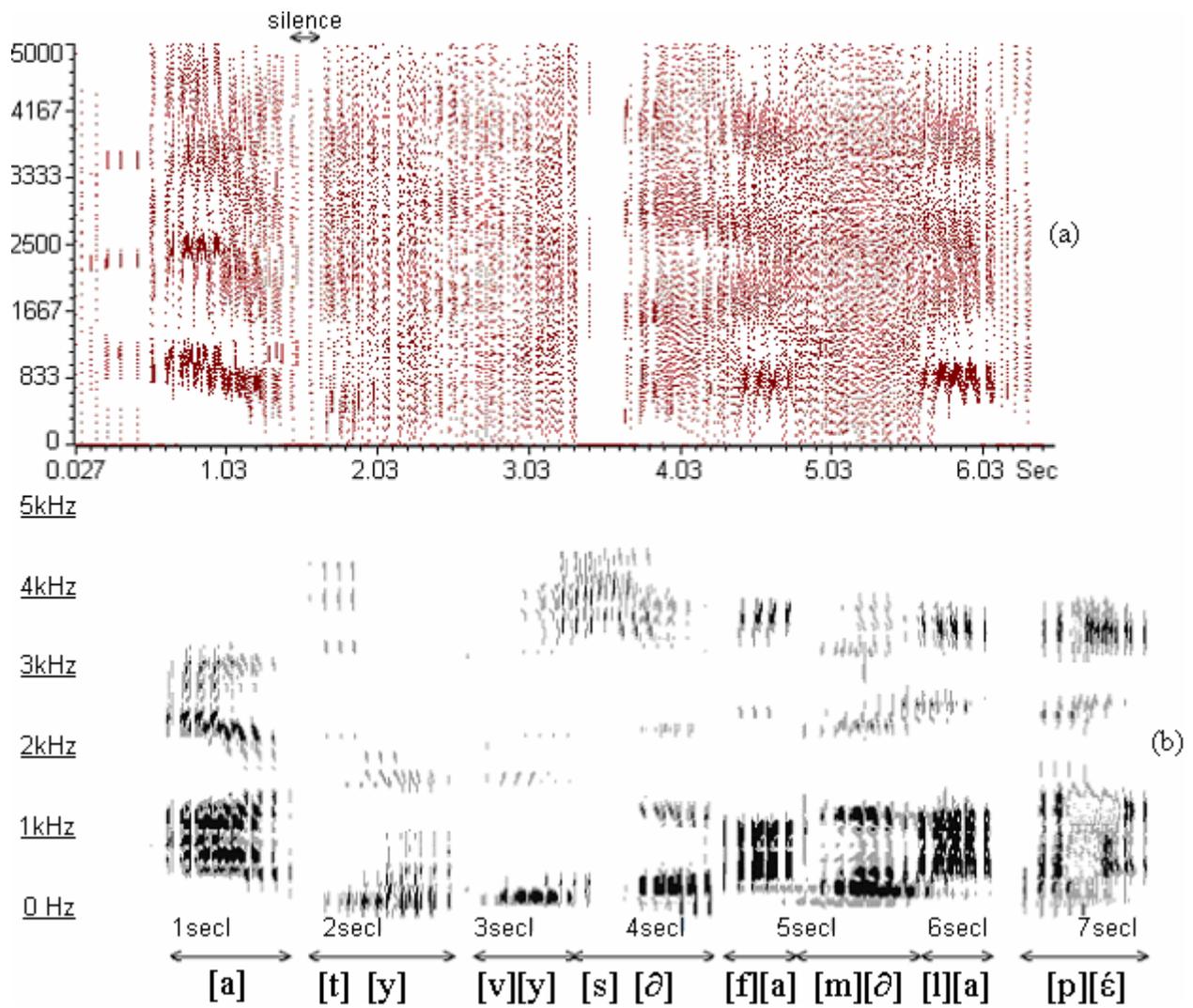


Figure 7-4: Spectre LPC (a) et sonagramme (b) propres au fichier Lapin.aud

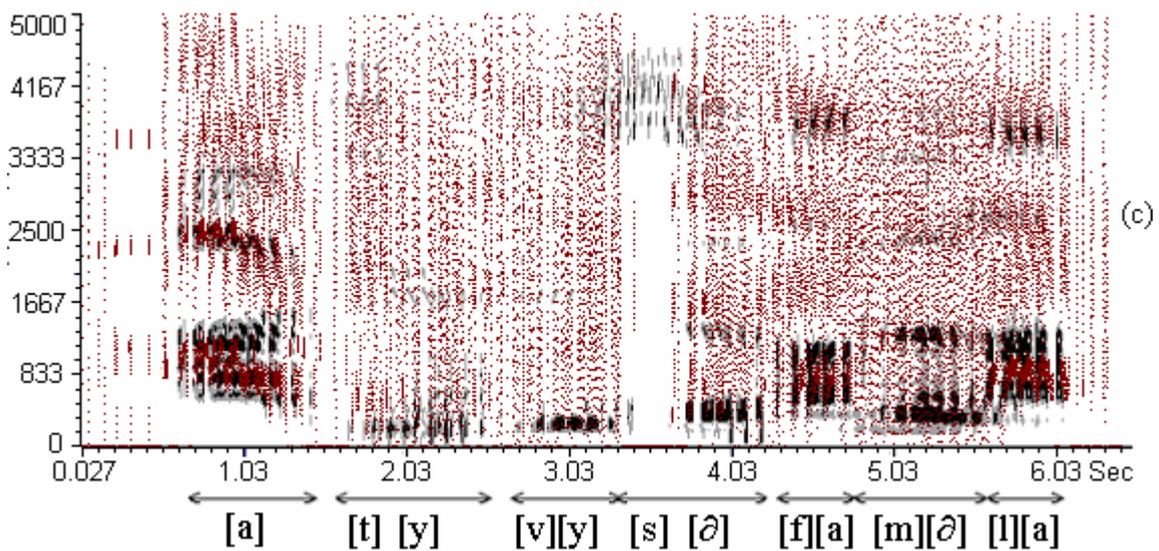


Figure 7-5: Superposition (c) des deux spectres de la phrase du fichier Lapin.aud

La comparaison des spectres, nous révèle les traits de similitude et de complémentarité suivants :

Phonèmes	LPC (en rouge)	Sonagramme (en noir)
[a] : voyelle orale.	quatre formants.	mêmes F ₁ , F ₂ , et F ₃ que LPC, avec F ₃ au début du LPC. F ₄ absent.
[t][y] : occlusive non voisée et voyelle orale.	[t] : zone diffuse le long de F, superposée aux formants de [y]. [y] : trois formants à peine visibles.	[t] : même zone qu'en LPC. [y] : mêmes F ₁ , F ₂ que LPC. F ₃ absent.
[v][y] : fricative voisée et voyelle orale.	[v] : zone diffuse le long de F, superposée aux formants de [y]. [y] : trois formants à peine visibles.	[v] : non détecté. [y] : mêmes F ₁ , F ₂ que LPC. F ₃ absent, (mêmes que [t][y])
[s][∂] : fricative non voisée et voyelle orale.	[s] : n'est pas détecté. [∂] : quatre formants.	[s] : zone diffuse à 4 kHz. [∂] : quatre formants aux bas des intervalles du LPC.
[f][a] : fricative non voisée et voyelle orale.	[f] : zone diffuse le long de F. [a] : quatre formants.	[f] : non détecté. [a] : mêmes F ₁ , F ₃ , et F ₄ qu'en LPC. F ₂ absent.
[m][∂] : consonne nasale et voyelle orale.	[m] : zone diffuse le long de F, et superposé aux formants de E. [∂] : formants masqués par la diffusion des deux consonnes les entourant ([m][∂][l]).	[m] : même zone qu'en LPC [∂] : quatre formants peu masqués par les diffusions, (mêmes que pour [s][∂]).
[l][a] : liquide et voyelle orale.	[l] : zone diffuse le long de F. [a] : quatre formants. (mêmes que pour [f][a]).	[l] : non détecté. [a] : mêmes F ₁ , F ₃ , et F ₄ qu'en LPC. F ₂ absent.

Table 7-2: Comparaison spectrale de la phrase du fichier Lapin.aud

Toutes les consonnes sont représentées par des zones diffuses le long de l'axe fréquentiel du pseudo-spectrogramme LPC, sauf pour [s] qui n'est pas détecté. Alors que sur le sonagramme, elles sont :

- soit non détectées comme [v], [f], et [l].
- soit représentées par une zone étroite de formants comme pour [t] et [s], ou zone diffuse comme pour [m].

Dans le pseudo spectrogramme LPC, les voyelles sont représentées par 4 formants pour [a] et [s][∂], 3 formants peu visibles pour [y], qui coïncident avec ceux détectés par le sonagramme. Les formants ne sont pas tous détectés par le sonagramme: exemple, F₂ absent pour [f][a] et [l][a], F₃ absent pour [y].

On constate que les bandes des formants dans le sonagramme sont :

- soit égales à celles du LPC, comme pour les formants élevés de [f][a], [l][a], et F₂ de [a],
- soit sous estimées et n'indiquent que la limite basse des bandes estimées par LPC, comme pour [s][ø] et F₃ de [a],
- soit surestimées et englobent les bandes du LPC, comme pour F₁ de [a].

On constate que les formants de [ø] dans le spectrogramme LPC, sont plus masqués par la diffusion des deux consonnes l'entourant ([m][ø][l]), que dans le sonagramme. D'un point de vue global, les segments des phonèmes analysés occupent les mêmes durées de traitement dans les deux spectres.

De la figure 7-6, on s'aperçoit que le signal RMS est maximum pour les phonèmes [a], [s][ø], [f][a], [m][ø], et [l][a], liés aux différentes voyelles associées. Par contre, la voyelle [y] des phonèmes [t][y] et [v][y] est moins énergétique que les voyelles précédentes. Les consonnes [t], [s], et [p] sont les moins énergétiques des phonèmes de la phrase.

Une erreur de mesure du pitch est constatée durant l'occlusive [t] et juste avant la voyelle [a] du début de phrase ainsi que du phonème [f][a]. Certaines parties de ces segments sont considérées donc à pitch élevé (correspondant à un ton aigu, comme celui d'un enfant). Le dédoublement du pitch se retrouve fréquemment en liaison avec des voyelles. La correction du pitch, en figure 7-6, met en évidence la même variation moyenne du pitch de ce locuteur, allant de 70 Hz à 115 Hz (valeur moyenne du pitch, estimée à 90). Cependant, certains dépassements vers 144 Hz se manifestent notamment à la fin de la phrase, en correspondance avec la voyelle du phonème [l][a]. On constate que le pitch n'est pas détecté durant la fricative non voisée [s], ce qui a interdit la mesure de ses transitions formantiques élevées.

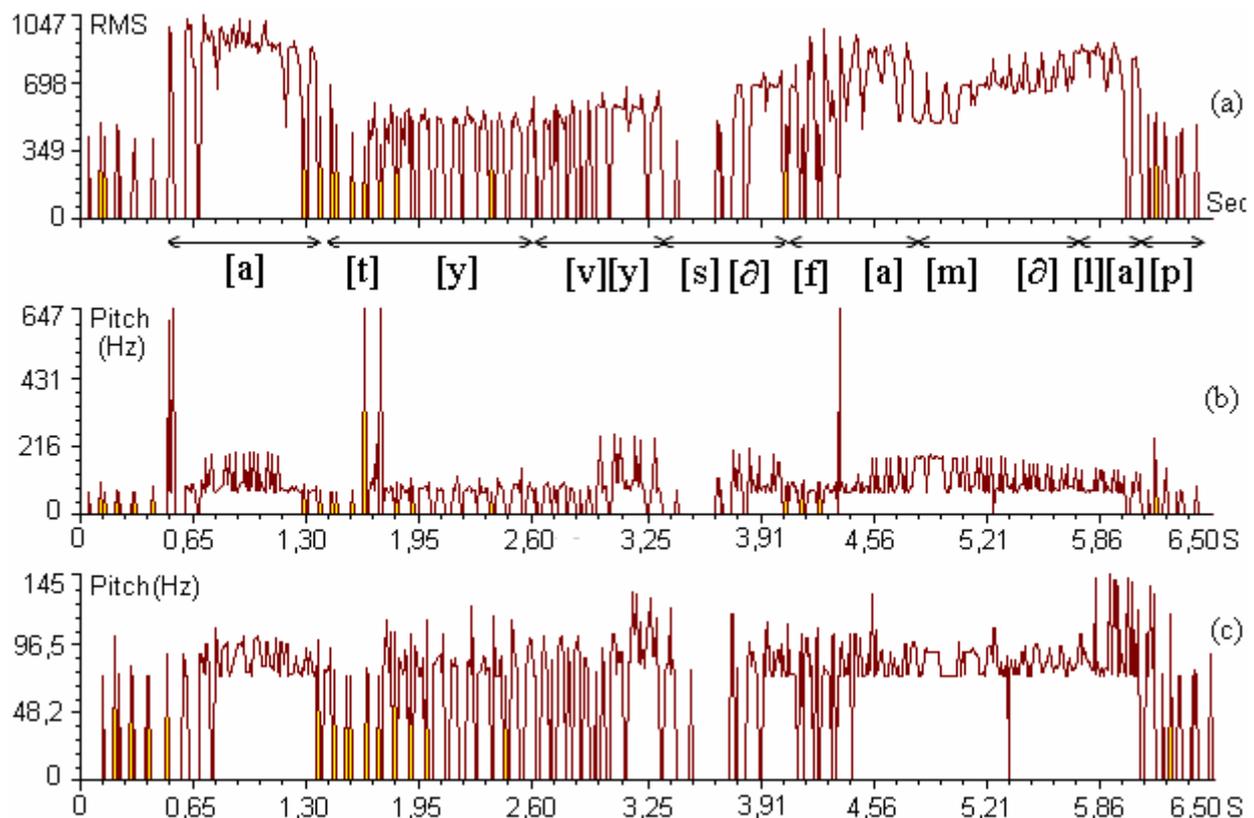


Figure 7-6 : Valeur efficace vraie (a), pitch non corrigé (b) et pitch corrigé (c), du fichier Lapin.aud

7.2. Performance des cartes réalisées

7.2.1. Mesure des performances d'une transmission de données

Parmi les moyens dont on dispose pour évaluer la qualité d'un système de transmission de données, on cite : le calcul du taux d'erreur binaire, l'obtention du diagramme en œil, et la mesure de la gigue de phase. On essaiera dans un premier temps de définir ces notions, puis de les appliquer sur notre système réalisé.

7.2.1.1. Taux d'erreur binaire [17] :

Le calcul du taux d'erreur est basé sur la probabilité P qu'un bruit gaussien de moyenne nulle et de tension efficace σ dépasse l'une des tensions de seuil symétriques $\pm d$ (dans le cas général de symboles M -aires, les valeurs seront dans l'alphabet $\{\pm d, \pm 3d, \dots, \pm(M-1)d\}$). En général, ce calcul peut se ramener à un rapport de puissance entre celle de la porteuse du signal modulé C , et celle du bruit thermique KTB , telle que K est la constante de Boltzman, T , la température équivalente du bruit et B , la bande du bruit.

$$P = 0,5 \operatorname{erfc}\left(\frac{d}{\sigma\sqrt{2}}\right) = 0,5 \operatorname{erfc}\left(\sqrt{\frac{C}{KTB}}\right) \quad 7.1$$

$\operatorname{erfc}(x)$ étant la fonction d'erreur complémentaire qui est dans le cas approché :

$$\operatorname{erfc}(x) = \frac{2}{\sqrt{\pi}} \frac{e^{-x^2}}{x + \sqrt{x^2 + 2}} \quad 7.2$$

La table 7-3 donne l'écart en décibels entre la tension efficace du bruit σ et la tension-crête du niveau d , en fonction du taux d'erreur sur les bits.

taux d'erreur	$20 \log(d / \sigma)$
10^{-1}	2,2
10^{-2}	7,3
10^{-3}	9,8
10^{-4}	11,4
10^{-5}	12,6
10^{-6}	13,5
10^{-7}	14,3
10^{-8}	15
10^{-9}	15,6
10^{-10}	16,1

Table 7-3: Rapport signal sur bruit en fonction du taux d'erreur

Le taux d'erreur de bit ou BER (Bit Error Rate) peut encore être défini, d'une manière plus simple, comme le rapport du nombre de bits erronés sur le nombre total de bits. Pour sa mesure, on utilise un signal test pseudo-aléatoire.

7.2.1.2. Diagramme en œil [47] :

La figure 7-7 illustre les signaux émis et reçus dans le cas d'un codage NRZ avec bruit capté négligeable.

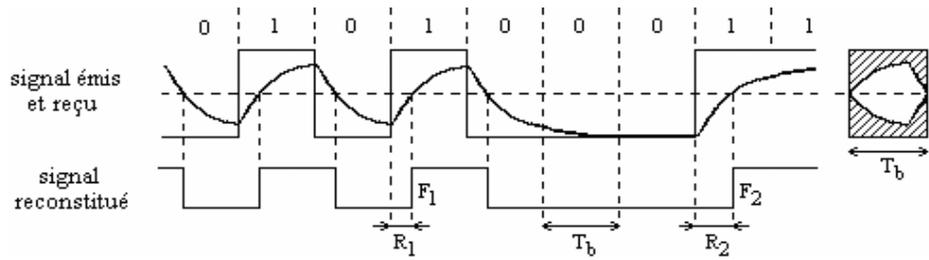


Figure 7-7: Signaux émis, reçu, et reconstitué, avec diagramme en œil du signal reçu

La ligne de transmission a une fonction de transfert de type passe-bas et transforme donc les fronts du signal émis en variations continues (réponse indicielle). Il en résulte un retard R des fronts du signal reconstitué sur ceux du signal émis. Ce retard présente un caractère aléatoire lié à l'information. Par exemple, le front F_2 qui termine une longue suite de « 0 », présente un retard plus important que le front F_1 , qui est inclus dans une séquence alternée de « 01010... ». D'une manière générale, les fluctuations des positions des fronts d'un signal constituent ce que l'on appelle la *gigue de phase* de ce signal (jitter en anglais).

On peut obtenir une représentation très parlante de la déformation du signal reçu en visualisant celui-ci sur un oscilloscope dont le balayage horizontal est synchronisé avec le rythme d'émission des données binaires. La largeur de l'écran représente alors la durée T_b de l'intervalle de temps élémentaire (ITE) affecté à chaque donnée (ou bit dans notre cas). De cette façon, toutes les contributions de nombreux ITE se superposent en remettant périodiquement à zéro l'échelle des temps. L'image obtenue, appelée *diagramme en œil*, est schématisée à la droite de la figure 7-7. Les ITE inclus dans de longues suites de « 0 » ou de « 1 », et dans lesquels le signal a atteint les niveaux crêtes, délimitent le diagramme horizontalement. Les ITE où le signal est le plus déformé, c'est-à-dire ceux qui appartiennent à des séquences alternées, déterminent les frontières courbées auxquelles le diagramme doit son nom. Les contributions des ITE n'entrant pas dans l'une de ces catégories couvrent de façon plus ou moins complète la zone hachurée. Lors de la reconstitution des données, on minimisera la gigue en choisissant le seuil de décision au niveau traversant l'œil (région non hachurée) dans sa plus grande largeur. Dans le cas où le bruit se superpose au signal capté, le diagramme en œil est moins ouvert que pour le signal exempt de bruit. Si la tension efficace du bruit σ est suffisamment faible, le bruit n'interviendra qu'au voisinage des instants de passages par zéro. Son effet essentiel sera de déplacer les fronts du signal reconstitué, autrement dit de contribuer à la gigue de phase. Si au contraire, σ est importante au point de fermer complètement l'œil, le système de reconstitution est inexploitable. Un comparateur à hystérésis, comme utilisé dans notre récepteur, constitue un bon choix pour diminuer considérablement l'effet du bruit.

En plus du bruit, il existe une seconde cause d'erreur sur la détermination de la valeur échantillonnée d'un symbole particulier (à un bit dans notre cas), qui est la distorsion intersymbole ou interférence entre symboles (IES), due aux symboles transmis avant et après le symbole considéré. Lorsqu'on applique une impulsion de Dirac à l'entrée du filtre de transmission, sa réponse impulsionnelle est retardée et élargie par la limitation de bande passante, et présente généralement des oscillations en forme d'un sinus cardinal dont le lobe principal correspond à l'impulsion d'entrée en signe et en amplitude. A la valeur obtenue à l'instant d'échantillonnage pour une impulsion particulière s'ajoute donc algébriquement, les contributions des autres impulsions précédentes et suivantes. Il existe une phase particulière du signal d'horloge d'échantillonnage pour laquelle la somme de ces contributions est minimale. Cet instant d'échantillonnage optimal est visible sur le diagramme en œil du signal de sortie du filtre ou de son image à la sortie du démodulateur. Il existe des familles de filtres pour lesquelles les contributions des autres impulsions sont nulles et donc l'œil complètement ouvert. Leur réponse

impulsionnelle s'annule périodiquement à des instants multiples d'une période de symbole T_s . La plus connue est celle des filtres dits en cosinus surélevé, figure 6-9.

7.2.1.3. Gigue de phase [17] :

L'horloge ou rythme servant à échantillonner le signal à la sortie du démodulateur doit être reconstituée ou récupérée à partir du signal lui-même. On utilise généralement pour ce faire ses passages à zéro. L'horloge est affectée d'un bruit de phase (gigue) si les passages à zéro ne sont pas espacés exactement d'un multiple de T_s . Une gigue aléatoire est donc surtout due au bruit de l'horloge, tandis qu'une gigue déterministe est surtout due aux limitations de bande dans le trajet des données. On constate que plus la gigue est importante plus l'ouverture de l'œil est étroite.

Les mesures de la gigue sont essentiellement basées sur le domaine temporel des diagrammes en œil. La construction du diagramme en œil ainsi que la mesure de la gigue de phase, sont schématisées en figure 7-8.

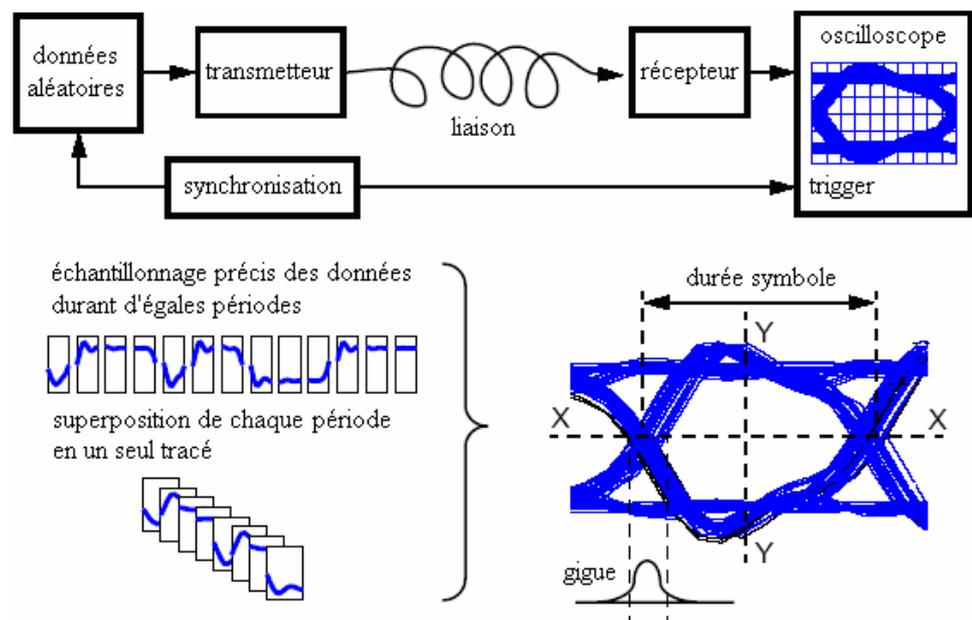


Figure 7-8: Mesure de la gigue à partir du diagramme en œil

Une source de données pseudo-aléatoires, est généralement utilisée dans la mesure de la hauteur de la gigue (épaisseur horizontale du trait) aux passages par zéro du diagramme en œil.

7.2.2. Tests et résultats pratiques des circuits d'implant

Un relevé de la réponse fréquentielle du démodulateur FM a été effectué en fonction de la fréquence variable d'un générateur sinusoïdal HF. Le discriminateur FM est composé d'une cellule LC à point milieu, chargée par un détecteur d'enveloppe, figure 6-31. Le relevé, présenté dans la figure 7-9, donne la transmittance V_s / V_e en fonction de la fréquence du signal d'entrée. La courbe en cloche du circuit résonant est plutôt centrée sur 1.96 MHz. Ceci est dû aux tolérances du couple L et C utilisé. Il est très difficile de fabriquer manuellement une bobine à point milieu et la faire accorder sur 2 MHz avec un condensateur à valeur normalisée la plus proche.

Selon la figure 7-9, le gain maximal du circuit n'excède pas 0.12 à sa résonance. Ainsi, sur l'un des flancs de la courbe, une excursion de fréquence est convertie en une variation relative d'amplitude de 0.12 seulement, au maximum.

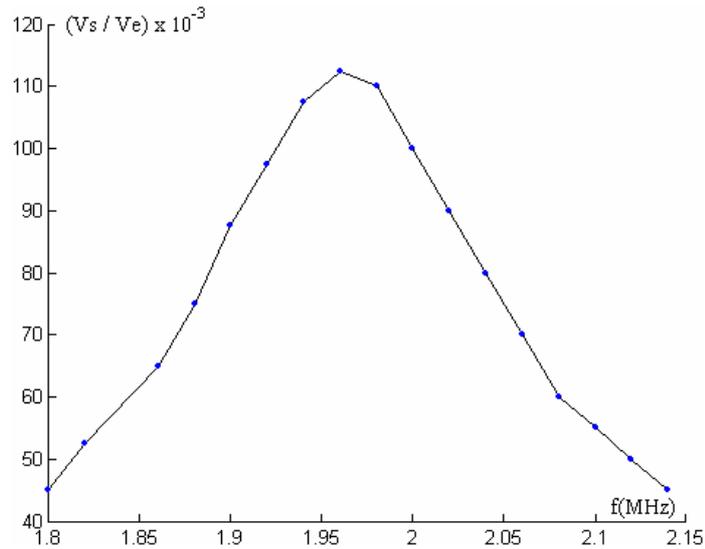


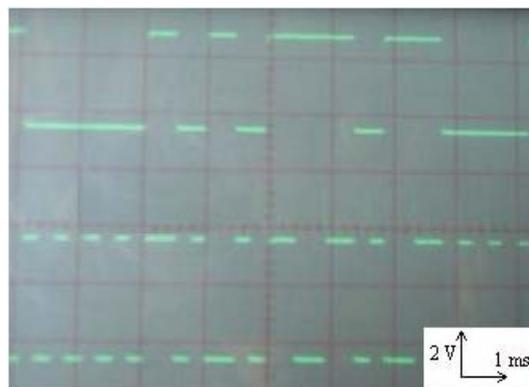
Figure 7-9: Relevé de la réponse fréquentielle du discriminateur FM

Certaines mesures ont été menées sur les deux cartes de transmission et réception, réalisées pour travailler conjointement avec la carte DSP. Ces mesures rendent compte du degré de performance de la liaison assurée par ces cartes.

7.2.2.1. Oscillogrammes :

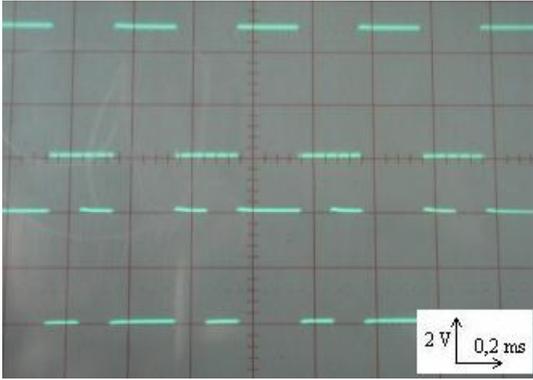
Différents signaux ont été relevés aux points clés des deux circuits. Les oscillogrammes correspondants et affichés sur un oscilloscope ont été photographiés.

Au niveau du circuit de transmission, nous avons sélectionné la fonction GPA (générateur pseudo aléatoire) du GAL. Les données en sortie du GPA (en haut de l'oscillogramme), sont modulées en MSK au rythme de 4800 bps à la sortie du GAL (en bas de l'oscillogramme), planche(1).

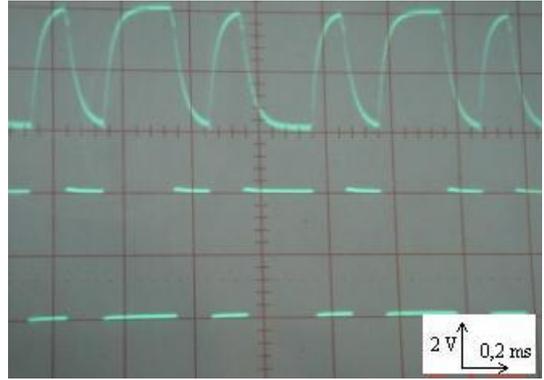


Planche(1): Données aléatoires du GPA et leur modulation MSK

Pour des données périodiques à l'entrée du GAL (en forme d'un signal carré, en haut de l'oscillogramme), ce dernier fournit le signal MSK correspondant (en bas de l'oscillogramme), planche(2). L'horloge de rythme est maintenue à 4800 bps.



Planche(2): Sortie modulateur MSK pour signal carré à l'entrée

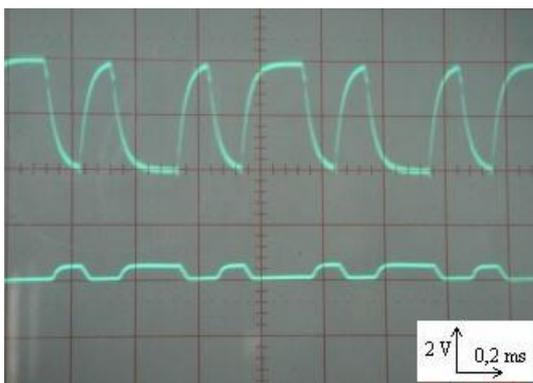


Planche(3): Sortie filtre de transmission comparée à son signal d'entrée

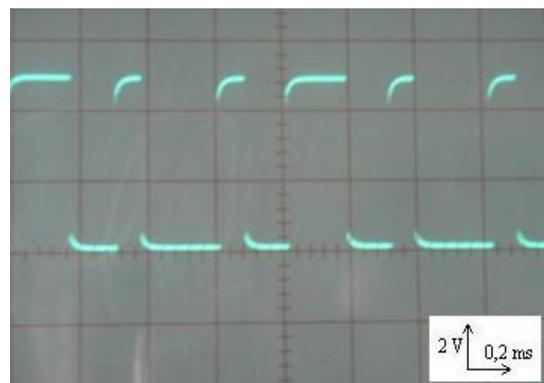
Le même signal MSK de la planche(2) est maintenu (en bas de l'oscillogramme), pour comparaison au signal filtré par le filtre de transmission (en haut de l'oscillogramme), planche(3). Le filtre de transmission présente un gain unitaire avec inversion de phase.

Au niveau de la carte de réception, le démodulateur FM permet de retrouver la forme du signal transmis avec cependant une tension crête à crête de 0.3V. Le signal démodulé est inversé en phase (en bas de l'oscillogramme), par rapport au signal de référence en sortie du filtre de transmission (en haut de l'oscillogramme), planche(4). Ceci est dû à ce que l'ampli cascade de réception est un inverseur de phase (les inversions de phase consécutives, dues à l'additionneur et l'ampli de puissance du transmetteur, se compensent).

Le signal à l'entrée de l'ampli cascade de réception, est amplifié au maximum (à cause du niveau élevé du signal d'entrée et le gain aussi élevé de l'ampli), soit donc 5V. Ce signal ne peut être aperçu au drain, car le circuit de démodulation FM y est connecté et converti l'onde FM en onde AM avant détection d'enveloppe. L'onde AM jouit donc de deux enveloppes à 5V crête à crête, dont il ne restera qu'une seule enveloppe à 2.5V crête à crête après détection d'enveloppe. Or, la transmittance du circuit résonant de démodulation est égale à 0.12 au maximum, figure 7-9. La tension détectée est donc de $0.12 \times 2.5V = 0.3V$ crête à crête. Ce qui confirme la mesure obtenue sur la planche(4).



Planche(4): Sortie démodulateur FM comparée au signal transmis

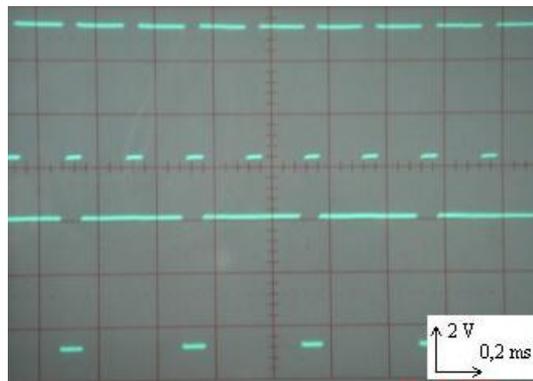


Planche(5): Mise en forme du signal démodulé en FM

Le signal démodulé en FM est mis en forme par la suite, grâce au trigger de Shmitt à seuils ajustables, planche(5).

Les deux monostables utilisés dans la démodulation MSK non cohérente, fournissent à leurs sorties les impulsions de la planche(6). Le monostable non redéclenchable délivre le signal

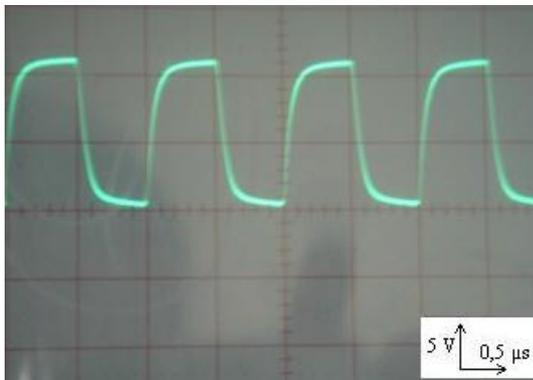
d'horloge (en haut de l'oscillogramme), tandis que le monostable redéclenchable délivre un signal de pseudo-données (ou données brutes) réclamant encore du traitement (en bas de l'oscillogramme).



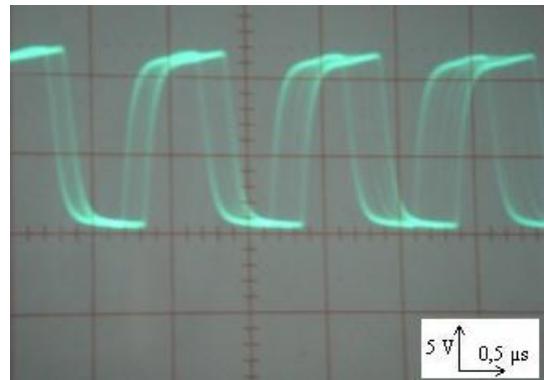
Planche(6): Signal horloge et données brutes

La lecture des données brutes sur fronts descendants de l'horloge récupérée de la planche(6), permet la restitution des vraies données, répliques de ceux en planche(2).

Ayant parcouru quelques signaux issus du traitement en bande de base, les planches suivantes donnent un aperçu sur les signaux HF mis en jeu. La planche(7) illustre la porteuse impulsionnelle de 1MHz non modulée en sortie de la PLL. La planche(8) visualise la modulation FM par les signaux MSK et dont la porteuse est stabilisée par PLL.

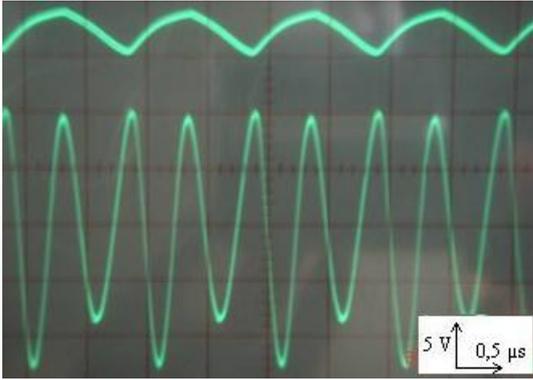


Planche(7): mi-porteuse en sortie PLL

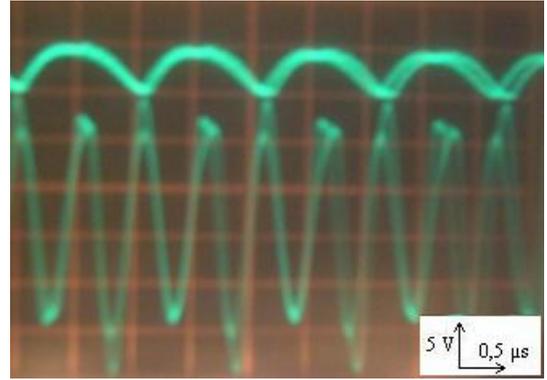


Planche(8): Modulation FM en sortie PLL

Au niveau du transistor amplificateur de puissance de la transmission sont montrés, dans la planche(9), les signaux de base et de collecteur en dehors de la modulation FM. Le signal de base est formé d'arcs exponentiels à la fréquence de 1 MHz (en haut de l'oscillogramme), alors que le signal du collecteur est une onde sinusoïdale filtrée à la fréquence de 2 MHz (en bas de l'oscillogramme). Il s'agit bien d'un multiplicateur de fréquence par 2. Cependant l'onde n'est pas parfaitement symétrique malgré de maintes efforts pour y remédier. Apparemment, c'est une addition de deux ondes sinusoïdales de fréquences voisines. La largeur de bande que présente l'ampli sélectif d'émission à coefficient de qualité estimée à $Q_p = 4.3$, est sans doute responsable du non filtrage de la fréquence parasite additionnelle.



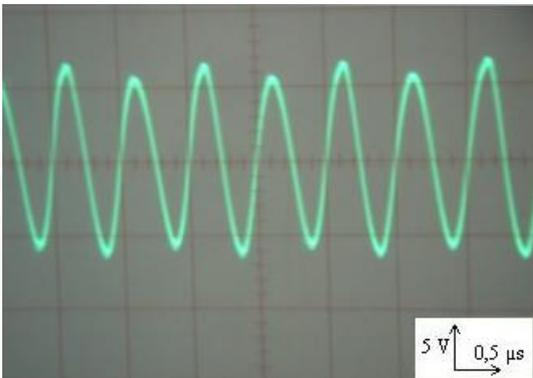
Planche(9): Tensions base et collecteur au repos de l'ampli de puissance



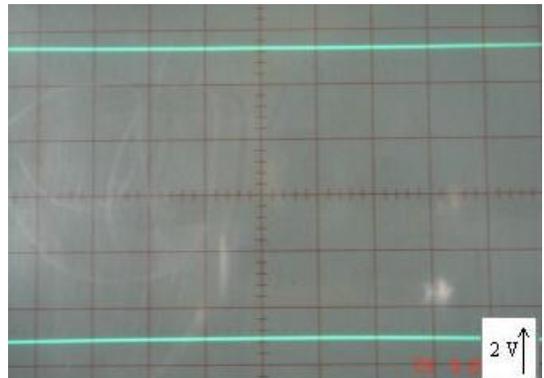
Planche(10): Tensions base et collecteur modulées en FM de l'ampli de puissance

En présence d'une modulation FM, les signaux à la base et au collecteur du transistor de puissance sont illustrés dans la planche(10). La modulation est un peu difficile à voir malgré l'illumination de l'écran de l'oscilloscope. Des pics parasites sont superposés à certaines alternances de la porteuse.

Pour ce qui est de l'alimentation à partir d'énergie HF, la planche(11) montre l'onde reçue sans modulation au secondaire (normalement implanté derrière l'oreille du patient). Il faut noter que le signal est limité dans ses alternances positives et négatives par les diodes Zener, et que le secondaire est chargé du circuit récepteur.



Planche(11): Onde reçue sans modulation

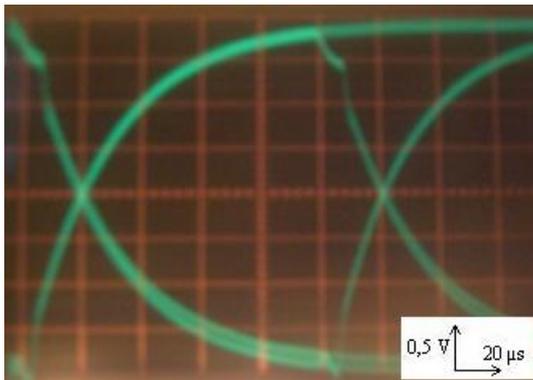


Planche(12): Tensions d'alimentation développées au récepteur

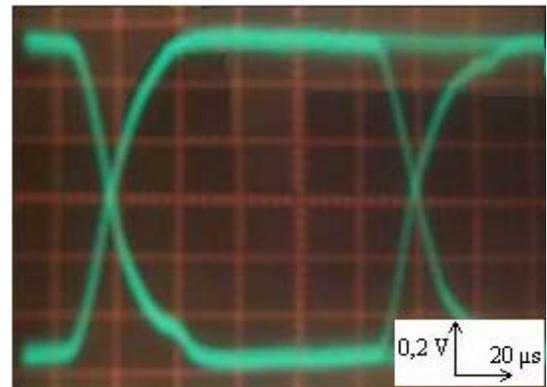
Séparément, les deux alternances de la planche(11) subissent un redressement simple alternance par diode rapide 1N4148, un filtrage par condensateur-goutte à $6.8 \mu\text{F}/6.3\text{V}$, et une régulation par diode Zener BZX460. Les tensions continues et symétriques, résultantes, sont montrées dans la planche(12).

Parmi les moyens dont on dispose pour évaluer la qualité d'un système de transmission de données, on cite : l'obtention du diagramme en œil, et la mesure de la gigue de phase, paragraphe 7-2-1. Ainsi, le signal MSK d'un signal carré filtré à la sortie du filtre de transmission, planche(3), est visualisé sur l'oscilloscope dont le balayage horizontal est synchronisé avec le rythme d'émission des données binaires. La largeur du diagramme en œil obtenu sur l'écran, représente alors la durée 0.1ms comme intervalle de temps élémentaire (ITE) affecté à chaque état du signal MSK (dans un signal MSK, l'impulsion attribuée à la donnée '0' change d'état toutes les 0.1ms).

De cette façon, toutes les contributions des nombreux ITE se superposent, formant le diagramme en œil de la planche(13). La reconstitution de l'horloge au récepteur s'effectue à partir du signal envoyé lui-même, et on exploite pour cela ses passages à zéro. Si les passages à zéro ne sont pas espacés exactement d'un multiple de période de bit, dû à un bruit de phase de l'horloge d'émission ou aux limitations de bande dans le trajet des données, l'horloge récupérée sera elle aussi affectée d'un bruit de phase (ou gigue de phase). Dans la planche(13), et à l'intersection avec l'axe temporel fictif, le diagramme en œil présente une gigue de phase d'au moins $\frac{1}{2}$ de graduation, soit $\frac{1}{2} \times 0.2 \times 20\mu\text{s} = 2\mu\text{s}$. Cette mesure peut au maximum s'étendre au double: $4\mu\text{s}$.



Planche(13): Diagramme en œil à la sortie du filtre de transmission



Planche(14): Diagramme en œil à la sortie du démodulateur FM

Un autre diagramme en œil est relevé à la sortie du démodulateur FM. Il est présenté dans la planche(14). Une gigue de phase similaire au cas précédent peut être mesurée: 2 à $4\mu\text{s}$. Pour la mise en forme du signal, la bonne ouverture de l'œil a facilité le positionnement à son intérieur, des deux seuils de décision propres au trigger.

7.2.2.2. *Mesure du taux d'erreur BER :*

Le taux d'erreur de bit ou BER (Bit Error Rate) est défini comme le rapport du nombre de bits erronés sur le nombre total de bits. Le programme mis au point au paragraphe 6-3, permet la mesure du BER en utilisant un signal de test pseudo-aléatoire, ainsi que l'affichage d'un diagramme en œil après le filtre de transmission. C'est le programme BER_DiagOeil.c.

Lors de l'exécution de ce programme, nous avons fixé un temps de mesure d'une heure pour obtenir les résultats suivants sur deux essais :

Le tableau d'une taille de 7714 de données aléatoires à 32 bits, et sauvegardé dans le fichier « gauss32.c », est transmis à $n = 43$ fois. Soit un nombre de bits envoyés de $\text{bitMax} = 10\,763\,952$ bits (en retranchant les temps morts, tel que les temps de transit et de calcul). Les bits erronés recensés sont $\text{bitCnt} = 6$ à 8 bits, d'où $\text{BER} = 10^{-6}$.

7.2.2.3. *L'alimentation par champ externe :*

L'analyse menée sur un transfert optimisé de puissance, allant d'une bobine d'émission vers une autre de réception dans le but d'alimenter un implant biomédical, nous a permis d'évaluer la puissance réellement reçue par ce dernier. Cette évaluation théorique a pris en compte les deux effets dus, à la présence du tissu biologique d'une part, et au désalignement éventuel des bobines de transfert d'autre part. En dehors du tissu biologique, un test pratique des rangées maximales de désalignement des bobines de transfert a été effectué afin de mesurer l'influence sur l'alimentation induite par voie HF [16], figures 7-10 et 7-11. Une chute de tension de 6V jusqu'à 5V est obtenue à partir des désalignements :

- axial seul : $\Delta = 7 \text{ mm}$, $\psi = 0^\circ$.
- angulaire seul : $\Delta = 0$, $\psi = 30^\circ$.

Dans le cas d'un désalignement combiné, axial et angulaire, par exemple: $\Delta = 5 \text{ mm}$, $\psi = 10^\circ$, Nous avons eu confirmation d'un accroissement de la tension de sortie (vers 6V) lorsque les désalignements se faisaient dans un même sens, et d'une diminution (à 5.4V) lorsque ces mêmes désalignements se faisaient dans des sens opposés, figure 3-10.

Les tensions parasites, pic à pic, observées sur la tension (9V) de la pile alimentant l'ampli émetteur, sont de 80 mV évoquant un taux de perturbation de $0.08 / 9V = 0.88 \%$. Tandis que celles observées sur la ligne d'alimentation induite en sortie, étaient de 20 mV, donnant un taux de perturbation de $0.02 / 6V = 0.33 \%$.

On précisera que la tension continue de 6V a été obtenue par filtrage uniquement, sans l'adjonction de diodes Zener, et ceci, sous une charge équivalente de 2.2 k Ω pour l'implant. On ajoutera aussi que dans cette expérience, le transistor de puissance de l'émetteur était excité par une onde sinusoïdale écrêtée (due à la polarisation en classe C), se présentant à la même fréquence d'accord que celle du circuit de son collecteur. En d'autres termes, cet ampli était accordé sur le fondamental du signal d'entrée et non sur son deuxième harmonique comme dans le cas du transmetteur de notre implant final.

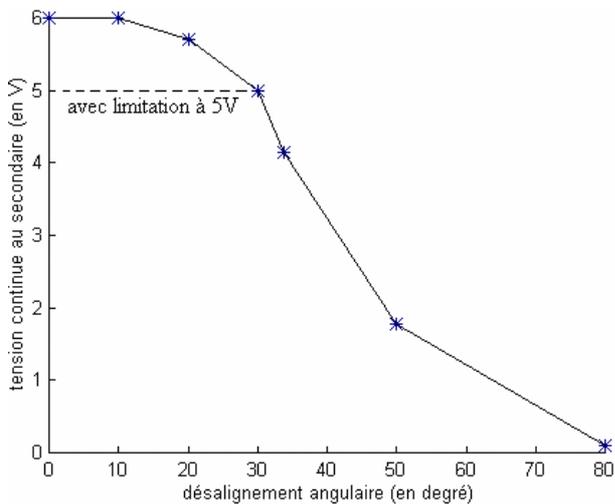


Figure 7-10: Relevé de la tension d'alimentation de l'implant en fonction du désalignement angulaire seul

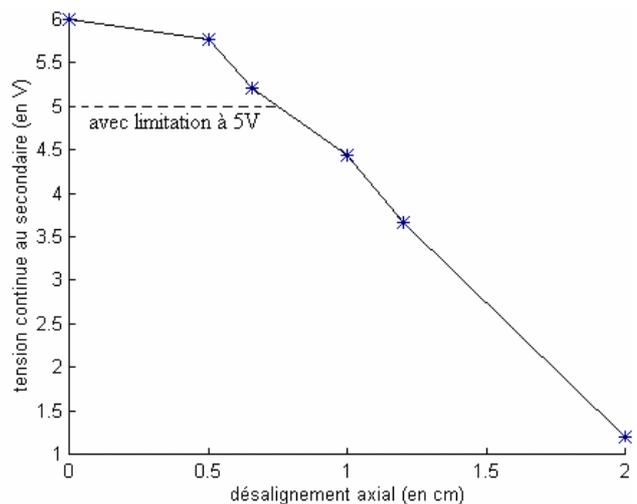


Figure 7-11: Relevé de la tension d'alimentation en fonction du désalignement axial seul

D'autre part, nous avons mesuré en complément, l'allure des tensions d'alimentation positive et négative de l'implant final, en fonction de la charge, figure 7-12. Dans cette expérience, les diodes Zener étaient bien placées et le secondaire encore chargé par le récepteur (entrée de l'ampli cascode).

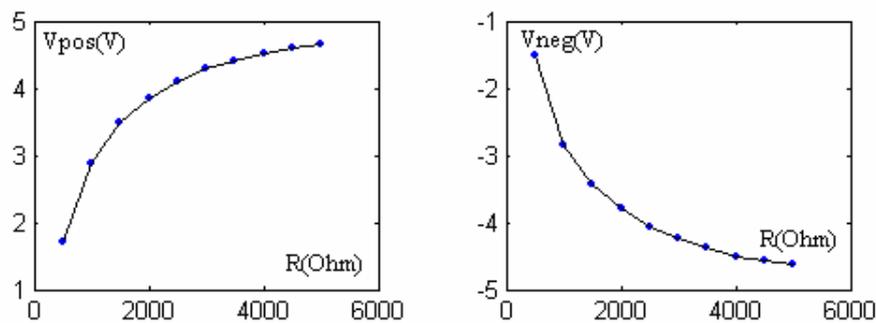


Figure 7-12: Tensions d'alimentation, positive et négative, en fonction de la charge

On constate que les deux tensions d'alimentation diminuent à fur et à mesure que la charge additionnelle augmente (diminution de R). Si on tolère une baisse de tension jusqu'à 4.7V, la résistance minimale est de 5 k Ω . On peut affirmer ainsi, que le courant maximal est de l'ordre de 1mA sous chaque alimentation, soit 2 mA au total. Nous arrivons donc à une puissance disponible au secondaire d'environ 9mW, un peu inférieure à celle prévue au chapitre 3, $P_s = 11mW$. La limitation de puissance ainsi mesurée, malgré l'omission de l'atténuation du canal de transmission représentée par le tissu biologique et le désalignement des bobines de transfert, est due surtout à l'accord de notre transistor d'émission sur le deuxième harmonique et non sur le fondamental du signal d'entrée. Le transistor doit donc réaliser deux fonctions (amplification + multiplication de fréquence) au détriment d'un transfert optimal de l'énergie HF.

Une autre cause peut être due à l'interaction mutuelle des deux fréquences porteuses très voisines, obtenues pratiquement et qu'on ne peut séparer, planche(9).

La présence du circuit récepteur comme charge, en plus de R, limite aussi le courant disponible du circuit d'alimentation.

7.3. Capture des paramètres de stimulation par RTDX

Pour visualiser les paramètres de stimulation de l'oreille interne d'un sourd, la mise en œuvre des fenêtres du logiciel développé, Sim-CI, est la suivante :

On lance le Sim-CI après avoir chargé le programme « implant.out » dans la mémoire du DSP. Dans la fenêtre de configuration du Sim-CI, figure 6-43, on a à choisir entre le mode de fonctionnement normal (envoi des paramètres par défaut : MCL=63, THR=0, et ajustage=0), ou le mode test (envoi de paramètres personnalisés). On clique sur « appliquer » pour valider la configuration en cours, puis sur « normal » pour quitter le mode test. En cliquant sur « suivant », la configuration retenue est envoyée au DSP par liaison RTDX, et la fenêtre d'acquisition s'affiche, figure 6-44. Un clic sur « lancer » démarre l'acquisition et la visualisation des paramètres de stimulation de la cochlée.

Les chronogrammes et l'électrodiagramme représentés dans les figures 7-13 et 7-14 correspondent à la phrase enregistrée « tous ce que apporte le coran demeure indéniablement constant jusqu'à la fin des temps ». La stimulation électrique porte sur les types d'électrode, Nucleus 22 et Clarion, respectivement.

L'Analyse des graphes, permet de les décomposer en deux parties (A et B) :

- **A** : est un segment de parole dont on distingue une variation de pitch, intensité, et position des formants (trajets formantiques).
- **B** : est un exemple de zone de silence, où le pitch et l'intensité doivent être nuls avec une absence totale de formants.

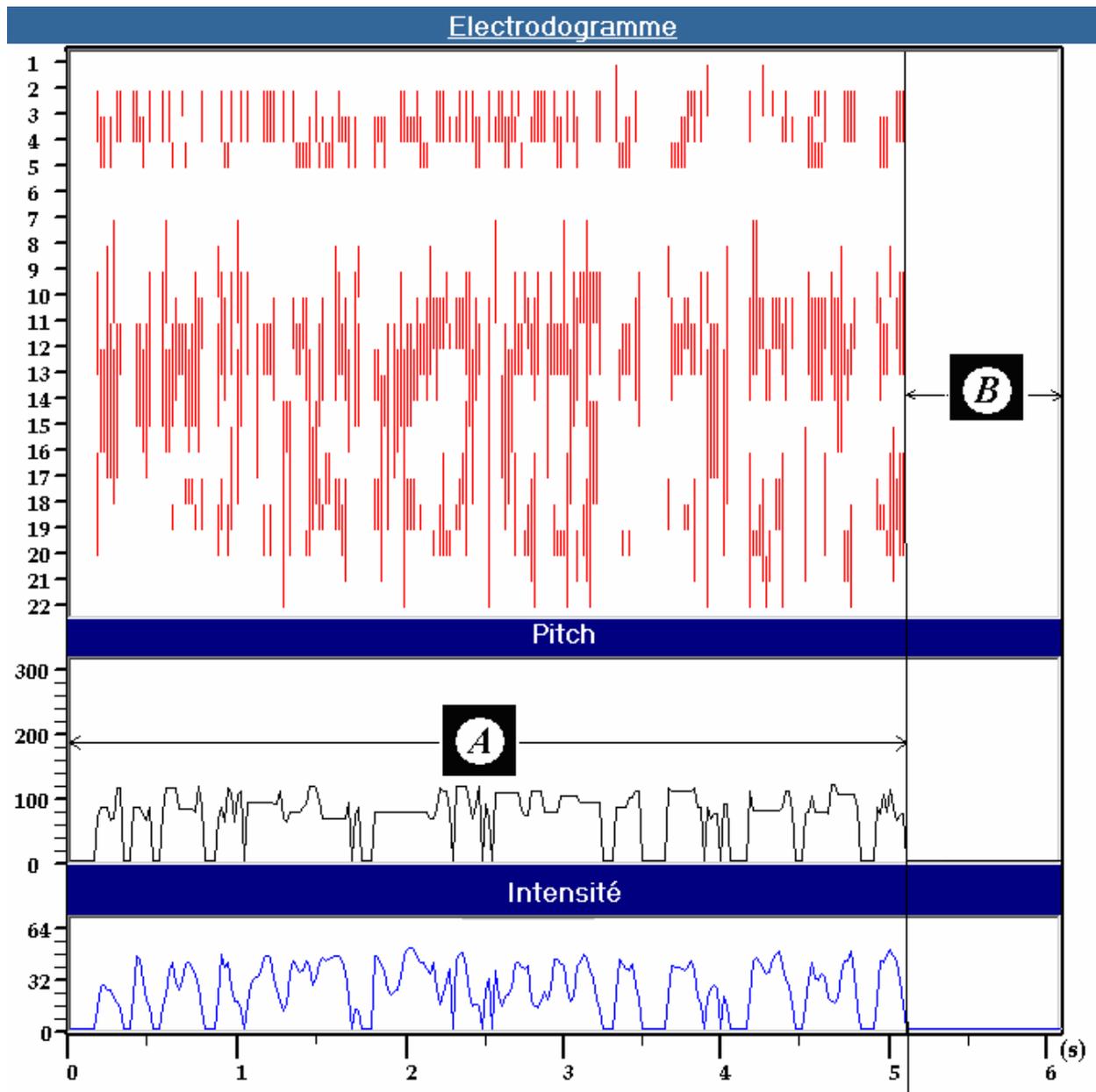


Figure 7-13: Chronogrammes et électrodogramme correspondant à la phrase « tout ce que apporte le coran demeure indéniablement constant jusqu'à la fin des temps » (Nucleus 22)

On constate bien que les trajets formantiques avec le type d'électrode Nucleus 22 sont plus visibles qu'avec le type de Clarion, à cause du nombre d'électrodes qui est presque deux fois plus élevé chez Nucleus 22 que Clarion. Mais une augmentation du nombre d'électrodes ne peut être liée de façon formelle à une meilleure perception du langage. Cette dernière est fonction de la surdité en termes de quantité des fibres nerveuses restantes et leur location.

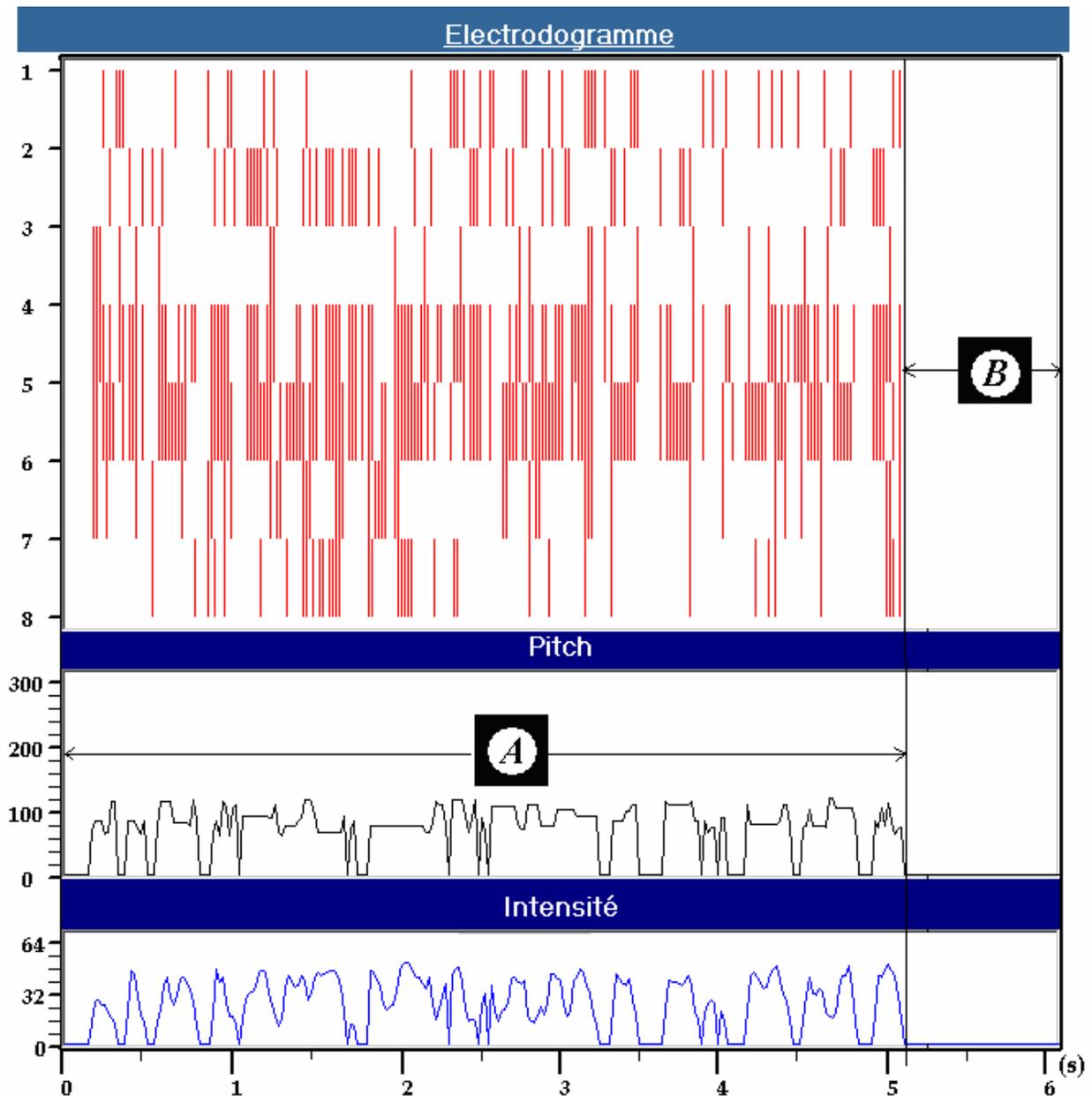


Figure 7-14: Chronogrammes et électrodogramme correspondant à la phrase « tout ce que apporte le coran demeure indéniablement constant jusqu'à la fin des temps » (Clarion)

Conclusion :

Notre stratégie d'implant a bénéficié d'une série de tests et de mesures, que ce soit sur le plan hard ou soft. Dans ce contexte, un pseudo spectrogramme de la parole obtenu par analyse LPC, est confronté à un sonagramme à large bande présenté au chapitre 4.

Une certaine similitude des spectres, notamment dans les segments propres aux voyelles, a été observée. La majorité des consonnes rencontrées dans les fichiers audio d'essai, se présente sous forme diffuse (aspect d'un bruit blanc) dans le spectre LPC, alors qu'elles sont généralement absentes des sonagrammes.

La principale différence des deux types de spectre utilisés, se rapporte à la remarque suivante :

Le spectre du sonagramme est basé sur l'information d'énergie ou amplitude relative des raies individuelles, traduite dans le tracé par un niveau de gris approprié. Alors que dans l'analyse LPC à court terme, le spectre résultant est basé directement sur l'information du couple formant-

bande. L'information d'énergie ou valeur RMS, connue aussi par tranche de la parole, correspond quant à elle, à celle de tous les formants réunis de la tranche.

La représentation LPC des bandes formantiques ne s'appuie nullement sur l'information d'énergie par formant mais sur une estimation directe, faite à partir du modèle prédicteur. Elle est par conséquent, plus exacte que celle du sonagramme.

Ceci dit, pour parfaire au contraste des formants obtenus sur un sonagramme, le spectre doit d'abord bénéficier d'un polissage efficace. Apparemment, le polissage effectué par fenêtrage Hanning, s'avère insuffisant puisqu'on rencontre par moment, un seul formant scindé en deux (cas du sonagramme du fichier parenthèse.aud).

L'aspect diffus des spectres LPC est dû à la représentation discrète en dix points de chaque bande formantique avec chevauchement éventuel entre les bandes de formants voisins.

Les fortes concentrations de ces points ne trahissent que les formants à faibles bandes correspondants aux voyelles (des dizaines de points concentrés sur de faibles étendues).

Les formants à bandes élargies restent quant à eux sous un aspect diffus et sont associés à la plupart des consonnes. En dépit de l'utilisation d'un critère de décision sur le pitch plus sévère (toute la courbe de la figure 5-12 est décalée de 0.1 vers le haut), la mesure du pitch est toujours effectuée pour les deux fichiers analysés.

Les différentes consonnes telles les occlusives non voisées, les fricatives voisées, les liquides voire les nasales, ont été modélisées par des formants à large bande causant les zones diffuses qui leur sont attachées dans les spectres. On rappelle qu'une zone diffuse d'une consonne, étalée le long de l'axe fréquentiel du spectre, est équivalente à celle d'un bruit blanc généré localement dans le modèle des sons non voisés, figure 4-7. De ce fait, il s'avère donc possible de supprimer la détection des segments non voisés au profit d'une utilisation exclusive des formants et leurs bandes dans la distinction entre les différents phonèmes perçus de la parole. Les consonnes seront donc représentées par des transitions formantiques plus ou moins courtes. Les segments parole-silence, détectés automatiquement dans les différents sonagrammes, sont de même, bien représentés dans les spectres LPC. Autre particularité de l'analyse LPC, est la mesure de la faible évolution du pitch au cours de la parole. L'erreur classique [74] du dédoublement de la mesure du pitch et se manifestant essentiellement en présence de voyelles, a été minimisée par un algorithme de correction. L'inconvénient de l'analyse LPC ainsi développée, est la non détection des fricatives non voisées comme le [s].

En connectant les cartes de transmission et de réception entre-elles, différentes mesures ont été effectuées telle que : le taux d'erreur binaire, le diagramme en œil après démodulation, et la gigue de phase à la réception. À cette fin, un fichier de données pseudo-aléatoires à 32 bits, a été utilisé. Des oscillogrammes ont été prélevés à différents points des deux cartes réalisées, ainsi que des mesures portant sur l'alimentation par champ externe, ont également été effectués. Ces différentes mesures sont en accord avec les calculs menés aux chapitres précédents.

Conclusion générale

L'implant cochléaire constitue une avancée majeure dans le traitement des surdités. C'est une réussite dans la restauration partielle de l'audition chez les personnes profondément sourdes. Dans l'intention de contribuer à la réhabilitation auditive des mal-entendants, une conception et une réalisation d'un prototype d'une prothèse cochléaire ont fait l'objet du présent travail. Des études anatomique et physiologique de l'oreille, suivies d'une étude sur les possibilités offertes et résultats obtenus par des implants cochléaires commerciaux les plus utilisés actuellement dans le monde, ont été menées. De ces études, il en est ressorti un certain nombre de critères qui nous ont permis de suggérer une nouvelle stratégie de stimulation électrique de la cochlée. Cette stratégie repose sur la norme GSM, qui a fait ses preuves de fiabilité et de robustesse vis à vis des bruits, dans le domaine de la téléphonie mobile. Les paramètres pertinents de la parole sont extraits à court terme par un DSP, puis transmis à faible débit vers l'implant. Dans ce contexte, un algorithme complet basé sur la modélisation paramétrique de la parole par prédiction linéaire autorégressive est développé. Les données utilisées dans la stimulation électrique de la cochlée, se ramènent ainsi à un jeu de paramètres de la parole tels que le pitch, l'intensité acoustique, les formants et leurs bandes, qui sont paquetisés et protégés par CRC (cyclic redundancy check) avant leur transmission HF vers le récepteur implanté. Des cartes de transmission-réception de données ont été conçues pour travailler conjointement avec la carte DSP qui s'occupe de l'essentiel du traitement. Pendant leur transmission à l'implant, les paramètres de la stimulation cochléaire sont capturés par liaison RTDX (real time data exchange) pour leur visualisation en temps réel sur une interface graphique. Les fenêtres d'affichage développées sous Visual Basic permettent une représentation graphique sous forme de chronogrammes propres à la fréquence fondamentale (pitch) et à l'intensité acoustique. L'information, des formants et leurs bandes, est traduite en électrodiagramme, montrant le lieu et l'étendue de la stimulation électrique à travers les électrodes employées.

Une fenêtre de configuration est introduite en vue de dresser une carte du patient à partir des premiers essais post-opératoires. Dans cette carte, figurent les seuils d'audition minimal (THR) et maximal (MCL), inhérents au patient, ainsi que l'ajustage éventuel du lieu d'insertion de l'électrode. Cette dernière correction est nécessaire en cas d'un défaut commis lors de la phase opératoire, relativement au décalage de l'électrode par rapport à son emplacement optimal dans la cochlée.

Notre conception d'implant cochléaire a été dotée de travaux en parallèle, en vue de la consolider par une réalisation optimale, vérification de ses données et mesure de ses performances. Ainsi, une étude approfondie est effectuée concernant l'alimentation de l'implant par champ électromagnétique externe, conjointement à la transmission des données par bobines couplées. Un outil de comparaison du spectre de la parole, basé sur le concept des sonagrammes à large bande, a été développé afin de valider la stratégie de stimulation suggérée. La circuiterie de transmission-réception a bénéficié d'une évaluation de ses performances en terme de mesure du taux d'erreur binaire, diagramme en œil, et gigue de phase. Différents oscillogrammes ont été photographiés, mettant en évidence le bon fonctionnement des circuits réalisés. Une gigue de phase de l'ordre de 4 μ s a été mesurée, n'altérant aucunement l'ouverture du diagramme en œil obtenu à la sortie du démodulateur FM du récepteur. Un taux d'erreur binaire inférieur à 10^{-6} , a également été mesuré pendant une heure de transmission de données aléatoires à 32 bits (quelques bits erronés pour plus de dix millions de bits transmis).

La limitation de la bande vocale à 3.6 kHz par le codec employé, a apporté une atténuation progressive des formants détectés à partir du 3^{ème} formant. En conséquence, une insuffisance a été constatée au niveau spectral concernant la non détection des fricatives non voisées. Cette limitation peut être levée en poussant l'extraction des formants jusqu'au 5^{ème} formant. Cela apportera plus d'information sur les fréquences élevées véhiculées par de telles consonnes, à

condition que la bande vocale traitée s'étend effectivement à 5 kHz, et que l'ordre du polynôme prédicteur soit égal à 12.

Pour trancher la question relative à notre apport effectif, notre stratégie doit être essayée sur un implant cochléaire programmable du commerce. Seule l'expérimentation sur des patients pourra nous renseigner sur l'efficacité de la stratégie proposée.

La plupart du succès des implants cochléaires est dû aux techniques avancées de traitement du signal. En dépit du succès très encourageant, il reste beaucoup à apprendre sur la stimulation électrique du nerf auditif et répondre ainsi aux questions non encore résolues. Il y a des recherches actuellement en cours sur les stratégies de codage qui seraient susceptibles de procurer une bonne écoute de la musique. Il y aurait deux programmes possibles : un pour la parole, et un pour la musique. La plupart des stratégies en étude ou avec peu d'années d'utilisation, s'orientent vers l'emploi de la stimulation simultanée, comme : SPS (Simultaneous Pulsatile Sampler), MPS (Multiple Pulsatile Sample), ASR (Adaptative Stimulation Rate), HAP (Hybrid Analog Pulsatil).

Bien que ces nouvelles technologies soient préférées chez certains patients, d'autres trouvent que la stimulation simultanée de toutes les électrodes reste difficile à comprendre. Les recherches doivent aboutir au développement de nouvelles techniques de traitement du signal, capable de transmettre plus d'information au centre auditif. Le développement d'algorithmes robustes vis à vis du bruit, aidera mieux les patients implantés dans les communications en milieu bruité. Pour cela, il faut découvrir totalement le codage naturel de la parole dans l'oreille saine et le comparer aux effets d'une stimulation électrique du nerf auditif.

On peut espérer que les futures recherches dans les prothèses cochléaires arriveront à un stade qui permettra aux patients d'atteindre les meilleures performances, et que d'autres générations d'implants vont progressivement élargir les indications pré-opératoires vers certaines surdités sévères.

BIBLIOGRAPHIE

- [1] Abbas P., "Electrophysiology", Audiological Foundation, Singular Publishing Group inc, Tyler ed., pp. 317-356, 1993.
- [2] Achouri D., Sefra L., Batel N. (encadreur), "transmission-réception des données en MSK-FM", mémoire d'ingénieur d'état en électronique, C.U.Médéa, 2005.
- [3] Althans M., « PC/XT/AT/386 - le livre d'or », Ed. Sybex, 1992.
- [4] American speech-language-hearing association, "technical report: cochlear implants", ASHA supplement 24, in press, 2004.
- [5] Auvray J., « Electronique des signaux échantillonnés et numériques », Editions dunod, 1979.
- [6] Batel N., « Circuits HF et de Transmission », TPs du module TEC588, C.U.Médéa, 2003.
- [7] Batel N., Mehenni M., Dekar L., "Attenuation in transferred RF power to a biomedical implant due to the misalgnment coils", Proceedings of World Academy of Science, Engineering and Technology (PWASET), ISSN 1307-6884, vol. 10, pp 160-164, 2005.
- [8] Batel N., Mehenni M., Kouadik S., "Attenuation in transferred RF power to a biomedical implant due to the absorption of biological tissue", Proceedings of World Academy of Science, Engineering and Technology, ISSN 1307-6884, vol. 10, pp 165-168, 2005.
- [9] Baudoin G., Cernock J., Gournay P., Chollet G., "codage de la parole à bas et très bas débit", département signaux et télécommunications, ESIEE, Noisy le grand Cedex.
- [10] Benallal A., Ouchfoun M.L., Khelifi F., « Etude comparative des différentes méthodes de détection du pitch », mémoire d'ingénieur d'état en électronique, UST Blida, 1997.
- [11] Besson R., « Technologie des composants électroniques », Editions Radio, tome n°1, 5^{ème} édition, 1984.
- [12] Blamey P., Dowell R., Clark G., "Acoustic parameters measured by a formant-estimating speech processor for a multiple-channel cochlear implant", Journal of the Acoustical Society of America, vol. 82, pp. 38-47, 1987.
- [13] Boex C., Pelizzone M., Montandon P., "Improvements in speech recognition with the CIS strategy for the Ineraid multichannel intracochlear implant", Advances in cochlear implants, Hochmair ed., pp. 136-140, 1994.
- [14] Boite R., Bourlard H., Dutoit T., Hang J., Leich H., « traitement de la parole », Presses polytechniques et universitaires romandes, 1999.
- [15] Bonfils P., Uziel A., « Anatomie des voies auditives », Editions techniques, Encyclopédie médico-chirurgicale, Paris ORL 20022 A10, 1990.
- [16] Boucherit A., Boucherit A., Batel N. (encadreur), « Transfert transcutané optimal d'une puissance radio-fréquence destinée à l'alimentation d'implant », mémoire d'ingénieur d'état en électrotechnique, C.U.Médéa, 2000.
- [17] Brémenson C., « Faisceaux hertziens – Traité d'électronique », Techniques de l'ingénieur, 1993.
- [18] Bussey H.E., « Measurement of RF properties of materials: a survey », Proceedings of the IEEE, vol. 55, n°6, pp 1046-1053, 1967.
- [19] Chaoui C., Benrebaya M., Batel N. (encadreur), « Modélisation autorégressive de la parole pour une transmission à faible débit », mémoire d'ingénieur d'état en électronique, C.U. Médéa, 2002.
- [20] Charachon R., Uziel A., Gratacap B., Genin J., « Physiologie de l'audition », Encyclopédie médico-chirurgicale, Paris ORL 20030 A10, 1986.
- [21] Chossat M., « Aide-mémoire de mathématique », Ed. Dunod, 1977.
- [22] Clark G., "The university of Melbourne-Nucleus multi-electrode cochlear implant", Advances in ORL, vol. 38, pp. 1-189, 1987.

- [23] Clark G., "The university of Melbourne: Nucleus multi-electrode cochlear implant", Editions Karger, 1987.
- [24] Cours, « Digital Modulation and GSMK », University of Hull.
- [25] Data Bulletin, « MSK modem : MX469 1200/2400/4800bps », MX.COM. INC, 1998.
- [26] Danhauer J., Ghadialy F., Eskwitt D., Mendel L., "Performance of 3M/House cochlear implant users on tests of speech perception", Journal of the American Academy of Audiology, vol. 1, pp. 236-239, 1990.
- [27] Detlaf A., Yavorski B., « Aide-mémoire de physique », Editions de Moscou (5^{ème} édition), 1986.
- [28] Dorman M., Hannley M., Dankowski K., "Word recognition by 50 patients fitted with the Symbion multichannel cochlear implant", Ear and Hearing, vol. 10, pp. 44-49, 1989.
- [29] Dorman M., Loizou P., "Changes in speech intelligibility as a function of time and signal processing strategy for an Ineraid patient fitted with CIS processors", Ear and Hearing, vol. 18, pp. 147-155, 1997.
- [30] Dorman M., Loizou P., "Mechanisms of vowel recognition for Ineraid patients fit with CIS processors", Journal of the Acoustical Society of America, vol. 102, pp. 581-587, 1997.
- [31] Dorman M., "Speech perception by adults", Audiological Foundation, Singular Publishing Group inc, Tyler ed., pp. 145-190, 1993.
- [32] Dowell R., Mecklenburg D., Clark G., "Speech recognition for 40 patients receiving multichannel cochlear implants", Archives of Otolaryngology, Head and Neck Surgery, vol. 112, pp. 1054-1059, 1986.
- [33] Dowell R., Seligman P., Blamey P., Clark G., "Evaluation of a two-formant speech processing strategy for a multichannel cochlear prosthesis", Annals of ORL, vol. 96, pp.132-134, 1987.
- [34] Eddington D., "Speech discrimination in deaf subjects with cochlear implants", Journal of the Acoustical Society of America, vol. 68, pp. 885-891, 1980.
- [35] Edgerton B., Brimacombe J., "Effects of signal processing by the House-3M cochlear implant on consonant perception", Acta Otolaryngologica, pp. 115-123, 1984.
- [36] Encinas J., « systèmes à verrouillage de phase (PLL) », Masson, 1990.
- [37] Eyris C., Chouard C.H., Peytral C., « Systématisation des voies cochléaires », Encyclopédie médico-chirurgicale, Paris ORL, 20020 A20, 1990.
- [38] Fearn R.A., "Music and pitch perception of cochlear implant recipients", thesis of PHD, University of New South Wales, 2001.
- [39] Flack F.C., James E.D., Schlapp D.M., « Mutual inductance of air-cored coils : Effect on design of radio-frequency coupled implants », Pergamon press, Med. & Biol. eng., vol.9, 1971.
- [40] Fontolliet P.G., « Traité d'électricité: systèmes de télécommunication », Presses polytechniques et universitaires romandes, vol.18, 1996.
- [41] François R., « Modulation-Démodulation – Traité d'électronique », Techniques de l'ingénieur, 1993.
- [42] Gueulle P., « Composants électroniques programmables sur PC », ETSF, 1997.
- [43] Gueulle P., « Logiciels PC pour l'électronique », ETSF, 1996.
- [44] Gyenot P., Hans T., « Régulation et Asservissement », Ed. Eyrolles, 1990.
- [45] Hack H., « Le manuel des GAL », PubliTronic, 1994.
- [46] Helms J., Muller J., Schon F., "Evaluation of performance with the COMBI 40 cochlear implant in adults: A multicentric clinical study", ORL, vol. 59, pp. 23-35, 1997.
- [47] Hervé J., « Electronique pour les transmissions numériques », Editions ellipses, 1993.
- [48] Hikmet S., « Transmission des signaux numériques – Traité d'électronique », Techniques de l'ingénieur, 1993.

- [49] Hinojosa R., Marion M., "Histopathology of profound sensorineural deafness", Annals of New York Academy of Sciences, vol. 405, pp. 459-484, 1983.
- [50] Hochmair-Desoeyer I., Hochmair E., "Percepts elicited by different speech-coding strategies", Annals of New York Academy of Sciences, vol. 405, pp. 268-279, 1983.
- [51] Hochmair D., Hochmair E., Stiglbanner H., "Psychoacoustic temporal processing and speech understanding in cochlear implant patients", New York Raven Press, Schindler and Merzenich eds., pp. 291-304, 1985.
- [52] House W., Berliner K., "Cochlear implants: Progress and perspectives", Annals of ORL, vol. 91, pp. 1-124, 1982.
- [53] Jeutter D., Josse F., « Design of a radio-linked implantable cochlear prosthesis using surface acoustic wave devices », IEEE trans. ultrasonics and frequency control, vol. 40, n°5, pp 469-477, 1993.
- [54] Jeutter D., « Telemetry of biomedical signals », Medical & Life sciences eng., vol. 13, pp 131-145, 1994.
- [55] K0 W.H., Liang S.P., Fung C.D.F., « Design of radio-frequency powered coils for implant instruments », Med. & Biol. eng. & Computing, 1977.
- [56] Koch D.B., "Cochlear Implants: An Overview", Advanced Bionics Corporation, Sylmar, CA, 2000.
- [57] Lecerf A., « Physique des ondes et des vibrations », 2^{ème} édition, Lavoisier Tec & Doc, 1996.
- [58] Legent F., Perlemuter L., Vandenbrouck C., « Cahiers d'anatomie ORL », Editions Masson 4^{ème} édition, 1984.
- [59] Loeb G., "Cochlear prosthetics", Annual Review in Neuroscience, vol. 13, pp. 357-371, 1990.
- [60] Loeb G., Kessler D., "Speech recognition performance over time with the Clarion cochlear prosthesis", Annals of ORL, vol. 104, pp. 290-292, 1995.
- [61] Loizou P.C., « Mimicking the human ear », IEEE signal processing magazine, pp 101-130, 1998.
- [62] Maazouz M., Batel N. (encadreur), « Extraction des formants de la parole par la méthode prédictive avec confrontation au sonagramme à large bande », mémoire d'ingénieur d'état en électronique, C.U.Médèa, 2004.
- [63] McDermott H., MKay C., Vandali A., "A new portable sound processor for the University of Melbourne/Nucleus Limited multielectrode cochlear implant", Journal of the Acoustical Society of America, vol. 91, pp. 3367-3371, 1992.
- [64] Mehenni M., « Optimisation des éléments d'une chaîne de télémesure », Thèse de doctorat d'état, Ecole Nationale Polytechnique d'Alger, 1992.
- [65] Mehenni M., Yvroud E., Abignoli M., Marchal C., « Alimentation par champ électromagnétique externe pour télémesure : conception d'un banc expérimental », Ed. RBM, vol.14, n°2, 1992.
- [66] Memmi F., Batel N. (encadreur), « Visualisation par capture HF des paramètres de stimulation auditive chez un sourd », mémoire d'ingénieur d'état en électronique, C.U.Médèa, 2006.
- [67] Meyer Y., « Les ondelettes - algorithmes et applications », Armand colin, 2^{ème} édition, 1994.
- [68] Morsli A., Tahraoui L., Batel N. (encadreur), « Analyse spectrale d'un signal audio-fréquence », mémoire d'ingénieur en électrotechnique au C.U. de Médèa, 2001.
- [69] Nasar S.A., Paul C.R., Whites K.W., « Introduction to electromagnetic fields », Mac Graw-Hill, 1997.
- [70] Ouhabi A., « fondements théoriques du traitement du signal », Collection scientifique de l'ingénieur électronique, 1993.

- [71] Patrick J., Clark G., "The Nucleus 22-channel cochlear implant system", *Ear and Hearing*, vol. 12, pp. 3-9, 1991.
- [72] Press W.H., Tenkolsky S.A., Vetterling W.T., Flannery B.P., « numerical recipes in C », Cambridge University press, second edition, 1997.
- [73] Prieur G., « Radiofréquences et télécommunications analogiques », Ed. Masson, 1996.
- [74] Rabiner L. R., Schafer R. W., "digital processing of speech signals", Prentice-Hall, Inc, 1978.
- [75] Resweber E. J., « A DSP GMSK Modem for Mobitex and other Wireless Infrastructures », note technique sur site www.ti.com.
- [76] Rosen S., Walliker J., Brimacombe J., Edgerton B., "Prosodic and segmental aspects of speech perception with the House/3M single-channel implant", *Journal of Speech and Hearing Research*, vol. 32, pp. 93-111, 1989.
- [77] Sahraoui T., Douifi A., « Etude comparative entre la transformation de FOURIER, la transformation de FOURIER à court terme et la transformation en ondelette : application à l'analyse des HITS », mémoire d'ingénieur en électrotechnique au C.U. de Médéa, 2000.
- [78] Schwan H.P., « Electrical and acoustic properties of biological materials and biomedical applications », *IEEE trans. biomedical eng.*, vol. 31, n°12, pp 872-878, 1984.
- [79] Schwan H.P., Foster K.R., « RF-field interactions with biological systems: electrical properties and biophysical mechanisms », *Proceedings of the IEEE*, vol. 68, n°1, pp 104-113, 1980.
- [80] Seligman P., McDermott H., "Architecture of the Spectra 22 speech processor", *Annals of ORL*, pp. 139-141, 1995.
- [81] Seligman P., Patrick J., Tong Y., Clark G., "A signal processor for a multiple-electrode hearing prosthesis", *Acta Otolaryngologica*, pp. 135-139, 1984.
- [82] Servin C., « Réseaux et Télécommunications », Dunod, 2003.
- [83] Shannon R., "Multichannel electrical stimulation of the auditory nerve in man", *Hearing Research*, vol. 11, pp. 157-189, 1983.
- [84] Skinner M., Clark G., Whitford L., Seligman P., Staller S., "Evaluation of a new spectral peak coding strategy for the Nucleus 22 channel cochlear implant system", *American Journal of Otology*, vol. 15, pp. 15-27, 1994.
- [85] Soma M., Galbraith D.C, White R.L., « Radio-frequency coils in implantable devices : Misalignment analysis and design procedure », *IEEE Trans. on Biomedical Eng.*, vol.34, n°4, 1987.
- [86] Sterkers O., Ferrary E., Tran ba huy P., « Physiologie des liquides labyrinthiques », *Encyclopédie médico-chirurgicale*, Paris ORL 20030 B10, 1986.
- [87] Stuchly M.A., Stuchly S.S., « Dielectric properties of biological substances », *Journal of microwave power*, vol. 15, n°1, pp 19-26, 1980.
- [88] Terman F.E., « Radio Engineers handbook », Mc Graw-Hill, 1943.
- [89] Texas Instruments, « Code Composer user's guide », SPRU296, 1999.
- [90] Texas instruments, « TMS320C5x DSP starter kit user's guide », SPRu101A, 1996.
- [91] Texas Instruments, « TMS320C6000 Chip Support Library API user's guide », SPRU401C, 2001.
- [92] Texas Instrument, "TMS320C6000 CPU and instruction set reference guide", SPRU189F, 2000.
- [93] Texas Instrument, "TMS320C6000 peripherals reference guide", SPRU190D, 2001.
- [94] Texas Instrument, "TMS320C6000 programmer's guide", SPRU198F, 2001.
- [95] Texas Instruments, « TMS320C6000 DSP/BIOS Application Programming Interface reference guide », SPRU403C, 2001.
- [96] Texas Instruments, « TMS320 DSP/BIOS user's guide », SPRU423A, 2001.
- [97] Texas instruments, « user's guide TMS320C5x », SPRu56c Digital Signal Processing

Products, 1997.

- [98] Texas Instruments, « User's manual of PROLOGIC », 1989.
- [99] Tischer M., « la bible du PC : programmation système », Ed. da, Data Becker, 5^{ème} édition.
- [100] Tran ba huy P., Bastian D., Ohresser M., « Anatomie de l'oreille interne », Encyclopédie médico-chirurgicale, Paris ORL 20020 A10, 1990.
- [101] Truchetet F., « Ondelettes pour le signal numérique », Ed. Hermes, 1998.
- [102] Tye-Murray N., Lowder M., Tyler R., "Comparison of the F0/F2 and F0/F1/F2 processing strategies for the cochlear implant", *Ear and Hearing*, vol. 11, pp. 195-200, 1990.
- [103] Tyler R., "Speech perception by children", *Cochlear Implants: Audiological Foundations*, Singular Publishing Group inc, Tyler ed., pp. 191-256, 1993.
- [104] Van den Honert C., Stypulkowski P., "Single fiber mapping of spatial excitation patterns in the electrically stimulated nerve", *Hearing Research*, vol. 29, pp. 195-206, 1987.
- [105] Vern J. L., « Génération de séquences pseudo-aléatoires », *Radio Plans* n° 555, 1995.
- [106] Wallenberger E., Battmer R., "Comparative speech recognition results in eight subjects using two different coding strategies with the Nucleus 22 channel cochlear implant", *British Journal of Audiology*, vol. 25, pp. 371-380, 1991.
- [107] White M., Merzenich M., Gardi J., "Multichannel cochlear implants: Channel interactions and processor design", *Archives of Otolaryngology*, vol. 110, pp. 493-501, 1984.
- [108] Wilson B., Finley C., Lawson D., "Better speech recognition with cochlear implants", *Nature*, vol. 352, pp. 236-238, 1991.
- [109] Wilson B., Lawson D., Zerbi M., "Advances in coding strategies for cochlear implants", *Advances in Otolaryngology: Head and Neck Surgery*, vol. 9, pp. 105-129, 1995.
- [110] www.allhear.com, "AllHear Cochlear Implant", AllHear Inc., Aurora, OR.
- [111] www.cochlear.com, "Nucleus Multichannel Cochlear Implant System", Cochlear Corporation, Englewood, CO.
- [112] www.cochlearimplant.com, "Clarion Multi-Strategy Cochlear Implant System", Advanced Bionics Corporation, Sylmar, CA.
- [113] www.cochlearimplants.com, "Med-El Cochlear Implant", Med-El Corporation, Research Triangle Park, NC.
- [114] www.latticesemi.com: Site de la firme Lattice sur les GAL
- [115] www.mxcom.com, « MSK and its Application to Wireless Data Transmission », MX.COM, Inc. at Winston-Salem USA, 1997.
- [116] www.ti.com: Site de la firme Texas Instruments.
- [117] Zeng F., Shannon R., "Loudness balance between acoustic and electric stimulation", *Hearing Research*, vol. 60, pp. 231-235, 1992.
- [118] Zierhofer C., Hochmair D., Hochmair E., "Electronic design of a cochlear implant for multichannel high-rate pulsatile stimulation strategies", *IEEE Transactions on Rehabilitation Engineering*, vol. 3, n°1, pp. 112-116, 1995.
- [119] Zierhofer C., Peter O., Bril S., Pohl P., Hochmair D. E., "A multichannel cochlear implant system for high-rate pulsatile stimulation strategies", *Advances in cochlear implants*, Hochmair eds., pp. 204-207, 1994.