

ECOLE NATIONALE POLYTECHNIQUE

DEPARTEMENT : ELECTRONIQUE

المدرسة الوطنية المتعددة التقنيات  
BIBLIOTHEQUE — المكتبة  
Ecole Nationale Polytechnique

**PROJET DE FIN D'ETUDES**

**S U J E T**

CONTRIBUTION A LA  
REALISATION D'UN  
VOCODEUR A CANAUX

Proposé par :

Mr. BOUSSEKSOU

Etudié par :

S. MEKAOUI

K. SEMMAR

Dirigé par :

Mr. BOUSSEKSOU

PROMOTION : JUIN 87

ECOLE NATIONALE POLYTECHNIQUE

DEPARTEMENT D'ELECTRONIQUE

PROJET DE FIN D'ETUDE

المدرسة الوطنية المتعددة التقنيات  
المكتبة — BIBLIOTHEQUE  
Ecole Nationale Polytechnique

THEME

CONTRIBUTION A LA REALISATION  
D'UN VOCODEUR A CANAUX

DIRIGE PAR :Mr B.BOUSSEKSOU

REDIGE PAR:Mr S.MEKAQUI  
Mr K.SEMMAR

PROMOTION JUIN 1987

\* REMERCIEMENTS \*

---

NOUS TENONS A EXPRIMER NOS PLUS VIFS REMERCIEMENTS A NOTRE PROMOTEUR MONSIEUR BOUSSEKSOU POUR SON AIDE , SES CONSEILS ET SON SUIVI TOUT AU LONG DE L'ELABORATION DE CE PROJET DE FIN D'ETUDE. ILS VONT EGALEMENT A NOS PARENTS ,QUI ONT CONSENTI TANT D'EFFORTS ET DE SACRIFICES .

QUE TOUS CEUX QUI ONT CONTRIBUE DE PRES OU DE LOIN A L'EDIFICATION DE CE TRAVAIL , TROUVENT EN CES QUELQUES LIGNES L'EXPRESSION DE NOTRE PROFONDE GRATITUDE .

S. MEKAOUI

K. SEMMAR

\* DEDICACES \*

A la memoire de mes grands parents

Je dedie ce modeste travail

A mon pere

A ma mere

A mon beau frere : LEBAH ABDELKADER

A toute ma famille

A tous mes ami(e)s

Avec mes remerciements

Slimane .MEKAOUI

\* DEDICACES \*

---

A la memoire de mes grands parents

Je dedie ce modeste travail

A mon pere et a ma mere

A mes freres Youcef et Touffik

A mes neveux Nadir Salim Djazil

A ma niece Sihem

A monsieur et madame A.MEKAQUI

A tous mes amis en particulier Hakim

A toute ma famille

Avec mes remerciements

Karim .SEMMAR

المدرسة الوطنية المتعددة التقنيات  
BIBLIOTHEQUE — المكتبة  
Ecole Nationale Polytechnique

\*\*\*\*\*  
\*\*  
\*\* PLAN \*\*  
\*\*  
\*\*\*\*\*

CHAPITRE I      PHYSIOLOGIE DE LA PAROLE ET ETUDE DE LA PRODUCTION  
DES SONS

I.1 - DESCRIPTION DE L'APPAREIL PHONATOIRE

I.1.1 - PARTIE SUBGLOTTIQUE

I.1.2 - PARTIE GLOTTIQUE OU LARYNX

- a - LE CARTILAGE ORICOIDE
- b - LE CARTILAGE THYROIDE
- c - LE CARTILAGE EPIGLOTTIQUE
- d - LES CARTILAGES ARYTENOIDES
- e - L'ARTICULATION

I.1.3 - PARTIE SUPRAGLOTTIQUE

- a - LE CONDUIT VOCAL
- b - LE CONDUIT NASAL
- c - ROLES DE CES DEUX CAVITES

I.1.4 - PRINCIPE DE LA PRODUCTION DES SONS

- a - LES SONS VOISES
- b - LES SONS SOURDS OU LES " SOURDES "
- c - LES PARAMETRES ARTICULATOIRES

I.2 - LES DIFFERENTS SONS DE LA PAROLE

I.2.1 - LES VOYELLES

- a - LES VOYELLES ORALES
- b - LES DIFFERENTES VOYELLES ORALES DU FRANCAIS
- c - LES VOYELLES NASALES
- d - LES VOYELLES NASALES DU FRANCAIS

I.2.2 - LES CONSONNES

- a - LES CONSONNES DITES " CONSTRUCTIVES "
- b - LES CONSTRUCTIONS DU FRANCAIS
- c - LES CONSONNES DITES " OCCLUSIVES "
- d - LES LIQUIDES

- e - LES SEMI-VOYELLES OU LES SEMI-CONSONNES
- f - CONCLUSION
  
- II - SONS ELEMENTAIRES ET PHONEMES
- II.1 - CARACTERISTIQUES ACOUSTIQUES PRINCIPALES  
DES SONS DE LA PAROLE
- II.1.1 - TIMBRES ET SONS
- II.1.2 - PRESENTATION DU SIGNAL DE LA PAROLE
- II.1.3 - TIMBRE EN REGIME TRANSITOIRE
- II.1.4 - PARAMETRES ACOUSTIQUES
  - a - L'INTENSITE ACOUSTIQUE
  - b - NIVEAU D'INTENSITE
- II.2 - ROLE DE L'OREILLE
- II.2.1 - INTRODUCTION
- II.2.2 - L'APPAREIL AUDITIF PERIPHERIQUE
  - a - DESCRIPTION GENERALE
  - b - L'OREILLE EXTERNE
  - c - L'OREILLE MOYENNE
  - d - L'OREILLE INTERNE
- II.2.3 - REACTION DE L'OREILLE AU SON
  - a - SENSIBILITE DE L'OREILLE A L'INTENSITE
  - b - SEUIL DIFFERENTIEL D'INTENSITE
  - c - SENSIBILITE DE L'OREILLE A LA FREQUENCE
- III - THEORIE FORMANTIQUE
- III.1 - SON A UN FORMANT
- III.2.- SON A DEUX FORMANTS
- IV - INTELLIGIBILITE DE LA PAROLE
- IV.1 - INTRODUCTION
- IV.2 - METHODES DITES " SUBJECTIVES "



IV.3 - METHODES DITES " OBJECTIVES "

V - VARIABILITE DU SIGNAL DE LA PAROLE

CHAPITRE II TRAITEMENT DU SIGNAL DE LA PAROLE

I - INTRODUCTION

II - ANALYSE SPECTRALE

II.1 - INTRODUCTION

II.2 - TRANSFORMEE DE FOURIER

II.2.1 - TRANSFORMEE DE FOURIER DES SIGNAUX  
PERIODIQUES

II.2.2 - TRANSFORMEE DE FOURIER DES SIGNAUX  
APERIODIQUES

II.3 - PUISSANCE ET ENERGIE DES SIGNAUX

II.3.1 - INTRODUCTION

II.3.2 - PUISSANCE TEMPORELLE

a - PUISSANCE INSTANTANEE D'UN SIGNAL

b - PUISSANCE MOYENNE A UNE DUREE T

II.3.3 - ENERGIE D'UN SIGNAL

II.4 - ANALYSE NUMERIQUE

II.4.1 - INTRODUCTION

II.4.2 - THEOREME

II.4.3 - TRANSFORMEE DE FOURIER DISCRETE

II.4.4 - TRANSFORMEE DE FOURIER RAPIDE

III - ANALYSE CEPSTRALE

III.1 - INTRODUCTION

III.2 - TRAITEMENT HOMOMORPHIQUE DES SIGNAUX

III.2.1 - INTRODUCTION

III.2.2 - FORME CANONIQUE

III.2.3 - SYSTEMES CONVOLUTIFS

III.3 - FORMULATION

III.4 - CONCLUSION

IV - FILTRAGES

IV.1 - FILTRES NUMERIQUES

IV.1.1 - DEFINITION

IV.1.2 - STRUCTURE RECURSIVE

a - DEFINITION

b - MISE EN EQUATION

c - MODELISATION

d - FONCTION DE TRANSFERT

IV.1.3 - STRUCTURE NON RECURSIVE

a - DEFINITION

b - MISE EN EQUATION

c - MODELISATION

d - FONCTION DE TRANSFERT

IV.1.4 - LA PREDICTION LINEAIRE

a - DEFINITION

b - METHODE D'AUTO-CORRELATION

c - METHODE DU FILTRE EN TREILLIS

IV.2 - FILTRES ANALOGIQUES

V - CONCLUSION

CHAPITRE III LES VOCODEURS A TRAVERS L'HISTOIRE

I - INTRODUCTION

II - GENERALITES SUR LES VOCODEURS

II.1 - PRINCIPE DE BASE

II.2 - LES PRINCIPAUX TYPES DE VOCODEURS

II.2.1 - LE VOCODEUR A FORMANT

a - PRINCIPE

b - METHODES D'ANALYSE

c - LES SYNTHETISEURS A FORMANTS

#### 11.2.2 - LE VOCODEUR A PREDICTION LINEAIRE

a - PRINCIPE

b - L'ANALYSE

c - LA SYNTHESE

#### 11.2.3 - LE VOCODEUR A CANAUX

a - PRINCIPE

b - L'ANALYSE

c - LA SYNTHESE

### CHAPITRE IV METHODES GLOBALES POUR LA RECONNAISSANCE ET LA SYNTHESE DE LA PAROLE

#### I - LA RECONNAISSANCE DE LA PAROLE

##### 1.1 - DEFINITION

##### 1.2 - LE PROBLEME TEL QU'IL SE POSE

##### 1.3 - TECHNIQUES DE RECONNAISSANCE

##### 1.4 - APPRENTISSAGE DES SYSTEMES

##### 1.5 - LES PROBLEMES LIES AUX DEUX TYPES D'APPROCHES

a - L'APPROCHE GLOBALE

b - L'APPROCHE ANALYTIQUE

##### 1.6 - DESCRIPTION D'UN SYSTEME DE RECONNAISSANCE GLOBALE

###### 1.6.1 - L'ETAPE D'APPRENTISSAGE

###### 1.6.2 - L'ETAPE DE RECONNAISSANCE

##### 1.7.- METHODES MATHEMATiques

###### 1.7.1 - ETABLISSEMENT D'UNE DISTANCE ENTRE BLOCS DE PARAMETRES

#### II - SYNTHESE DE LA PAROLE

##### 11.1 - DEFINITION

11.2 - LES DIFFERENTES TECHNIQUES DE SYNTHÈSE

CHAPITRE V ETUDE DE LA REALISATION D'UN VOCODEUR A CANAUX

INTRODUCTION

I - ETUDE D'UN CANAL

1.1 - ETUDE D'UN CANAL

1.1.1 - ETUDE D'UN FILTRAGE

1.1.2 - NOTIONS SUR LES FILTRES ACTIFS

1.1.3 - PRINCIPAUX TYPES DE FILTRES

a - STRUCTURE DE BUTTERWORTH

b - STRUCTURE DE TCHEBYTCHEV

1.1.4 - PROTOTYPE PASSE BAS DE BUTTERWORTH

1.1.5 - PROTOTYPE PASSE HAUT DE BUTTERWORTH

1.1.6 - FILTRE PASSE BANDE ADOPTE

1.2 - DETECTION OU REDRESSEMENT

1.3 - SCHEMA GENERAL D'UN CANAL

1.4 - CHOIX DE L'ECHELONNEMENT DES FILTRES

1.4.1 - TEST DE COMPARAISON ET CHOIX  
DE L'ECHELONNEMENT

1.5 - NUMERISATION DU SIGNAL

1.5.1 - ECHANTILLONNAGE ET BLOQUAGE

a - ECHANTILLONNAGE

b - BLOQUAGE

c - ETUDE DU LF 398

d - LA CONVERSION ANALOGIQUE NUMERIQUE

II - ETUDE ET REALISATION DU DETECTEUR DE MELODIE

11.1 - SYNOPTIQUE

11.2 - FILTRAGE

11.2.1 - CALCUL DU FILTRE PASSE BAS D'ORDRE 2

11.2.2 - DETERMINATION DU PASSE HAUT D'ORDRE 2

11.3 - ETUDE DU BLOC COMPRESSEUR

11.3.1 - INTRODUCTION

11.3.2 - ETUDE DU FONCTIONNEMENT

11.4 - ETUDE DU BLOC DETECTEUR DE CRETES

11.4.1 - INTRODUCTION

11.4.2 - FONCTIONNEMENT DU DETECTEUR DE CRETES

111 - STOCKAGE DES DONNEES ET INTERFACAGE

111.1 - INTRODUCTION

111.2 - LE PIA 6821

111.2.1 - PRESENTATION

111.2.2 - SYNOPTIQUE

a - DESCRIPTION

b - ORGANIGRAMME

c - PROGRAMME

IV - DETECTEUR DE VOISEMENT

IV.1 - INTRODUCTION

IV.2 - SYNOPTIQUE

IV.3 - FONCTIONNEMENT

IV.4 - DECISION

CHAPITRE VI CALCULS TESTS ET REALISATION

I - CALCULS

1.1 - COMPOSANTS CONSTITUANTS LE BANC ANALYSEUR

1.2 - COMPOSANTS CONSTITUANTS LE DETECTEUR DE MELODIE

1.2.1 - FILTRAGE

1.2.2 - COMPRESSEUR

1.2.3 - DETECTEUR DE CRETES

11 - TESTS

II.1 - FILTRAGE

II.2 - COMPRESSEUR

II.3 - DETECTEUR DE CRETES

III - REALISATION

III.1 - SCHEMA ELECTRIQUE DU DETECTEUR DE MELODIE

III.2 - CIRCUIT IMPRIME ET SCHEMA D'IMPLANTATION

III.3 - CIRCUIT ELECTRIQUE GLOBAL DE L'ANALYSEUR

CONCLUSIONS GENERALES

## INTRODUCTION :

De Wolfgang Von KEMPLEN 1791 a Homer DUDLEY 1939 la machine parle. C'est un fait acquis ,de nos jours elle arrive a faire des phrases entieres pour une autonomie d'au moins cinq minutes .(Systeme d'annonces vocales de SILEC DSI ) .Le traitement automatique de la parole a pris un essor prodigieux ,et il devient presque normal aujourd'hui de parler de robots parlants et de robots capables de comprendre puis d'executer des ordres donnees . Mais comment s'y prennent-ils pour restituer la parole apres l'avoir entendue et comprise .

En realite cela tient d'un trilogisme constituant le fondement meme du traitement automatique de la parole .Il s'agirait simplement de donner a la machine une oreille pour entendre le message une partie de cerveau associee pour le comprendre ,et une bouche pour repondre Mais ceci n'est pas chose aisee , et reste encore au stade de la recherche ,le gros probleme etant de realiser ces trois operations en temps reel (ou record pour la machine ) afin de s'approcher de l'ideal (modele humain ) .

C'est la l'objectif que se fixent les chercheurs en creant le vocodeur a canaux dont le but est double car en plus d'une parole synthetique de bonne qualite, il permet la reduction du debit de cette derniere ,lors de la transmission telephonique par exemple .

Le present travail consiste en une modeste contribution a la realisation d'un vocodeur a canaux .

Ce dernier est essentiellement constitue de trois parties qui sont:

- L'analyseur
- Le reconnaisseur
- Le synthetiseur

Le plan de ce travail comporte six chapitres dont quatre sont reserves a l'etude theorique proprement dite du vocodeur a canaux ,et les deux dernieres non moindres a la realisation du detecteur de melodie .

- C H A P I T R E I -

PHYSIOLOGIE DE LA PAROLE ET ETUDE DE LA PRODUCTION DES SONS.

I) 1. Description de l'appareil phonatoire :

D'un point de vue fonctionnel, on divise l'appareil phonatoire en 3 grandes parties :  
- Partie subglottique.  
- Partie glottique.  
- Partie supraglottique.

1.1 Partie subglottique.

Elle est constituée des poumons et de la trachée artère, elle est à l'origine de la parole par envoi d'air vers les parties glottiques. Elle sert évidemment à la respiration. Elle constitue ainsi la source d'énergie de la parole ( voir schéma : 1 ).

1.2 Partie glottique ou larynx : ( schéma : 2 et 3 ).

Le larynx est un organe de phonation, faisant partie des voies ( AERIENNES ) respiratoires et situé entre le pharynx et la trachée. C'est un conduit cartilagineux recouvert intérieurement de muqueuse et comprenant des organes vibratoires, qui sont les cordes vocales.

Le larynx est un ensemble assez complexe de neuf cartilages et de muscles, permettant des configurations de mouvements variés et complexes. Ces mouvements sont l'élément essentiel dans le processus de la production de la parole, car déterminant le type de sons produits. Parmi ces neuf cartilages, trois qui sont les plus importants.

- Le cartilage thyroïdien : ou thyroïde située vers l'avant, et dont la partie antérieure constitue la pomme d'Adam.
- Les deux cartilages aryténoïdes situés vers l'arrière de la glotte, c'est sur cet ensemble de trois cartilages que viennent s'appuyer les cordes vocales. ( voir schéma 1 ).

cordes

- Les cordes vocales sont des replis blancs épais tendus horizontalement et qui selon leur position s'exposent plus ou moins au passage de l'air venant des poumons ( voir schéma 2 ).

Leur ensemble est assez complexe, mais en gros elles sont constituées d'un muscle qui permet d'en contrôler la tension interne.

Le schéma deux(2), montre clairement la position des cartilages du larynx, on distingue ainsi :

- a) Le cartilage cricoïde :  
Situé à la partie inférieure du larynx au dessus de la trachée ( à une forme de bague ). Il s'articule avec le cartilage thyroïde et avec les cartilages aryténoïdes ( voir schéma n° 2 ).
- b) Le cartilage thyroïde : ( voir schéma n° 2 ).  
C'est le plus volumineux, est formé de deux lames quadrilatères qui s'unissent en avant formant la pomme d'Adam. Ces lames se prolongent en arrière par des cornes supérieures et inférieures. Ces dernières s'articulent avec le cartilage cricoïde.
- c) Le cartilage épiglottique :  
Formant l'épiglottique, c'est une lame mince et élastique de forme ovale, située dans l'angle rentrant du cartilage thyroïde est fixée à la base de la langue, l'épiglotte obture l'orifice supérieure du larynx pendant la déglutition.



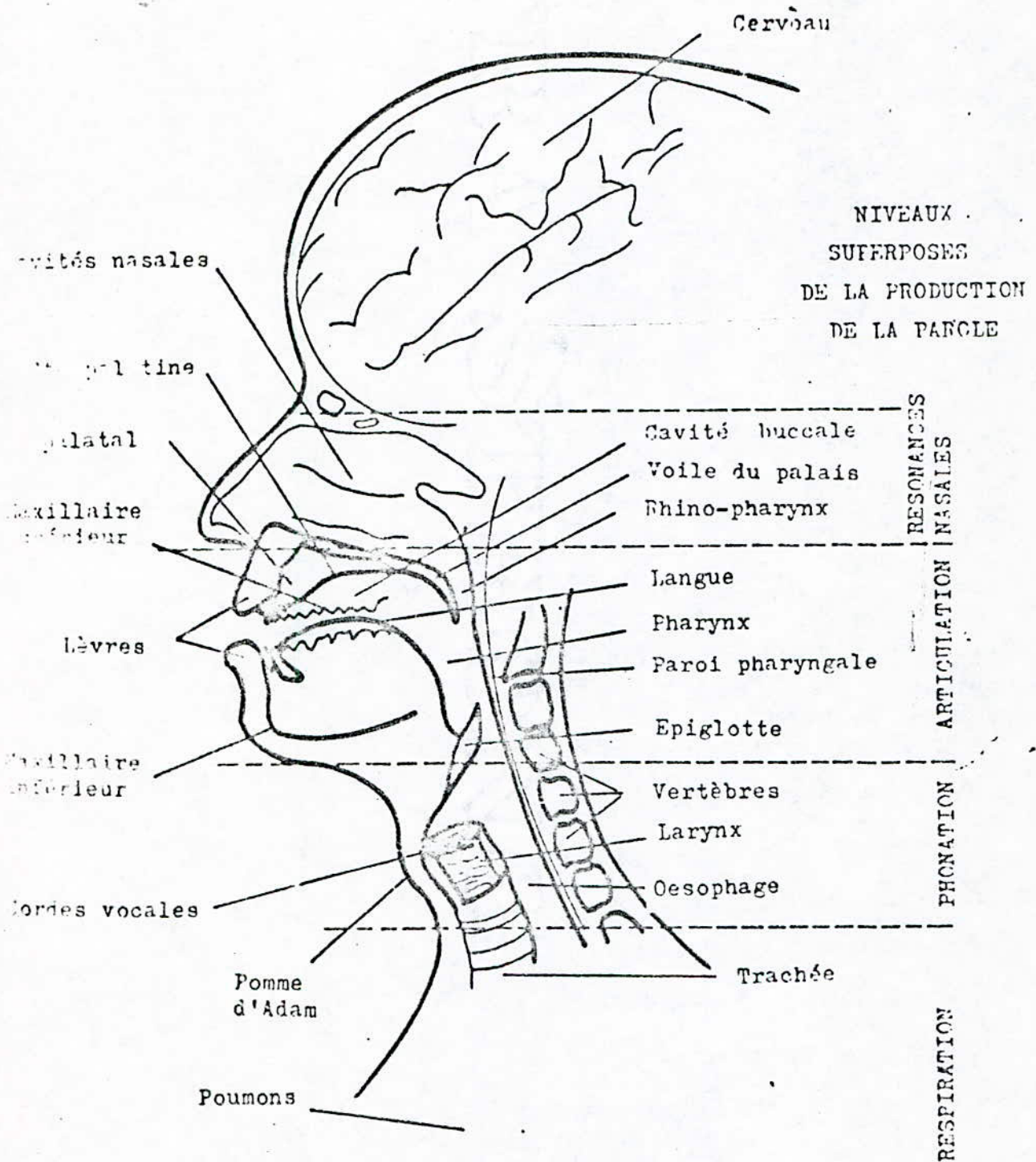


Figure 2-1 - Coupe médiane schématique des organes de la parole



d) Les cartilages arytenoïdes :

Ils sont au nombre de deux, ils ont la forme d'une pyramide triangulaire, sommet supérieur dont la base repose sur le cartilage cricoïde. Ils présentent une saillie dirigée en avant l'apophyse ou s'ingère la corde vocale.

e) L'articulation :

L'articulation du larynx se fait grâce aux différents cartilages qui le composent, en s'articulant entre-eux grâce aux ligaments et aux membranes qui les unissent

- Au niveau du larynx existent aussi des muscles qui assurent la mobilité des cartilages.

- La configuration intérieure du larynx exige aussi la présence d'une muqueuse ( tissu épithélial ) qui tapisse tous les cartilages, ligaments et muscles plus ou moins adhérente aux organes sous-jacents et pouvant se décoller en présence d'un œdème en certains points provoquant une obstruction et une dyspnée ( difficulté respiratoire aiguë ) de pronostic sévère chez l'enfant.

Elle exige aussi une vascularisation et une innervation. La vascularisation est assurée par les artères du larynx qui proviennent des artères thyroïdiennes. Les veines se jettent dans la jugulaire interne.

- L'innervation :

Les nerfs sont au nombre de deux, de chaque côté du larynx.

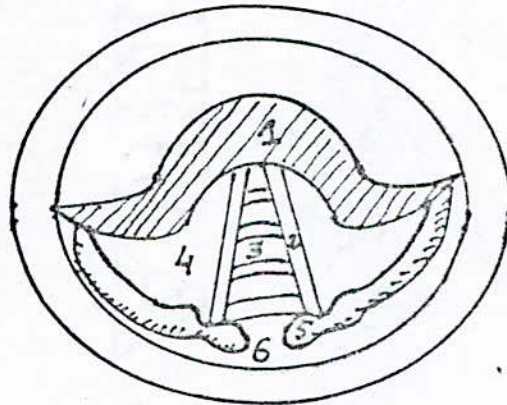
On distingue :

- Le nerf laryngé supérieur est sensitif et moteur du crico-thyroïdien. Le nerf laryngé inférieur ou récurrent est le nerf moteur de tous les autres muscles, né du pneumogastrique.

Il passe à droite sous l'artère sous-clavière et à gauche sous la crosse de l'aorte pour remonter de chaque côté du larynx.

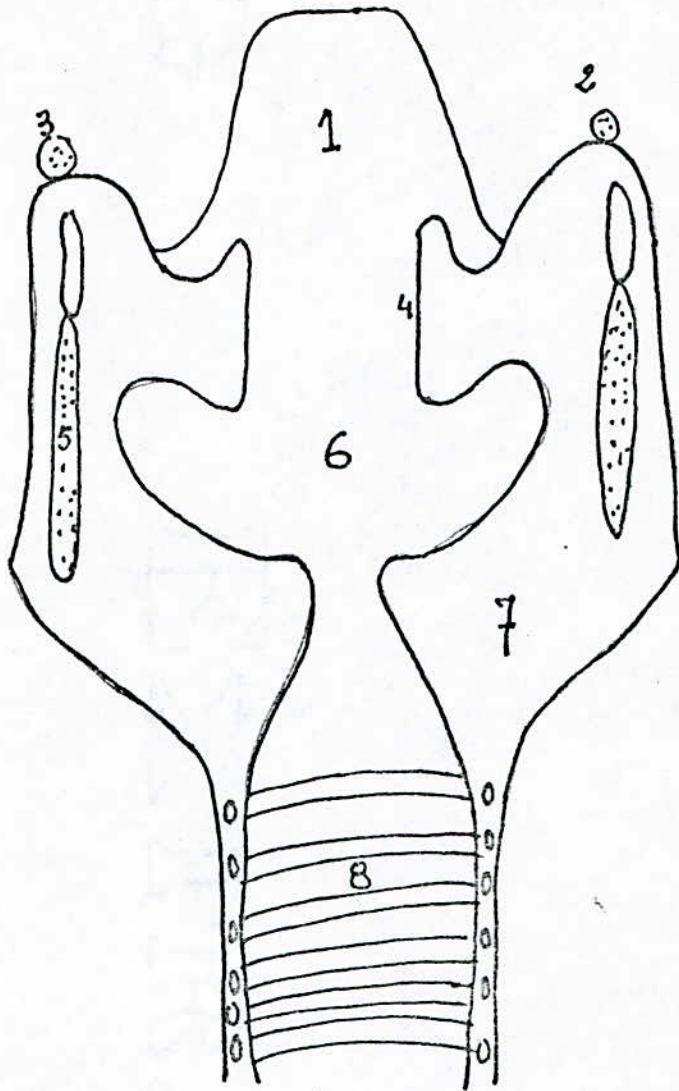
Lesion → Toute compression ou lésion des nerfs récurrents, entraîne une paralysie des cordes vocales et une aphonie ou dysphonie.

SCHEMA DE LA VUE LARYNGOSCOPIQUE DE LA GLOTTE : (2)



- 1- Epiglotte.
- 2- Cordes vocales inférieures.
- 3- Glotte.
- 4- Cordes vocales supérieures.
- 5- Tubercules de Santorini.
- 6- Échancre inter-aryténoïdien.

SCHEMA DU LARYNX : (3)



- 1- Epiglote.
- 2- Ostyroïde.
- 3- Cartilage épiglottique.
- 4- Cordes vocales supérieures.
- 5- Cartilage thyroïde.
- 6- Glotte.
- 7- Muscle thyro-aryténoïdien.
- 8- Cavité de la trachée.

### 1.3 La partie supraglottique.

- Elle est constituée par les cavités buccales et nasales, en appelle
- Les cavités buccales → le conduit vocal.
  - Les cavités nasales → le conduit nasal.

#### a) Le conduit vocal :

Commence par le larynx et se termine aux lèvres, sa longueur varie entre 12 cm et 20 cm. La section et sa longueur peuvent être constamment modifiées par le mouvement de haut en bas du larynx, par le mouvement de la langue, des lèvres et des mâchoires ( voir schéma récapitulatif ). ( N° 4 )

#### b) Le conduit nasal : ( voir schéma récapitulatif ) ( Schema:4 )

Il est constitué par les deux cavités nasales situées au dessus du velum ( luvette ). Sa longueur et sa section restent sensiblement constantes.

#### c) Rôles de ces deux cavités :

Essentiellement un effet de " Modulation " de " l'onde glottique ". La forme de ces cavités et leur couplage par l'intermédiaire du velum, qui peut boucher ou libérer l'entrée du conduit nasal, déterminent la nature du son produit "( Voyelles, consonnes, ... nasalités, ~~...~~ )".

### 1.4 Principe de la production des sons :

#### a) Les sons voisés.

Ce sont les sons pour lesquels les cordes vocales entrent en vibration, l'énergie étant fournie par les ~~cordes~~<sup>courents</sup> à travers la trachée artère sous forme de pression d'air. Pour la production des sons voisés, en début de phonation, les cordes vocales sont accélérées et s'opposent ainsi au passage de l'air. Sous l'action des forces de pression d'air, les cordes vocales vont se séparer l'une de l'autre, laissant ainsi entre elles un orifice que l'on appelle glotte.

Sous l'action de forces élastiques et d'un phénomène aérodynamique, " les cordes vocales " entrent en vibration ouvrant et fermant ainsi à intervalles réguliers la " glotte ".

C'est ainsi qu'elles laissent échapper vers les parties supraglottiques, un certain débit, ou flux d'air que l'on appelle   
 signal → " ~~signal~~ glottique ". Les parties supraglottiques moduleront alors le signal glottique pour délivrer le son voisé entendu.

L'accélération plus ou moins étroite des cordes vocales en début de " phonation " détermine le " registre " de la parole.

- Registre de poitrine ( Parole-normale ).

- Registre " FALSETTO " ( chuchètement ) ( faible accélération ).

La théorie qui vient d'être énoncée est la théorie " myoélastique " de la parole qui s'est longtemps opposée à la " neurochronaxique " qui s'avère de nos jours fautive. La première théorie est désormais universellement admise.

#### b) Les sons sourds ou les " sourdes " :

Les sons sourds, ou les " sourdes ", ont pour origine une source de bruit qui excite le conduit vocal. Ces sources de bruits sont de deux types :

##### - Constrictiens :

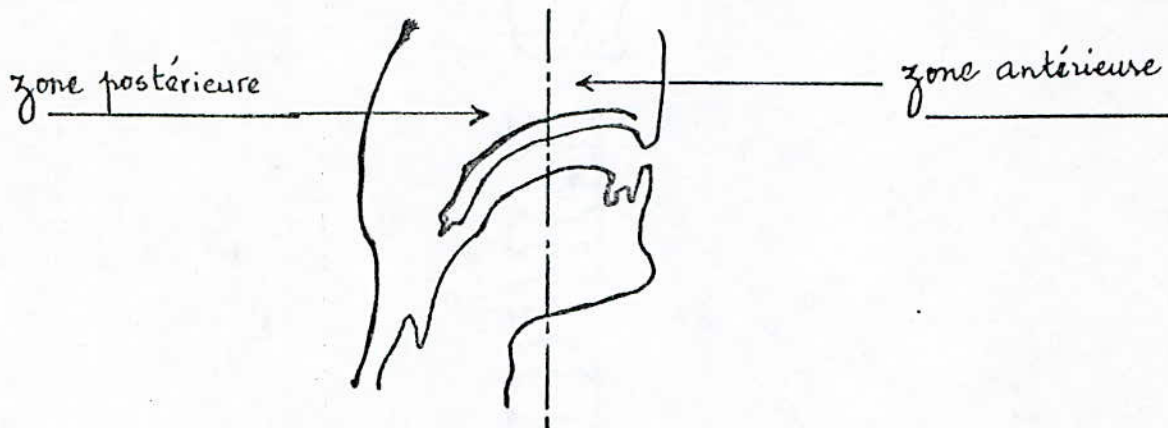
C'est à dire que cela correspond à un rétrécissement " très étroit " de la section du conduit vocal, grâce à la langue et aux lèvres, on assiste alors à un écoulement turbulent de l'air qui est à l'origine du bruit.

- Occlusiens ( occlusiens ), elles consistent en une " obturation du conduit vocal, suivi d'une brutale ouverture. Lorsque le conduit vocal est obturé la pression de l'air croît avec l'obturation. Lors de l'ouverture brutale, l'air s'engouffre dans le passage libéré, produisant ainsi un bruit comparable au bruit produit lorsqu'on débouche une bouteille, on a affaire à ce que l'on appelle une impulsion acoustique. Les cavités orales situées après la source de bruit moduleront le signal de bruit.

d) Les paramètres articulatoires :

Généralement les " phonétiques " caractérisent l'articulation d'un son produit par 5 paramètres articulatoires qui sont :

- Le point d'articulation, l'aperture ou l'ouverture, la labialisation, la nasalité, la latéralité.
- Le point d'articulation : C'est le point où le dôme de la langue est le plus proche de la voûte palatale ( du palais ). Pour préciser la position de ce point, il existe des zones bien précises que nous simplifierons pour nous ramener à deux grandes zones.
  - zone antérieure : c'est à dire vers l'avant du conduit vocal.
  - zone postérieure : c'est à dire vers l'arrière du conduit vocal.



- L'Aperture ou l'ouverture :

C'est le paramètre caractérisant la section du conduit vocal au point d'articulation. Si la section est faible, on parle de voyelle " fermée ". Si la section est grande on parle de voyelle " ouverte ".

- La labialisation :

Ce paramètre caractérise la position des lèvres. Si elles sont projetées vers l'avant, on dit qu'elles sont arrondies cas de la voyelle " O ", sinon elles sont écartées cas de la voyelle " I ".

- La nasalité :

Ce paramètre caractérise la position du velum (lucette) : son ouverture correspond à une nasalité, sa fermeture ne correspond pas à une nasalité.

- La latéralité :

Dans certains cas, le dôme de la langue touche la voûte palatine, sans pour autant obturer le conduit vocal qui se double sur les cotés. cas du [P.]

I ) 2. Les différents sons de la parole :

2.1 Les voyelles :

Les voyelles sont des sons voisés. Elles impliquent donc une vibration des cordes vocales selon que le velum (lucette) est abaissé (conduit nasal couplé avec le conduit oral) ou levé (conduit nasal bouché), on aura des voyelles nasales ou orales - les descriptions faites par la suite ne concerneront que la voix du registre de la poitrine qui correspond à une " phonation normale ".

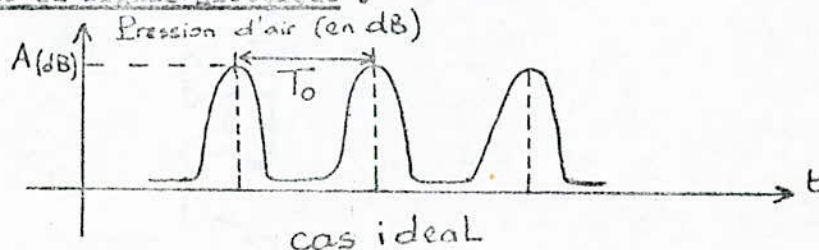
a) Les voyelles orales :

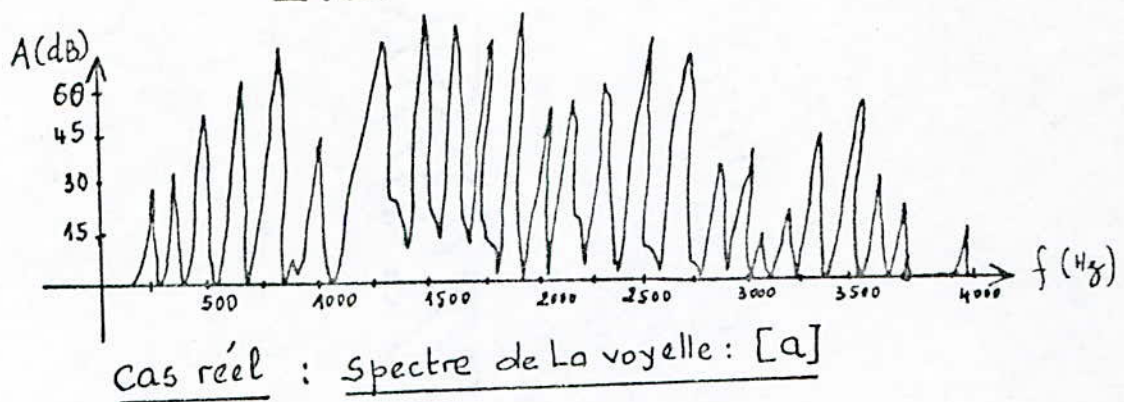
- le signal glottique :

Le mouvement vibratoire des cordes vocales laisse échapper vers les parties supraglottiques un flux d'air qui pour la production de voyelles est quasi-périodique.

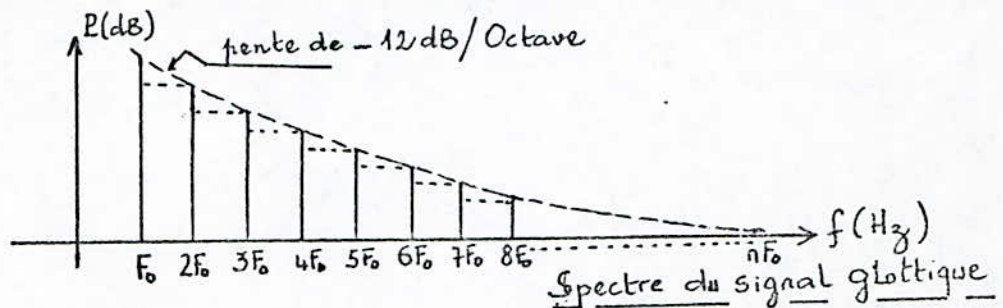
La fréquence du signal glottique ainsi obtenue est appelée fréquence fondamentale "du signal de la parole".

Forme du signal glottique :





une des caractéristiques importantes du signal glottique est son spectre. Il est caractérisé par une pente de  $-12$  dB/Octave, c'est à dire si on double la fréquence, l'intensité de l'harmonique correspondant diminue de  $+12$  dB.

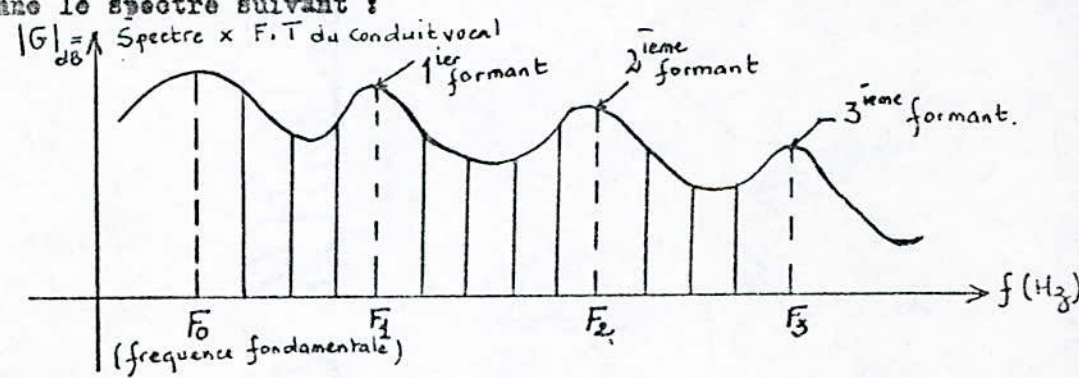


L'intensité du signal glottique est essentiellement fonction de la pression subglottique et de la tension des cordes vocales.

- La fréquence fondamentale  $F_0$  du signal glottique et donc de la parole est essentiellement fonction de la tension des cordes vocales et très faiblement de la pression subglottique.

- La pression subglottique varie entre 2 Pa et 20 Pa, ces valeurs étant mesurées par rapport à la pression atmosphérique.

A première approximation le spectre d'une voyelle orale est le produit du spectre du signal glottique par la fonction de transfert du conduit vocal, ce qui donne le spectre suivant :



Les 3 paramètres privilégiés permettant de décrire la forme du conduit vocal sont le point d'articulation, l'aperture et la labialisation, or, de la forme du conduit vocal dépendent les valeurs des différents formants et on a ainsi les relations suivantes :

\* Si le point d'articulation est antérieur

{	F1 décroît
	F2 croît
	F3 croît

\* Si le point d'articulation est postérieur :  $\left\{ \begin{array}{l} F1 \text{ croit} \\ F2 \text{ décroit} \\ F3 \text{ décroit} \end{array} \right.$

\* Si en avance les lèvres ( en arrondis ) les trois fréquences des trois formants diminuent.

\* Si en a une ouverture grande les formants se rapprochent de ceux de la voyelle neutre théorique [ E ]. Qui serait le signal de la parole qui aurait la forme ( géométrie ) d'un tube cylindrique de " section uniforme ".

b) Les différentes voyelles orales du français :

Le tableau ci-dessous donne les principales voyelles orales de la langue française.

VOYELLES	Que l'en retrouve dans le mot	Valeurs en Hz des deux premiers formants	
		F1 (Hz)	F2 (Hz)
[i]	<u>Li</u> t	F1 = 250 Hz	F2 = 2300 Hz
[e]	<u>é</u> té	F1 = 400 Hz	F2 = 2100 Hz
[ɛ]	ma <u>r</u> ais	F1 = 460 Hz	F2 = 1900 Hz
[y]	U <u>r</u> sule	F1 = 290 Hz	F2 = 1900 Hz
[œ]	<u>P</u> eur (son voisin de [ə] )	F1 = 540 Hz	F2 = 1400 Hz
[ə]	<u>P</u> etit ( Son voisin de [œ] mais souvent plus court )		
[ø]	<u>J</u> eu	F1 = 400 Hz	F2 = 1611 Hz
[ɑ]	<u>P</u> âte	F1 = 650 Hz	F2 = 1200 Hz
[o]	<u>S</u> aule	F1 = 400 Hz	F2 = 1800 Hz
[u]	<u>B</u> ijeu	F1 = 300 Hz	F2 = 800 Hz
[ɑ]	<u>P</u> atte		

T A B L E A U N° 1

Les valeurs des formants données ci-dessus sont approximatives et varient selon le locuteur, mais l'ordre de grandeur est ainsi donné. lorsqu'en modifie la fréquence d'un formant, il se produit alors une sensation de " hauteur formantique " qui n'est pas du même ordre de grandeur que celle de la " hauteur tonale ". Ainsi le [i] nous paraîtra plus aigu que le [e] et que le [e] nous paraîtra plus aigu que le [u].

c) Les voyelles nasales :

Comme pour les voyelles orales, les voyelles nasales impliquent une vibration des cordes vocales ( sens voisés ), mais cette fois l'abaissement du velum couple le conduit nasal et le conduit vocal. Le conduit nasal se comporte lui aussi comme un " RESONNATEUR " avec ces fréquences de résonances, mais produira en parallèle un phénomène d' "antirésonance", on observera une annulation, ou du moins une diminution de l'énergie acoustique à ces fréquences.

- Le couplage du conduit nasal induira donc une déformation du spectre, en va ainsi voir apparaître une déformation du spectre de nasalité à environ ( 250 Hz ). Le premier formant F1 sera plus élevé en fréquence et le plus faible en amplitude, comparé au premier formant de la voyelle orale correspondante.

Enfin les deuxième et 3ème formants F2 et F3 sont légèrement atténués.

d) Voyelles nasales du français :

- [ɛ̃] qu'en trouve dans bien
- [œ̃] qu'en trouve dans brun
- [ɑ̃] qu'en trouve dans chant
- [ɔ̃] qu'en trouve dans salon

2.2 Les consonnes :

Dans la production des consonnes intervient le phénomène de génération de bruit dans le conduit vocal - Si ce phénomène s'accompagne d'une vibration des cordes vocales, les consonnes seront dites "Voisées". Dans le cas contraire elles seront dites "Sourdes". Selon le type de perturbation intervenant dans le conduit vocal on classe les consonnes en différentes catégories.

a) Les consonnes dites " CONSTRUCTIVES ".

Sur un point du conduit vocal on observe un rétrécissement ou " Constriction " de la section du passage de l'air ceci est l'origine de la production d'un bruit car la section est suffisamment faible pour empêcher un écoulement " laminaire " de l'air, l'écoulement est turbulent. Ce bruit va ensuite être modulé par les cavités vocales situées en aval de la constriction - on conçoit donc bien que plus sur le point de constriction est situé vers l'arrière plus le spectre est complexe.

Si cette source de bruit s'ajoute à la source de " voisement " due à la vibration des cordes vocales, on aura donc une constriction voisée.

*n'a pas →* De manière générale le conduit nasal joue un rôle dans la production des " constructions ". En effet l'abaissement du voile du palais rend très difficile l'obtention du seuil de constriction.

b) Les " Constrictives " du français :

Le tableau suivant donne les principales consonnes " constructives du Français, on les classe en deux catégories " constructives sourdes " non voisées, et " Constructives voisées ".

Consonnes	Que l'en retrouve dans	Catégorie
[f]	→ f <u>ameux</u>	} Sourdes
[s]	→ s <u>enteur</u>	
[ʃ]	→ ch <u>at</u>	
[v]	→ v <u>ert</u>	} Voisées
[z]	→ z <u>ebre</u>	
[ʒ]	→ j <u>eur</u>	



Notons que dans le cas des "constrictives voisées", l'intervention de la vibration des cordes vocales diminue la pression d'air avant la constriction et par conséquent diminue le bruit de friction.

c) Les consonnes dites "OCCLUSIVES".

Parmi celles-ci, on en distingue deux catégories.

- Les occlusives orales ou plosives.
- Les occlusives nasales ou les consonnes nasales.

Elles sont produites par l'obturation du conduit oral ou vocal suivi d'une ouverture "brutale" ou "occlusion" sans intervention du conduit nasal. Ce phénomène d'occlusion est appelé "IMPULSION ACOUSTIQUE". On observe généralement trois (3) types d'impulsions acoustiques.

- 1) IMPULSION LABIALE :  
Écartement brutale des lèvres.
- 2) IMPULSION DENTALE :  
Retraction et abaissement de la langue vers l'arrière.
- 3) IMPULSION PALATALE :

Qui correspond à un abaissement brutal du dos de la langue.

De même que pour les "CONSTRUCTIVES", l'impulsion acoustique est modulée essentiellement par les cavités "SUPRAGLOTTIQUES" situées en aval du lieu d'occlusion. Donc plus le lieu d'occlusion sera vers l'arrière plus le spectre sera coloré. L'intensité ( en dB ) des impulsions est liée uniquement à des phénomènes articulatoires donc mécaniques, elle ne peut donc varier dans de grandes proportions. Les "plosives" peuvent être comme les "Constrictives" voisées ou sourdes.

Pour le Français les "OCCLUSIVES" orales ou, principales plosives sont les suivantes :

TABLEAU "9" :

<u>PLOSIVES</u>	<u>Que l'en retrouve dans :</u>	<u>Catégorie</u>
[P]	<u>Pari</u>	} Sourdes
[t]	<u>tenten</u>	
[k]	<u>Caramel</u>	
[b]	<u>benben</u>	} Voisées
[d]	<u>dinde</u>	
[g]	<u>gare</u>	

- Les occlusives nasales ou les consonnes nasales :

On observe une obturation totale du conduit vocal, mais l'air peut s'échapper librement par les cavités nasales ( voile du palais abaissé ). On observe alors la présence de fermants nasaux, de plus le phénomène d'antirésonnance déjà vu lors de l'étude des voyelles nasales intervient aussi ce qui diminue fortement l'énergie spectrale.

Enfin la suppression avant l'occlusion est faible puisque la majeure partie de l'air s'écoule par les cavités nasales. Par conséquent, l'intensité de l'impulsion acoustique est faible.

EXEMPLES :

- [n] comme dans Neuneu
- [m] dans montagne
- [ɲ] dans montagne

d) Les liquides : ( consonnes dites liquides ).

Le problème des liquides est un problème un peu difficile, en effet la production de ces consonnes ne fait pas véritablement intervenir la source de bruit, mais correspondant plutôt à une transition vocale dans le mode résonnant - c'est à dire une modification de la fonction d'air du conduit vocal accompagné d'une vibration des cordes vocales.

e) Les semi-voyelles ou les semi-censennes :

Ce sont les sons netés [w], [ɥ], [j] que l'on trouve respectivement dans LOUIS, LUI, PAILLE, ils correspondent à des transitions rapides du conduit vocal, ce qui les rend assimilables aux censennes, mais dans la mesure où le conduit vocal continue à fonctionner en mode résonnant, on se rapproche alors des voyelles d'où le nom de "semi-voelles - semi-censennes."

f) Conclusion :

L'étude précédente abouti à une classification générale, des différents sons de la parole, le tableau (3) de la page suivante illustre cette classification.

II - Sons élémentaires et phonèmes :

Introduction :

La différence fondamentale entre sons et phonèmes réside dans " l'aspect phonétique ". En effet jusqu'ici la notation des sons sous forme de voyelles ou de censennes entre [ ] crochets indiquait la " réalisation acoustique " de ces voyelles ou censennes, cette notation porte le nom de " Notation phonétique ". Ce point de vue est à distinguer de " l'aspect phonologique " entre / / parenthèses que l'on attribue le plus souvent aux " phonèmes ".

Les phonèmes sont les unités sonores minimales porteuses de signification. Ainsi le signe de durée ( : ) est un phonème dans certaines langues ( Arabe, allemand, ... ) et ne l'est pas en Français.

Le tableau de la page suivante donne les 37 phonèmes de la langue Française. (4) (page n°: 13)

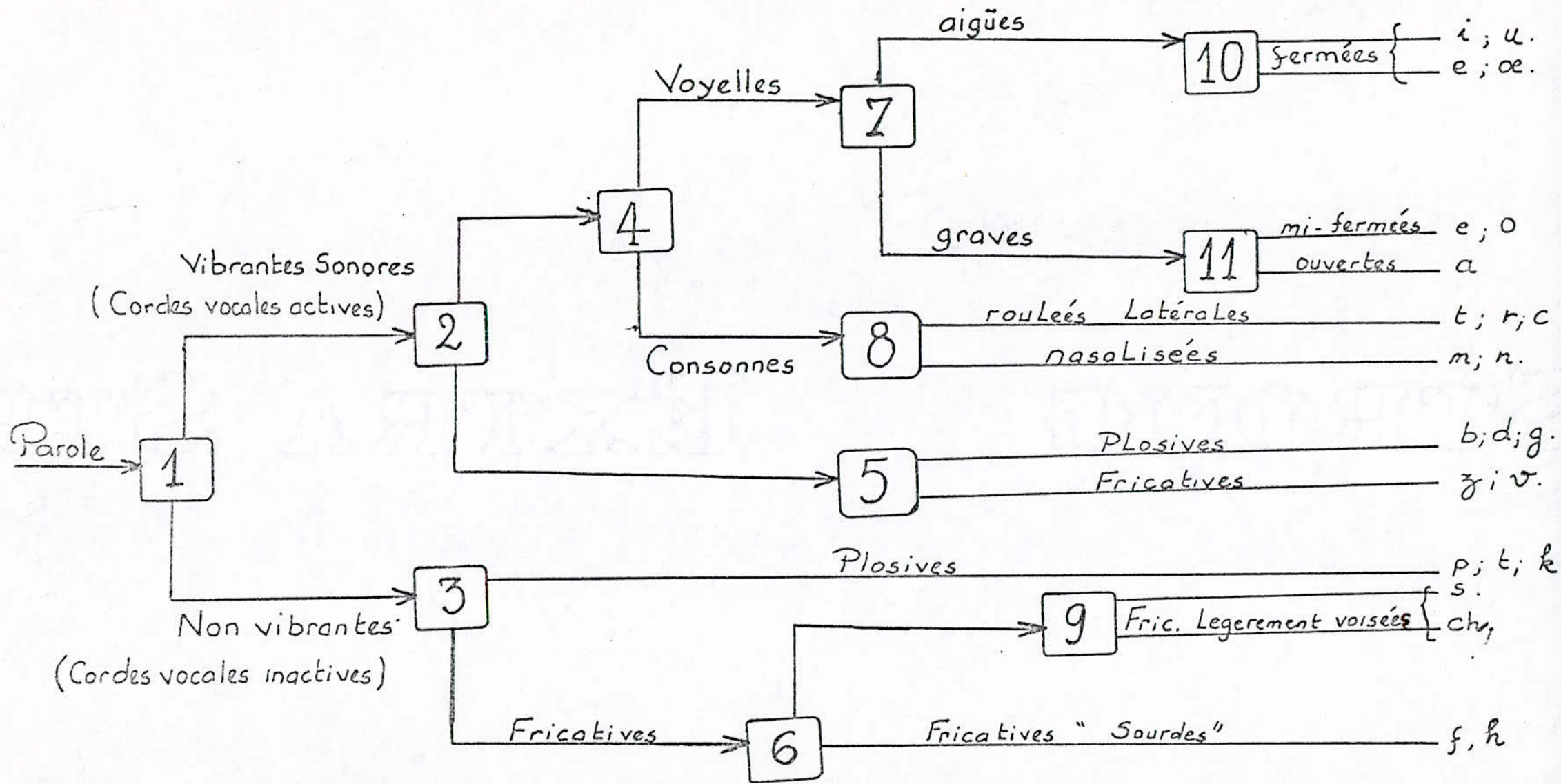


Tableau N° 3 ; Classification des différents sons de La Parole

Phonèmes	Terminologie Pratique	mot clé
/i/	i	Lit
/e/	e fermé	été
/ɛ/	e ouvert	forêt
/a/	a antérieur	Papa
/ɑ/	a postérieur	âne
/o/	o ouvert	port
/ø/	o fermé	peau
/u/	ou	Loup
/y/	u	tu
/ø/	eu fermé	Creuse
/œ/	eu ouvert	peur
/ɛ̃/	ẽ sourd	petit
/ɪ/	in	brin
/œ̃/	un	brun
/ɑ̃/	an	blanc
/ɔ̃/	on	blond
/j/	} SEMI-VOYELLES	jeux
/y/		Lui
/w/		Louis
/p/	} CONSONNES PLOSIVES OU OCCLUSIVES	pan
/b/		banc
/t/		rente
/d/		rude
/k/		car
/g/	baque	
/f/	} CONSONNES-Fricatives	faux
/v/		veau
/s/		Coussin
/z/		cousin
/ʃ/		chou
/ʒ/		gour
/m/	} CONSONNES NASALES	mer
/n/		banal
/ɲ/		agneau
/ŋ/		camping
/l/	} Consonnes Liquides	Lire
/ʎ/		rive

Tableau N°4 :

Les 37 Phonèmes de La Langue Française

1. Caractéristiques acoustiques principales des sons de la parole :

Introduction :

Le signal de la parole, représente une vibration donnée du conduit vocal est caractérisée par sa " fréquence fondamentale ", et un nombre fini de formants d'une part, et par son amplitude correspondant à la pression du débit d'air rejeté par les poumons et, mesuré en décibels ( dB ).

La fréquence fondamentale que l'on appelle aussi fréquence de mélodie est comprise entre 100 et 150 Hz pour les hommes, entre 200 et 300 Hz environs pour les femmes, et se situe aux environs de 400 Hz pour les enfants.

1.1 Timbres et son :

Le son est dû principalement à la propagation dans l'air, ou dans un quelconque milieu élastique d'une vibration mécanique.

Alors que le timbre du son lié à cette vibration définit sa " qualité globale ", on entend par qualité globale :

( son haut, son bas ) qualité : ( volume : sourd, grinçant, strident etc ... ).

D'un point de vue spectral, nous avons vu qu'une vibration sonore était caractérisée par sa fréquence fondamentale "  $F_0$  " et un certain nombre de formants. Les formants ne constituent que les maxima de la courbe de réponse du conduit vocal. Il s'avère que le timbre d'un son est directement lié aux caractéristiques spectrales du signal à savoir :

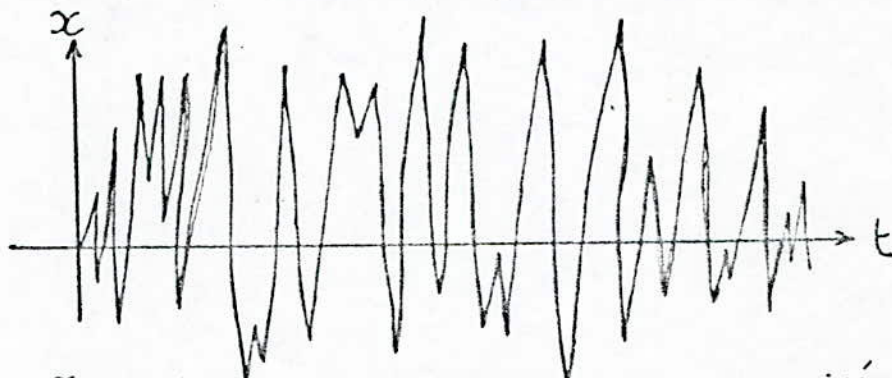
- La stabilité de la fréquence  $F_0$
- Nombre de formants
- La richesse en harmoniques de façon générale.

Par ailleurs le timbre des sons a fait l'objet de récentes recherches, qui ont prouvé qu'en régime permanent, il dépendait de caractéristiques spectrales, précédemment cités, et qu'en régime transitoire, il était fonction de sa période, de la durée et de l'évolution de ces composantes fréquentielles. On a su ainsi faire la différence entre des sons à 1 formant, 2 formants, nous verrons plus loin le détail de cette étude.

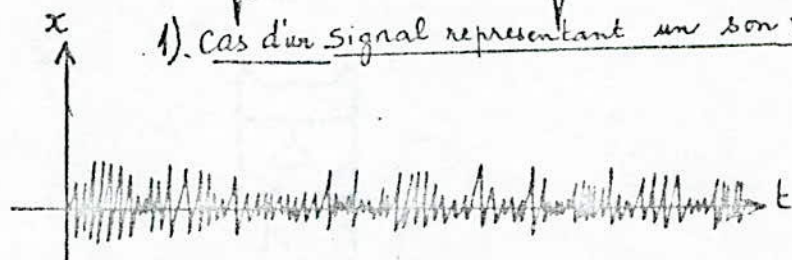
1.2 Présentation du signal de la parole :

Nous donnerons dans ce paragraphe deux formes de réponses du signal de la parole, les plus souvent rencontrées.

Il ne s'agit pas d'enregistrement pratique, mais de formes théoriques, donnant une image " virtuelle " de ce que l'on pourrait rencontrer comme signal, lors de l'étude pratique de ce dernier.



1). Cas d'un signal représentant un son voisé



2). Cas d'un bruit représentant un son non voisé (Bruit)

1.3 Timbre et régime transitoire :

Un chercheur du nom de STUMPF a montré l'importance de l'effet du régime transitoire sur le timbre d'un son. Il a mis en évidence l'influence de la suppression de la période de formation d'une note de musique, sur la reconnaissance d'un instrument donné. La reconnaissance de l'instrument concerné s'avéra impossible.

Cette expérience a montré que la période du transitoire agissait de plusieurs façons sur le timbre. Deux causes essentielles ont été dégagées ; à savoir :

- a) Une cause proprement physique liée à la composition spectrale de la période du transitoire et à l'évolution des harmoniques.
- b) Une cause physiologique liée à la perception du signal, c'est le rôle de l'oreille qui est ici incriminé et précisément à la constante de temps de réaction de l'oreille.

1.4 Paramètres acoustiques :

a) L'intensité acoustique :

Elle correspond à la quantité d'énergie transportée par l'onde sonore par unité de temps et par unité de surface perpendiculairement à la direction de propagation.

Elle est définie par la relation :

$$I = \frac{P^2}{\rho v} \quad \text{ou} \quad \begin{array}{l} P : \text{pression acoustique en (Pa)} \\ \rho : \text{masse spécifique de l'air en kg/m}^3 \\ v : \text{vitesse du son} \end{array}$$

La valeur de l'intensité acoustique en un point donné dépend de l'énergie fournie au départ par la source.

b) Niveau d'intensité - niveau de pression sonore :

L'oreille est <sup>un</sup>organe sensible à des pressions variant entre  $2 \cdot 10^{-5}$  Pa à  $2 \cdot 10^2$  Pa,

Des écarts d'intensité acoustique ne correspondraient plus à la réalité de la perception des sons. C'est pourquoi on a été amené à caractériser l'intensité sonore par une grandeur qui rend mieux compte de la réalité et en prenant le logarithme de l'intensité, on a ainsi défini le niveau d'intensité sonore qui s'écrit :

$$NI = 10 \log \frac{I}{I_0} \quad (\text{s'exprimant en dB}).$$

et correspondant à un niveau de pression :

$$NPS = 20 \log \frac{P}{P_0} \quad (\text{dB}) \quad (\log = \text{logarithme décimal})$$

(  $P_0, I_0$  )  $P_0$  - valeur normale de pression prise dans les conditions normales.

$I_0$  - choisi pour que dans les conditions normales on ait : M.P.S  $\approx$  N I.

## 2. Rôle de l'oreille :

### 2.1 Introduction :

Après avoir vu le son d'un point de vue objectif, c'est à dire en tant que phénomène physique aux propriétés mesurables à l'aide d'appareils ( Intensité acoustique, pression acoustique, fréquence ) nous allons désormais nous intéresser à l'aspect subjectif du son, c'est à dire à la perception que l'homme a, à travers son appareil auditif, et dans un premier temps nous décrivons sommairement, la physiologie interne de l'oreille en mettant en relief les organes et les processus mis en oeuvre lors de l'audition.

Nous nous limitons à l'essentiel dans l'étude pour dégager uniquement les propriétés physiques de l'oreille. ( voir schéma: 4 )

### 2.2 L'appareil auditif périphérique :

#### a) description générale :

L'appareil a pour but de nous faire percevoir les sons, c'est à dire de transformer les ondes sonores en informations analysables et décodables par le cerveau ( voir coupe schématique de l'oreille fig n° 4 ). Ce processus de transduction que l'on distingue en trois grandes parties qui correspondent aux trois parties que l'on distingue communément dans l'oreille :

- L'oreille externe: constituée du conduit auditif externe.
- L'oreille moyenne: constituée elle du tympan, et d'osselets ( marteau et enclume ).
- L'oreille interne: ( cette dernière est constituée de canaux semi-circulaires et cochlée )

#### b) L'oreille externe :

Le pavillon est essentiellement constitué de cartilage, son rôle est d'amener l'onde sonore, par l'intermédiaire de ses replis vers le conduit auditif, ce dernier permet l'acheminement de cette onde vers le tympan.

#### c) L'oreille moyenne :

Elle est constituée du tympan et de l'ensemble des trois osselets le marteau, l'enclume et l'étrier, son rôle est de transmettre les vibrations de l'onde sonore aux liquides de la cochlée ( système hydraulique de la cochlée ). Sous l'effet des vibrations de la pression, la membrane du tympan entre en vibration. L'amplitude des oscillations n'est pas négligeable, et se trouve amplifiée par les osselets ( M,E;E ), avant d'être transmise aux liquides intracochléaires.

#### d) L'oreille interne :

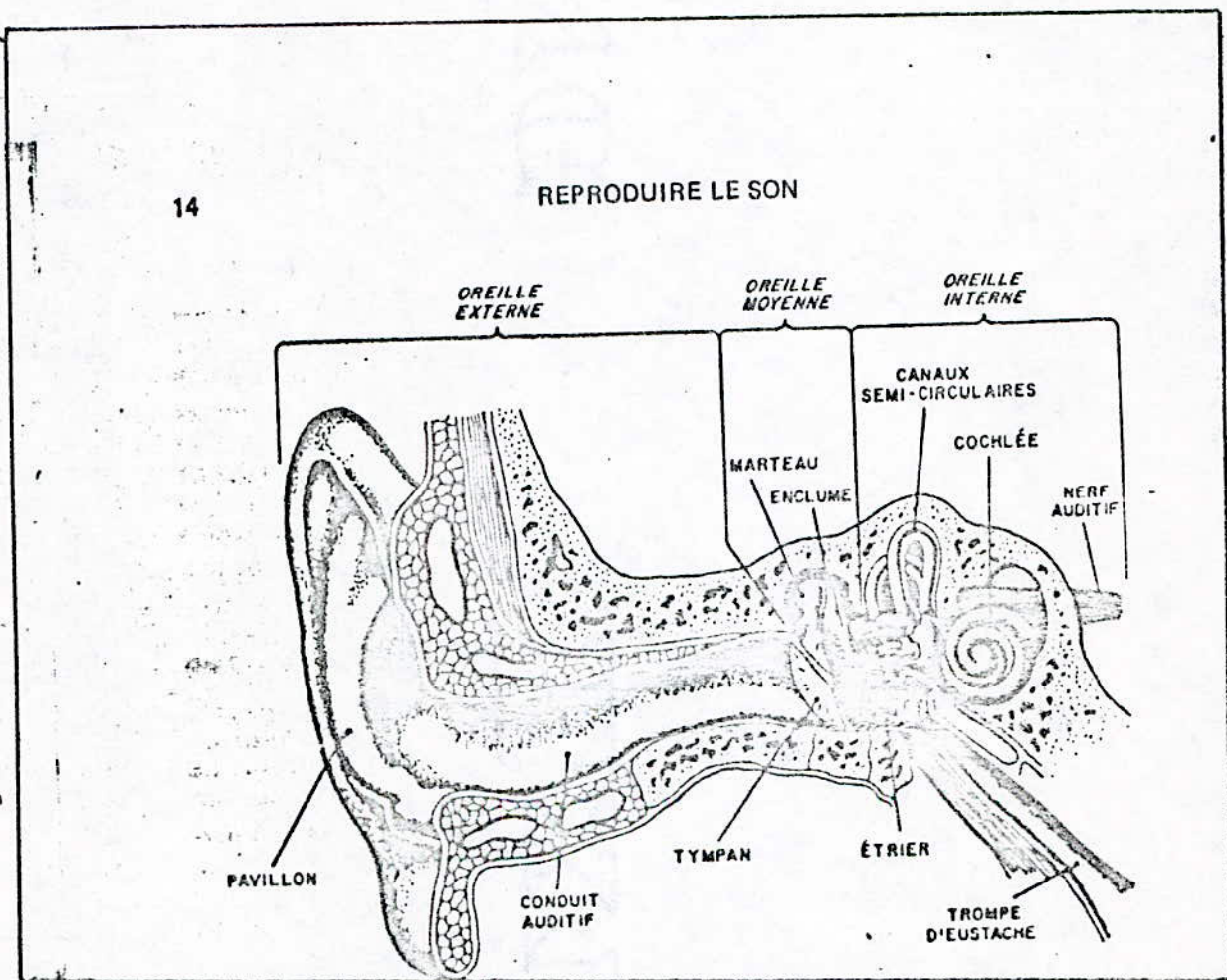
Situé à l'intérieur de l'os temporal, elle est divisée en deux grandes cavités, une de ces cavités contient les canaux circulaires qui jouent un rôle fondamental ( équilibre ).

La deuxième cavité contient l'oreille interne proprement dite, la cochlée.

La cochlée, elle même constituée de cavités, contient des liquides, dont la composition varie d'une cavité à une autre, ce sont les liquides intracochléaires. Sous l'action de l'étrier qui crée des compressions ou des dépressions, ces fluides, parcourus par des vibrations.

sort

Coupe Schematique de l'oreille 4





2.3 Réaction de l'oreille au son :

a) Sensibilité de l'oreille à l'intensité.

La loi de Weber-Fechner nous apprend que la sensibilité de l'oreille c'est à dire le pouvoir qu'à l'oreille pour percevoir un son varie comme le logarithme de l'intensité de l'excitation, c'est à dire : si l'on a deux sons d'intensité (  $I_1$  ,  $I_2$  ) et de sensibilités (  $S_1$  ,  $S_2$  )

$$( S_2 - S_1 ) = \log ( I_2 ) - \log ( I_1 ) = \log \frac{I_2}{I_1}$$

On comprend donc, à l'énoncé de cette loi, l'introduction de l'échelle logarithmique pour mesurer l'intensité sonore ou le niveau de pression sonore.

b) Seuil différentiel d'intensité ou sensibilité différentielle de l'oreille à l'excitation :

C'est la plus petite différence du niveau d'intensité qui est détectée par l'oreille ( noté  $\Delta N I$  ). Cette valeur dépend de la vitesse à laquelle les sons sont perçus c'est à dire plus le temps séparant deux sons est court, plus petit sera le seuil (  $\Delta N I$  ), détecté.

c) Sensibilité de l'oreille à la fréquence :

L'oreille d'un homme jeune doué d'une bonne audition perçoit les sons de fréquence, comprise, entre 20 HZ et 20 KHZ.  
Rappelons pour mémoire que la fréquence fondamentale du signal de parole est pour l'homme 70 HZ - 160 HZ.  
pour la femme 130 HZ - 250 HZ

III - Théorie formantique :

1. Son à 1 formant :

Tels  $\rightarrow$  ~~les~~ sont appelés certains sons naturels qui n'ont qu'une bande de fréquence dominante, ( présence d'un seul formant ).  
- Si le formant est aigu, on perçoit un sifflement,  
- S'il est très étroit, on perçoit un son coloré,  
- Si la bande de fréquence du formant est large, on a l'impression confuse dite de bruit ( bruit blanc ).

2. Sons à deux formants :

De tels sons sont assez répandus, telles que les voyelles et la plupart des notes de musique. Ces observations spectrales ont permis de bâtir une théorie formantique basée sur les critères suivants :

- a) On admet que ces sons ont deux zones formantiques caractéristiques.
- b) Leur timbre dépend de l'emplacement de ces zones dans l'échelle des fréquences.

Cette théorie a permis d'expliquer certains phénomènes musicaux telle que l'impossibilité de chanter la voyelle " @ " en texture aiguë. De définir un seuil différentiel de perception, hauteur tonale qui se mesure en SAVANTS.

De reconstituer dans les synthétiseurs de fréquences tels que le " VOCODER ", des timbres existants et d'en créer de nouveaux. Pierre Delattre a montré en dessinant ce qu'on appelle le diagramme formantique des voyelles en portant en abscisse la fréquence moyenne du formant (1), en ordonnée celle du formant (2) que les points représentatifs de toutes les voyelles se groupaient le long de deux côtés d'un vague triangle, dit triangle des voyelles de Delattre.

Le triangle vocalique dessiné en page suivante montre les valeurs moyennes des deux premiers formants ( voir fig n.° 1 ).  
Cette théorie était cependant insuffisante car elle ne tenait pas compte de :

- a) La répartition des harmoniques.
- b) La période de la formation.

Elle négligeait ainsi les sons de plus de 2 formants qui restaient mal étudiés ( consonnes, etc ... ).

Un autre chercheur G. Straka ( 1965 ) a déterminé une figure donnant les paramètres d'articulation ( voir fig n° 2 ), qui porte le nom de trapèze articulatoire de G. Straka. Cette figure donnait la position du point d'articulation et l'ouverture, lors de la prononciation des voyelles.

En étudiant parallèlement ces deux tracés on s'aperçoit que :

- L'ouverture → 1) Si (L'ouverture) du conduit vocal augmente, la valeur du premier formant augmente (  $F_1$  ).
- 2) Le déplacement de la langue vers l'arrière correspond à une diminution du deuxième formant ( 2ème formant ) (  $F_2$  ).

IV - Intelligibilité de la parole :

1 . Introduction :

On appelle intelligibilité de la parole le pourcentage de <sup>phonèmes</sup> ~~phénomènes~~ perçus correctement, c'est à ~~une~~ <sup>une</sup> propriété sémantique de la parole, son importance est fondamentale notamment dans les techniques audiovisuelles - pour la mesurer, deux groupes de méthodes peuvent être utilisés.

2 . Méthodes dites "subjectives" :

Elles consistent à faire rejeter des messages par des "cobayes". Pendant que d'autres les écoutent. Le nombre de messages correctement perçus est alors exprimé en pourcentages.

Les messages peuvent être des logatomes ( mots composés d'une syllabe porteurs d'aucune signification ) tels que : ter, neut, biss. Cette méthode est assez rigoureuse, elle impose un grand nombre d'essais pour arriver à des valeurs statistiques acceptables.

3 . Méthodes dites objectives :

a) Courbes de FLETCHER et MUNSON.

Fletcher et Munson ont étudié les réactions simultanées de l'oreille aux variations de fréquences et aux variations d'intensité acoustiques, pour cela, ils ont demandé à un certain nombre d'expérimentateurs de se prêter à l'expérience suivante :

On fait entendre un son à 1000Hz d'une intensité acoustique donnée, puis en faisant varier la fréquence du signal sans en modifier l'intensité acoustique, on demande à la personne de modifier l'intensité acoustique jusqu'à ce qu'elle ressente la même sensation auditive qu'avec le signal à 1000 Hz.

L'expérience est répétée pour plusieurs niveaux d'intensité acoustiques et donne les courbes d'isophonie ( même sensation auditive ) de la fig 3 ( ~~1940~~ ), appelées courbes de Fletcher et Munson. ( voir fig n° 3 ).

A partir des courbes on constate que plus l'intensité est faible, plus la sensibilité aux basses et hautes fréquences est faible.

V - Variabilité du signal de la parole :

Introduction :

Les sons de la parole, s'ils ont un certain nombre de points communs dans leurs principes de production, n'en sont pas moins variables selon la personne, où les conditions de locution en distingue 2 types de variabilité de la parole :

- Variabilité intra locuteur.
- Variabilité inter locuteur.

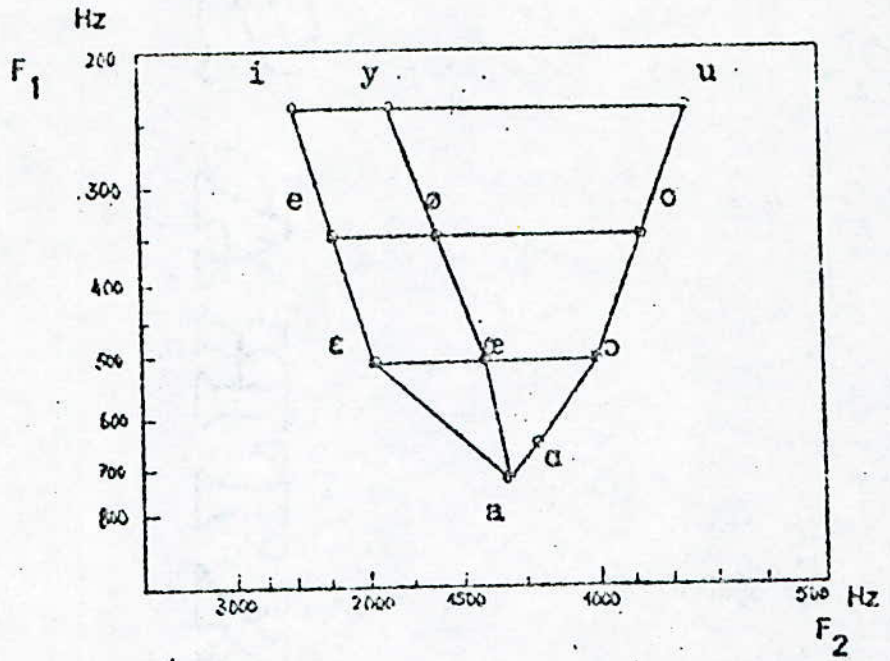


Figure 1 - Triangle acoustique des voyelles du français d'après P. DELATTRE (1948).

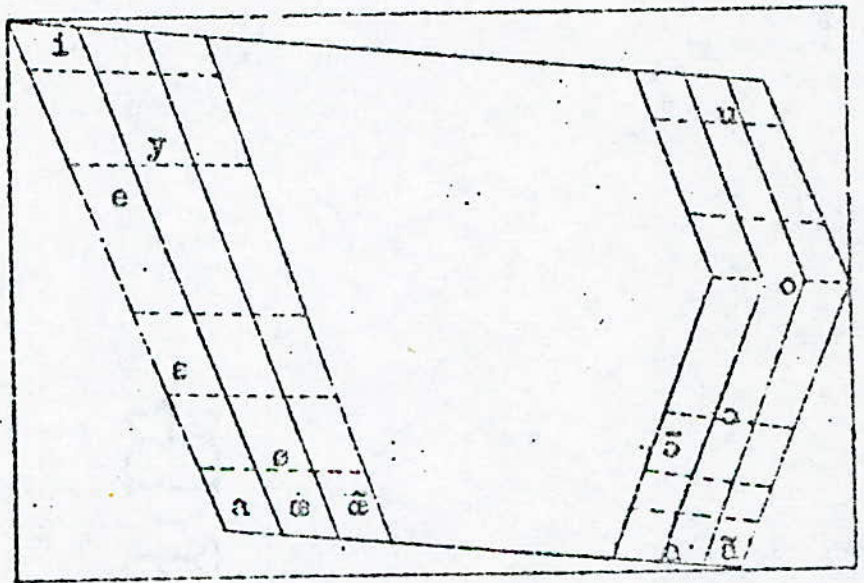
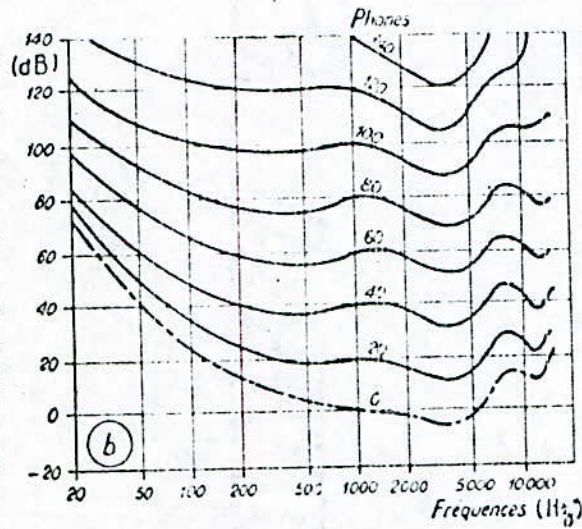
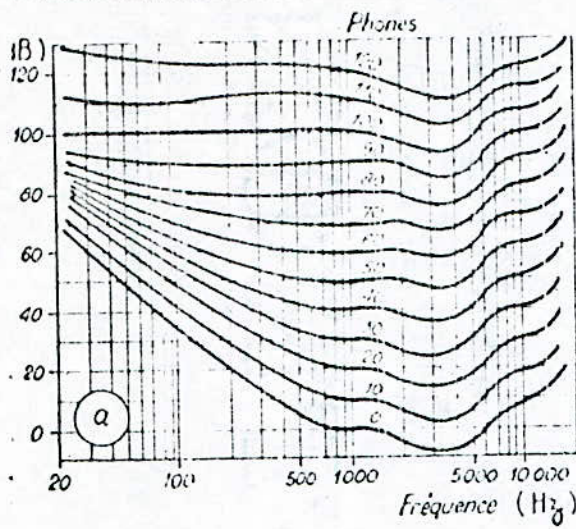


Figure 2 - Quadrilatère articulaire des voyelles du français d'après G. STRAKA (1965).



- - - - - seuil d'audibilité  
 courbes de Fletcher et Munson (servant de normes internationales)  
 courbes de Robinson  
**Fig. 3** — Courbes d'égalité de sensations d'intensité sonore.

1 . Variabilité Intra :

Elle concerne les différences de production du signal de la parole chez un même locuteur. Elle dépend :

- a) des soins apportés à l'articulation.
- b) vitesse d'élocution.
- c) force de la parole ( type de voix ).
- d) variation de mélodie.

2 . Variabilité interlocuteur :

Elle concerne les différences due aux caractéristiques des différents locuteurs qui sont :

- Variation de la mélodie ( suivant le locuteur )
- Variation de l'échelle formantique.

- CHAPITRE XI -

TRAITEMENT DU SIGNAL DE LA PAROLE :

I - Introduction :

La parole est un signal très redondant, et sert à véhiculer non seulement l'information propre au message à transmettre, mais aussi celle relative au locuteur.

En tant que phénomène physique, l'étude de la parole nécessite l'utilisation des systèmes numériques qui consiste en l'extraction des paramètres fondamentaux tel que le pitch, les formants. A cet effet, plusieurs méthodes analogiques et numériques d'analyse de parole ont été décrites.

Les analyseurs utilisent toutes les notions d'analyse spectrale du signal rayonné par la bouche pour le traitement fréquentiel, et d'analyse par auto-corrélation pour le traitement temporel.

II - Analyse spectrale :

1. Introduction :

Parmi les méthodes d'analyse spectrales de type numérique, on trouve la transformée de Fourier.

Celle-ci consiste à décomposer un signal  $x(t)$  sous la forme d'une somme infinie de sinusoides de fréquence  $f$ . et d'amplitude  $x(f)$ :

2. Transformée de Fourier :

Lorsqu'on considère un signal quelconque, il est indispensable d'avoir présent à l'esprit deux représentations possibles de ce signal, une représentation temps, c'est à dire une représentation de la forme  $y = f(t)$  dans laquelle la variable indépendante  $t$  est la durée qui s'écoule, et une représentation fréquence de la forme  $Y = F(\gamma)$  dans laquelle la variable indépendante est la fréquence dont la dimension est l'inverse d'un temps.

Ces deux représentations sont reliées entre elles par la transformation de Fourier et il est indispensable, pour pouvoir utiliser les méthodes de traitement du signal de savoir manipuler ces deux représentations complémentaires.

2.1. Transformée de Fourier des signaux périodiques :

Soit  $x(t)$  une fonction périodique de période  $T$ , on montre que  $x(t)$  peut s'écrire :

$$x(t) = \frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos \frac{2\pi}{T} nt + b_n \sin \frac{2\pi}{T} nt)$$

$$x(t) = \frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos 2\pi \nu_0 nt + b_n \sin 2\pi \nu_0 nt)$$

en posant :  $\nu_0 = \frac{1}{T}$ , les coefficients  $a_n$  et  $b_n$  sont données par :

$$a_n = \frac{2}{T} \int_{-T/2}^{T/2} x(t) \cos 2\pi \nu_0 nt dt$$

$$b_n = \frac{2}{T} \int_{-T/2}^{T/2} x(t) \sin 2\pi \nu_0 nt dt$$

Si l'on pose :  $X(n\nu_0) = \frac{1}{2} (a_n - j b_n)$

on a :  $X(n\nu_0) = \frac{1}{T} \int_{-T/2}^{T/2} x(t) \cdot e^{-2\pi j n \nu_0 t} dt$

$X(n, \nu_0)$  est le spectre de fréquence, grandeur général<sup>valeur</sup> complexe, qui peut se décomposer en :

- spectre d'amplitude :

$$|X(n, \nu_0)| = \frac{1}{2} \sqrt{a_n^2 + b_n^2}$$

- spectre de phases :

$$\varphi(n, \nu_0) = \text{Arctg} \left( -\frac{b_n}{a_n} \right)$$

réciroquement on aura :

$$X(\nu) = \sum_{n=-\infty}^{\infty} |X(n, \nu_0)| \cdot e^{-j\varphi(n, \nu_0)}$$

$$x(t) = \sum_{n=-\infty}^{\infty} e^{2\pi j n \nu_0 t} \frac{1}{T} \int_{-T/2}^{+T/2} x(\sigma) e^{-2\pi j \nu_0 n \sigma} \cdot d\sigma$$

les valeur négatives de  $n$  sont introduites pour rendre les équations plus simples. On vérifie aisément que si  $x(t)$  est réel :

$$a_{-n} = a_n \quad \text{et} \quad b_{-n} = -b_n$$

il est important de remarquer que le spectre d'une fonction périodique de période  $T$  est composée de raies dont l'écart minimum est sur l'axe des fréquences :

$$\nu_0 = \frac{1}{T}$$

Le spectre d'une fonction periodique est donc essentiellement discontinu. Il n'existe que pour les valeurs de la fréquence multiples de  $\frac{1}{T} = \nu_0$

2. 2 T. de F des signaux non periodiques :

Du sens de la communication, la serie de Fourier ne présente pas un grand intérêt vu qu'elle est appliquée à des signaux periodiques. Ainsi l'intégrale de Fourier permet le traitement et l'analyse quelque soit le signal (à periodique), la non periodicité peut être considérez comme résultant d'une extention à l'infini de la periode  $T$ . L'intervalle de fréquence  $\nu_0 = 1/T$  tend alors vers zéro et le spectre devient alors une fonction qui peut être continue on a :

$$x(t) = \int_{-\infty}^{+\infty} e^{2\pi j \nu t} d\nu \int_{-\infty}^{+\infty} x(\sigma) e^{-2\pi j \nu \sigma} \cdot d\sigma$$

$$X(\nu) = \int_{-\infty}^{+\infty} x(t) \cdot e^{-2\pi j \nu t} \cdot dt$$

est appelée la T. de Fourier de  $x(t)$ ,  
on écrit habituellement :

$$x(t) \iff X(\nu)$$

$X(\nu)$  est une fonction de  $\nu$ , en général complexe qui comprend donc une partie réelle (ou en phase), et une partie imaginaire (ou en quadrature).

$$\mathcal{R}[X(\nu)] = \int_{-\infty}^{+\infty} x(t) \cdot \cos(2\pi \nu t) \cdot dt$$

$$\mathcal{J}[X(\nu)] = \int_{-\infty}^{+\infty} x(t) \sin(2\pi \nu t) \cdot dt$$

le spectre d'amplitudes est :

$$|X(\nu)| = \sqrt{\mathcal{R}[X(\nu)]^2 + \mathcal{J}[X(\nu)]^2}$$

le spectre de phase est :

$$\varphi(\nu) = \text{Arctg} \left( -\frac{\mathcal{J}[X(\nu)]}{\mathcal{R}[X(\nu)]} \right)$$

3. Puissance, Energie des signaux :

3.1 Introduction :

Toute transmission d'information est liée à une transmission d'énergie. Lorsqu'en fait une mesure le processus, objet de la mesure, subit toujours un prélevement d'énergie de la part du dispositif de mesure. Cette notion de puissance d'un signal est donc très importante.

3.2 Puissance temporelle :

a) Puissance instantanée d'un signal :

Considérons un signal  $x(t)$  quelconque, donc pouvant être complexe on peut écrire :  $x(t) = a(t) + j b(t)$   
 $a(t)$  et  $b(t)$  étant alors deux fonctions réelles.

La P.i. de  $x(t)$  est :  $[a(t)]^2 + [b(t)]^2$

$$P(t) = [a(t) + j b(t)] \cdot [a(t) - j b(t)] = x(t) \cdot x^*(t) \quad (1)$$

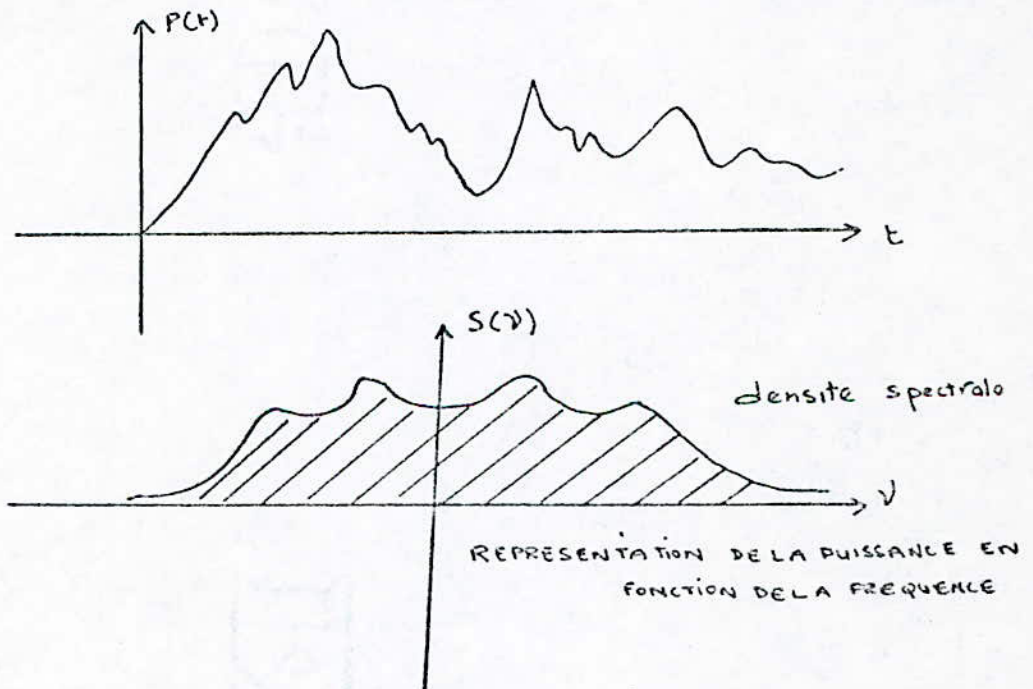
$x^*(t)$  est la quantité complexe conjuguée de  $x(t)$   
b) — P moyenne sur une durée T : on la définit par:

$$P(t_0, T) = \frac{1}{T} \int_{t_0}^{t_0+T} x(t) \cdot x^*(t) dt \quad (2)$$

si  $x$  est réel comme c'est le cas des signaux physiques alors (2) va devenir :

$$P(t_0, T) = \frac{1}{T} \int_{t_0}^{t_0+T} [x(t)]^2 \cdot dt \quad (3)$$

et on dira que l'énergie d'un signal est l'intégrale de sa puissance dans tout le domaine d'existence du signal



3.3 Energie d'un signal :

a) - dans un intervalle de temps  $\Delta t$  autour d'un instant  $t_0$

Cette énergie a pour valeur :

$$E(t_0, \Delta t) = \int_{t_0 - \frac{\Delta t}{2}}^{t_0 + \frac{\Delta t}{2}} P(t) \cdot dt = \int_{t_0 - \frac{\Delta t}{2}}^{t_0 + \frac{\Delta t}{2}} |x(t)|^2 \cdot dt$$

b) - Energie totale du signal : elle est donnée par

$$E_x = \int_{-\infty}^{+\infty} P(t) \cdot dt = \int_{-\infty}^{+\infty} |x(t)|^2 \cdot dt$$

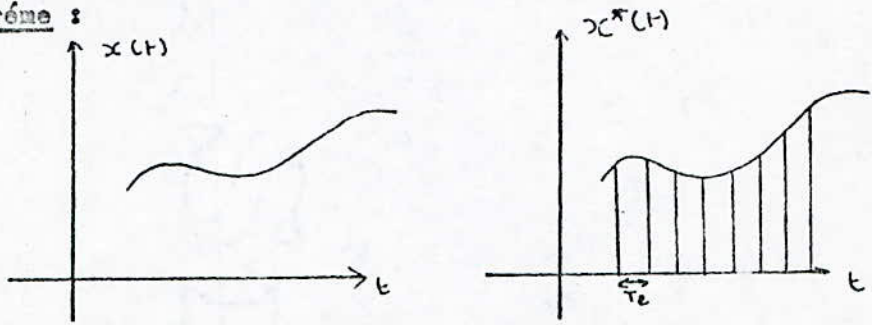


4. Analyse numérique :

4.1 Introduction :

Très souvent, on ne traite pas directement les signaux analogiques ( donc continus ) fournis par les capteurs de mesure, mais on les échantillonne, c'est à dire que l'on observe ces signaux d'une manière continue mais à certains instants seulement. <sup>non pas</sup>  
 Dans le cas général, ces observations sont périodiques : si elles sont séparées de  $T_e$ , on dit que l'on échantillonne à une fréquence  $F_e = 1/T_e$ . Il est donc indispensable de considérer l'effet de cet échantillonnage sur des signaux continus.

4.2 Théorème :



Le signal  $x^*(t)$  est appelé signal échantillonné de  $x(t)$  à la fréquence  $F_e = 1/T_e$  par le train d'impulsions de Dirac  $\delta(t-nT)$  on notera :

$$x(nT) = x^*(nT) = x(t) \sum_{-\infty}^{\infty} \delta(t-nT)$$

ou  $\delta(t-nT) = \begin{cases} 1 & \text{si } t=nT \\ 0 & \text{si } t \neq nT, \quad n \in \mathbb{N} \end{cases}$

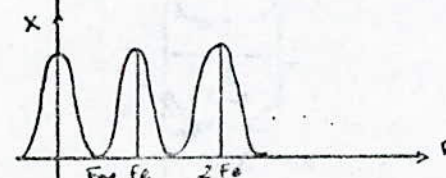
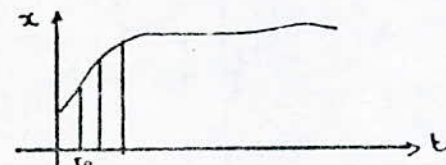
Donc l'échantillonnage est réalisé en multipliant le signal analogique  $x(t)$  par une suite périodique d'impulsions de Dirac de période  $T$ .

Dans le domaine fréquentiel, la Transformée de Fourier  $X_e(f)$  d'un signal échantillonné est donné par le produit de convolution.

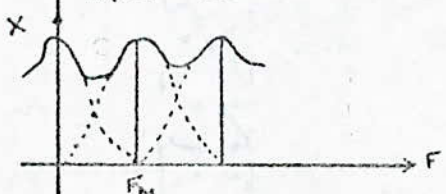
$$X_e(f) = X_a(f) * E(f) = \frac{1}{T} \sum_{-\infty}^{\infty} X_a(f - \frac{n}{T})$$

Donc on peut procéder à un échantillonnage si le théorème de Shannon est vérifié qui dit :

- on peut reconstituer une fonction à partir de ces échantillons pourvu que la fréquence d'échantillonnage soit supérieure ou égale à deux fois la plus haute fréquence contenue dans le spectre.



Cas ou  $F_e > 2F_M$



Cas ou  $F_e < 2F_M$

**4.3 Transformé de Fourier discrète :**

On appelle transformé de Fourier discrète, une représentation fréquentielle de suites temporelles périodiques, elle possède trois caractères fondamentaux :

- le signal est une suite périodique finie.
- son spectre est également une suite périodique finie.
- les périodes du signal et les périodes des spectres contiennent le même nombre d'échantillons :

- Notation :

$$X(k) = \sum_{k=0}^{N-1} x(n) \cdot e^{-j 2\pi \frac{kn}{N}} \quad ; \text{ Pour } k=0,1, \dots, N-1$$

la transformée inverse existe et a pour expression :

$$x(n) = \frac{1}{N} \sum_{k=0}^{N-1} X(k) e^{+j 2\pi \frac{kn}{N}} \quad ; \text{ Pour } n=0,1, \dots, N-1$$

l'inconvénient de cet algorithme est qu'il est long et fastidieux pour cela on fait appel à un autre algorithme beaucoup plus rapide celui-ci est appelé : F F T. (Fast, Fourier, Transforms)

**4.4 Transformé de Fourier rapide :**

La transformée de Fourier rapide est un algorithme de calcul permettant d'obtenir la transformé de Fourier discrète après un nombre d'opérations arithmétiques inférieur à celui nécessaire dans le cas d'un calcul direct.

En effet, la transformé de Fourier discrète d'une séquence ~~faire~~ finie

$$\{x(n)\}$$

a pour expression

$$X(K) = \sum_{n=0}^{N-1} x(n) \cdot e^{-j \frac{2\pi}{N} nK} \quad \begin{matrix} 0 \leq n \leq N-1 \\ 0 \leq K \leq N-1 \end{matrix}$$

si on pose :

$$W = e^{-j \frac{2\pi}{N}}$$

$$X(K) = \sum_{n=0}^{N-1} x(n) \cdot W^{nK}$$

or on a :

$$e^{-j 2\pi} = 1 \Rightarrow W^N = 1$$

$$e^{-j \pi} = -1 \Rightarrow W^{N/2} = -1$$

$\Rightarrow$  La (DFT)<sup>-1</sup> s'écrit alors :

$$x(n) = \frac{1}{N} \sum_{k=0}^{N-1} X(k) \cdot W^{-nK}$$

Du fait que  $\omega^{n+1}$  ce qui veut dire que les coefficients  $\omega^{nk}$  sont périodiques et de périodes  $N$ . En exploitant les propriétés de symétrie de  $\omega^{nk}$  on aboutira à une série d'algorithmes dont le 1er a été élaboré par : "Cooley-Tukey". Celui-ci nécessite  $N \log N$  multiplications alors que le calcul direct de la T.F.D par  $N$  Valeurs de  $\omega^{nk}$  nécessite  $N^2$  multiplication et  $N(N-1)$  addition, ce qui est très long et pas commode lorsque  $N$  augmente.

III - Analyse cepstrale :

1. Introduction :

On pourrait se demander, pourquoi ne filtre-t-on pas tout simplement le signal d'entrée pour éliminer les ondulations non souhaitées du signal de la parole ? C'est que le filtrage élimine tout ce qui a un caractère additif et non multiplicatif. Transformons alors les multiplications en additions grâce aux logarithmes et filtrons alors le signal - ceci est le principe de base de l'analyse cepstrale.

2. Traitement homomorphique des signaux :

2.1 Introduction :

Contrairement aux systèmes linéaires faciles à analyser, les systèmes non linéaires présentent quelques difficultés, celles-ci sont résolues par le traitement homomorphique qui est basé sur le principe de superposition généralisé

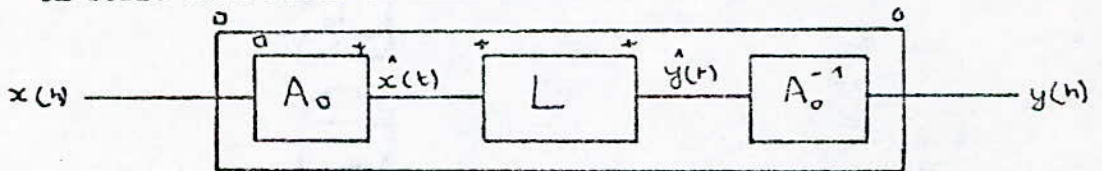


$\square$  = symbole d'addition ou de convolution ( pour les signaux d'entrée ).

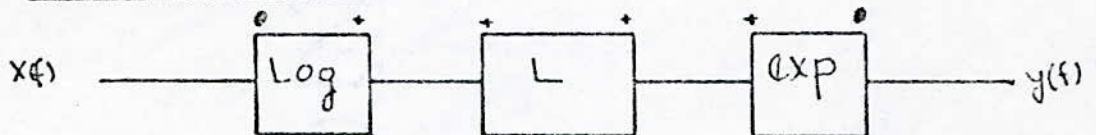
$\circ$  = "-----" ( "-----" sortie )

2.2 Forme canonique :

Tout système homomorphique peut être représenté par trois systèmes en série dont celui du milieu est linéaire :



2.3 Systèmes convolutifs :



Dans les applications très variées il est très fréquent de se trouver devant le problème d'isoler une composante utile du signal ou de mesurer certaines de ses caractéristiques, lorsque celui-ci est combiné par convolution avec d'autres signaux en supposant que les lois à l'entrée et à la sortie du système homomorphique sont identiques

$$X(f) \xrightleftharpoons{\quad} x(k) ; \quad x(k) = x_1(k) * x_2(k)$$

$$X(f) = X_1(f) \cdot X_2(f)$$

3 . Formulation :

Soit  $|S(f)|$  le spectre en amplitude d'un signal de parole et  $|S(f)|^2$  son spectre de puissance.

Pour un son voisé, le spectre de puissance peut se mettre sous la forme :  $|S(f)|^2 = |G^2(f)| \cdot |C(f)|^2$

ou  $|G(f)|$  est le spectre de la source glottique et la réponse en fréquence du conduit vocal.

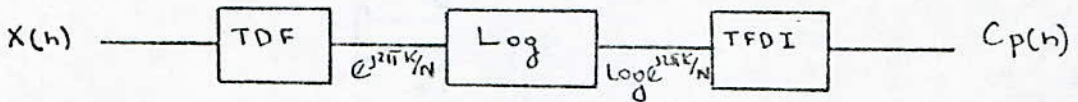
En prenant le logarithme de l'expression précédente il vient :

$$\text{Log} [ |S(f)|^2 ] = \text{Log} |G^2(f)| + \text{Log} |C(f)|^2 \quad (1)$$

en appliquant la T F on aura :

$$\text{T F } a = \text{T F } b + \text{T F } c$$

T F a au carré nous donne le cepstre :



Le cepstre est donc la T F inverse du logarithme du spectre et il est fonction d'un variable appelé quefrencce ayant la dimension du temps.

4 . Conclusion :

le principe de la méthode cepstrale, comme l'indique l'expression (1), est de séparer les effets de la source d'excitation et la réponse du conduit vocal.

Cette séparation est bien effectuée, ce qui représente l'avantage de cette méthode, utilisée en reconnaissance de la parole, pour la détermination du pitch. Cependant, elle reste non employée, du fait que le volume des calculs nécessaires est trop important.

IV - Filtrages :

1 . Filtres numériques :

1 . 1 Définition :

Un filtre numérique est un système numérique qui, recevant à son entrée, les échantillons quantifiés  $\{x(n)\}$  d'un signal  $x(t)$  délivre à sa sortie une suite de nombres  $\{y(n)\}$

1 . 2 Structure recursive :

a) définition :

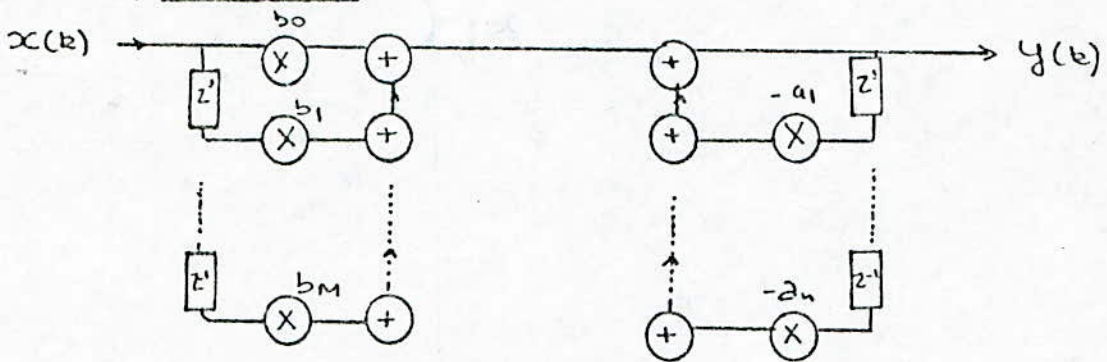
C'est un filtre dont chaque échantillon de sortie dépend des échantillons présents et passés d'entrée ainsi que des échantillons de sortie.

Ce type de filtre est caractérisé par l'équation de récurrence suivante :

b) Mise en équation :

$$y(k) = \sum_{m=0}^M \frac{b_m}{a_0} x(k-m) - \sum_{n=1}^N \frac{a_n}{a_0} y(k-n)$$

c) Modelisation :



d) Fonction de Transfert :

$$H(z) = \frac{Y(z)}{X(z)} = \frac{\sum_{m=0}^M \frac{b_m}{a_0} \cdot z^{-i}}{1 + \sum_{n=1}^N \frac{a_n}{a_0} z^{-i}}$$

1.3 Structure non recursive :

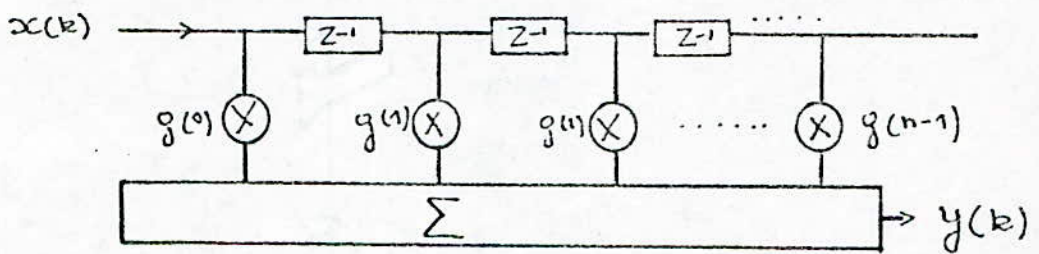
a) Définition :

C'est un filtre dont chaque échantillon <sup>de sortie</sup> dépend des échantillons présents et passés d'entrée, il est caractérisé par l'équation suivante :

b) Mise en équation :

$$y(k) = \sum_{l=0}^M g(l) \cdot x(l-k)$$

c) Modelisation :



d) Fonction de Transfert : il suffit de mettre les  $a_n = 0$

$$H(z) = \sum_{l=0}^M \frac{b_m}{a_0} \cdot z^{-i}$$

1.4 La prédiction linéaire :

a) Définition :

L'analyse par prédiction linéaire constitue une méthode très précise pour la détermination des coefficients prédictifs, et consiste à exprimer la valeur du nième échantillon en fonction des p échantillons et de l'entrée  $u(n, T)$  donc il s'agit d'obtenir une formule de type

$$(2) \hat{S}(n, T) = \sum_{k=1}^P a(k) \hat{S}(nT - kT) + Au(nT)$$

Les coefficients  $a(k)$  et  $A$  reproduit le mieux <sup>son</sup>  $-7,5$  possible le comportement du canal vocal de l'entrée  $u(nT)$

On obtient la formule de type (2), pour chaque fenêtre, ce qui signifie que les coefficients  $a(k)$  et  $A$  doivent être recalculés sur chaque fenêtre.

$$\hat{S}(n) = \sum_{k=1}^P a(k) \cdot \hat{S}(n-k) + Au(n)$$

Transformons en  $Z$  cette equation:

$$\hat{S}(z) = \sum_{k=1}^P a(k) \cdot \hat{S}(z) \cdot z^{-k} + Au(z)$$

$$\hat{S}(z) \left[ 1 - \sum_{k=1}^P a(k) \cdot z^{-k} \right] = Au(z)$$

$$H(z) = \frac{\hat{S}(z)}{U(z)} = \frac{A}{1 - \sum_{k=1}^P a(k) \cdot z^{-k}}$$

Le  $\hat{S}(z)$  apparait comme la reponse du filtre  $A/H(z)$  ne possédant que des pôles à l'entrée  $U$ .

Les principales méthodes de prédiction linéaire sont :

- La méthode d'auto-corrélation.
- La formulation du filtre en treillis.

b) Méthode d'auto-corrélation :

L'hypothèse de base dans cette méthode est de considérer que le signal est stationnaire dans un intervalle fini d'où l'intérêt du fenêtrage qui a pour rôle d'annuler le signal à l'extérieur de l'intervalle en question.

$$e(n) = S(n) - \sum_{k=1}^P a(k) S(n-k)$$

$$E = e^2(n) = \sum [S(n) - S'(n)]^2$$

$$= \sum [S(n) - \sum a(k) \cdot S(n-k)]^2$$

le principe consiste à <sup>obtenir</sup> faire  $E$  et à le minimiser :

$$dE / da(k) \rightarrow 0$$

$$\frac{\partial E}{\partial a(k)} = \sum_n [S(n) - \sum_{k=1}^P a(k) S(n-k)] S(n-k) = 0$$

$$\sum_{h=1}^P \sum_{k=1}^P a(k) S(n-k) \cdot S(n-i) = \sum S(n) \cdot S(n-i)$$

$$\sum S(n) \cdot S(n-i) = \sum_{k=1}^P a(k) \sum S(n-k) \cdot S(n-i)$$

ce qui constitue un système de P équation à P inconnu mais à partir de P = 12 on arrive à  $S(n) \cdot \hat{S}(n) = 0$   
 Posons  $m = n - 1$

$$\sum a(k) \sum_{m=-\infty}^{\infty} S(m) \cdot S(m+|i-k|) = \sum S(m+|i|) \cdot S(m)$$

on définit la fonction d'auto-correlation R(i) comme suit :

$$R(i) = \sum_{n=-\infty}^{\infty} S(n) \cdot S(n-i) \quad R(i) = R(-i)$$

$$R(i) = \sum_{n=0}^{N-1-|i|} S(n) \cdot S(n+|i|)$$

$$R(i) = \sum_{k=1}^P a(k) \cdot R(|i-k|)$$

$$R(i) = \sum_{k=1}^P a(k) \cdot R(|i-k|)$$

$$\begin{bmatrix} R(1) \\ R(2) \\ R(3) \\ \vdots \\ R(P) \end{bmatrix} = \begin{bmatrix} R(0) & R(1) & & & R(P-1) \\ & R(1) & & & \\ & & \ddots & & \\ & & & \ddots & \\ & & & & R(0) \end{bmatrix} \times \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ \vdots \\ a_P \end{bmatrix}$$

M : est une matrice carré qui a la particularité d'être symétrique et sa diagonale est composée d'éléments identique, on l'appelle Toeplitz.

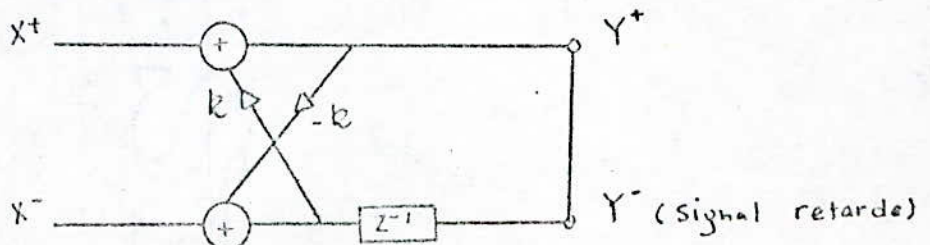
c) Méthode du filtre en treillis :

La modélisation se fait à partir d'un filtre à P étages ou P est l'ordre du filtre dans chaque étage, pour cela dans chaque étage, on réalise deux multiplications des signaux par les coefficients de réflexions appropriés, les signaux traités se déplacent en sens inverse l'un de l'autre.

Considérons maintenant un étage du filtre, il se présente sous forme d'un quadripôle avec les signaux :

$X^+$  et  $X^-$  à gauche,  $Y^+$  et  $Y^-$  à droite, les signes + et - en indice se réfèrent aux signaux directs et rétrogrades respectivement.

Modèle du filtre :



Considérons cet étage et écrivons les équations en aura :

$$Y^+ = Y^-$$

$$Y^+ = X^+ + k z^{-1} \cdot Y^-$$

$$X^+ = Y^+ - k z^{-1} Y^-$$

$$X^- = -k Y^+ + z^{-1} Y^-$$

$$\Rightarrow \begin{cases} X^+ = Y^+ - k z^{-1} Y^- = S \\ X^- = -k Y^+ + z^{-1} Y^- \end{cases} \Rightarrow \begin{pmatrix} X^+ \\ X^- \end{pmatrix} = \begin{pmatrix} 1 & -k z^{-1} \\ -k & z^{-1} \end{pmatrix} \begin{pmatrix} Y^+ \\ Y^- \end{pmatrix}$$

dans un filtre mono-étage, les bornes  $Y^+$  et  $Y^-$  sont connectées ensemble et donnent le signal de sortie  $S$ , tandis que la borne  $X^+$  où pénètre le signal d'entrée  $E$ , l'équation matricielle précédente devient :

$$\begin{pmatrix} E \\ X^- \end{pmatrix} = \begin{pmatrix} 1 & -k z^{-1} \\ k & z^{-1} \end{pmatrix} \begin{pmatrix} S \\ S \end{pmatrix} ; \text{ soit } E = (1 - k z^{-1}) S$$

$$\text{d'où } \frac{S}{E} = \frac{1}{1 - k z^{-1}}$$

on retrouve ainsi la fonction de transfert du circuit de synthèse par prédiction linéaire, et en pourra généraliser en mettant plusieurs filtres en cascade le résultat sera :

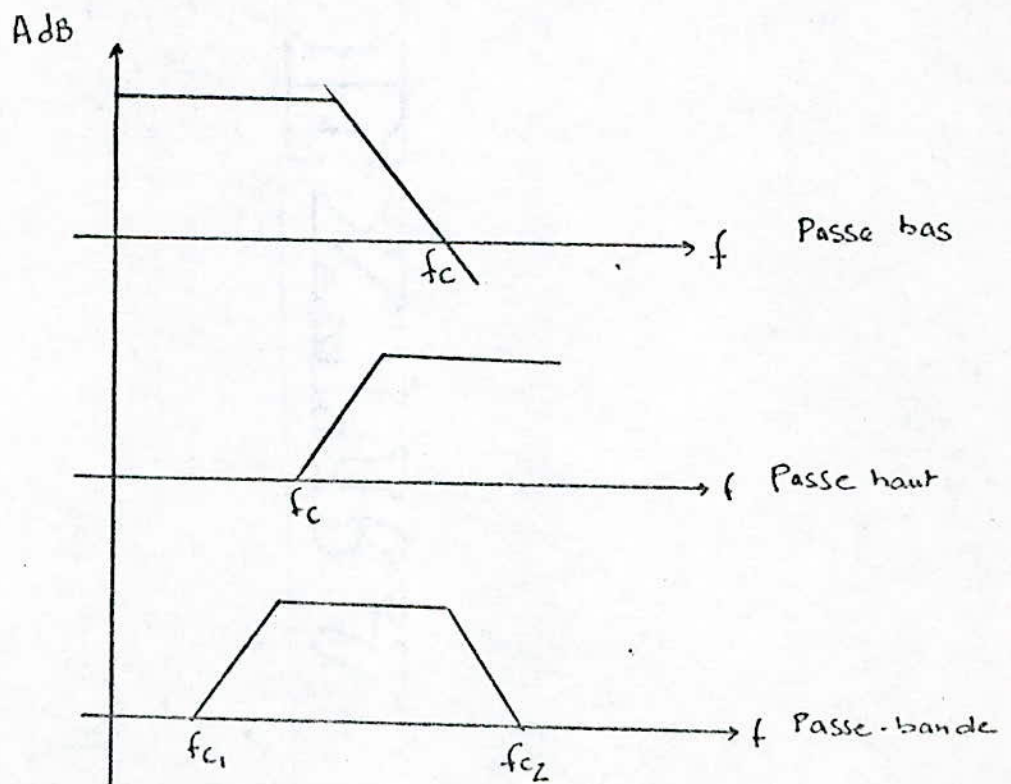
$$\frac{1}{1 - \sum a_k \cdot z^{-k}}$$

## 2. Filtres analogiques :

Dans le domaine analogique, le rôle d'un filtre est de sélectionner et choisir parmi toutes les composantes spectrales du signal considéré. Celles qui remplissent les conditions imposées :

- fréquence inférieure à une limite " $f_c$ " pour les passe-bas
- fréquence supérieure à une limite " $f_c$ " pour les passe-haut
- fréquence comprise dans un certain intervalle " $f_c$ " pour les "passe-bande".

Ces différents filtres sont caractérisés par la courbe représentant les variations de la fonction de transfert " $T$ " en fonction de la fréquence " $f$ ", appelée courbe de réponse du filtre :





V - Conclusion :

Le point le plus important que nous avons tenté de mettre en évidence dans ce chapitre est le lien étroit existant entre l'analyse spectrale et la modélisation des signaux. Ceci est directement lié à l'impossibilité pratique d'un signal sans hypothèses supplémentaires. En effet, en ajoutant des hypothèses et que si l'on connaît parfaitement la fonction de corrélation en un support limité, une analyse spectrale parfaite du signal est théoriquement possible.

LE CEPSTRE COMPLEXE :

Dans ce qui suit, nous considérons de séquences  $X(n)$  absolument sommables. De telles séquences ont une transformée en  $z$  dans le domaine de convergence comprend le cercle unité et qu'il est alors légitime d'introduire les transformées de Fourier associées.

Le cepstre complexe d'une séquence  $X(n)$  associée à la transformée  $X(z)$  est alors défini comme la séquence  $\hat{X}(n)$  dont la transformée de Fourier  $\hat{X}(e^{j\omega})$  est le logarithme complexe de la transformée de Fourier  $X(z)$ .

$$\begin{aligned} (1) \quad \hat{X}(e^{j\omega}) &= L_{\omega} X(e^{j\omega}) = L_{\omega} |X(e^{j\omega})| + j \text{Arg} [X(e^{j\omega})] \\ (2) \quad \hat{X}(n) &= \frac{1}{2\pi} \int_{-\pi}^{+\pi} \hat{X}(e^{j\omega}) \cdot e^{j\omega n} \cdot d\omega \end{aligned}$$

Dans l'expression (1) la partie imaginaire du logarithme complexe, qui n'est autre que la phase de  $X(z)$  évaluée sur le cercle unité est défini comme fonction impaire et continue de  $\omega$  sur  $]-\pi, +\pi[$  ( problème de l'unicité de la décomposition ).

Dans le cas des séquences à phase minimale ( pôles et zéros sont à l'intérieur du cercle unité ) associées à des transformées en  $Z$  rationnelles du type  $X(z)$  :

$$X(z) = A \prod_{k=1}^M (1 - a_k \cdot z^{-1}) / \prod_{l=1}^N (1 - c_l \cdot z^{-1})$$

on montre que le cepstre complexe est nul pour  $n$  négatif et qu'il est identique à un facteur multiplicatif 2 près au cepstre ( défini comme la transformée de Fourier inverse du logarithme du module de la transformée de Fourier  $X(n)$  ).

Cette propriété permet de s'affranchir de la manipulation délicate d'un logarithme complexe, on devine alors les relations suivantes :

$$\begin{aligned} \hat{X}(n) &= 0 \quad \text{Im} \quad n < 0 \\ \hat{X}(n) &= C(n) = L_{\omega} [A] : n = 0 \\ \hat{X}(n) &= 2 C(n) = \sum c_k^n / n - \sum a_k^n / n \quad (3) \end{aligned}$$

ou  $C(n)$ , dit cepstre est par définition donné par :

$$C(n) = \frac{1}{2\pi} \int_{-\pi}^{+\pi} L_{\omega} |X(e^{j\omega})| e^{j\omega n} \cdot d\omega$$

PROPRIETES DU CEPSTRE :

L'équation (3) met en évidence la décroissance du cepstre en  $1/n$ , pour un modèle sans zéros (prédiction linéaire), on remarque ainsi que le rythme de décroissance est directement lié au pôle de plus fort module (Bande la plus étroite).

La propriété essentielle est celle de déconvolution que l'on peut résumer ainsi :

- Si une séquence  $X(n)$  est la convolution de deux séquences
- $X_1(n)$  et  $X_2(n)$
- $$X(n) = X_1^*(n) \cdot X_2(n)$$

On aura par définition :

$$\hat{X}(e^{j\omega}) = \hat{X}_1(e^{j\omega}) + \hat{X}_2(e^{j\omega})$$
 Si les domaines d'existence des réponses fréquentielles  $X_1(e^{j\omega})$  et  $X_2(e^{j\omega})$  sont séparés, on aura ainsi la possibilité par transformation inverse, d'isoler la réponse associée à  $X_1(n)$ , en ayant fait un filtrage.

La troisième propriété est que le cepstre d'une série d'impulsions de période  $T$ , est une série d'impulsions de même période.

$$(4) \begin{cases} X_P(k) = \sum X(m) \exp(-j \frac{2\pi}{N} km) & 0 \leq k < N-1 \\ \hat{X}_P(k) = \text{Log} |X_P(k)| & 0 \leq k < N-1 \\ C_P(m) = \frac{1}{N} \sum X_P(k) \exp(j \frac{2\pi}{N} km) & 0 \leq m < N-1 \end{cases}$$

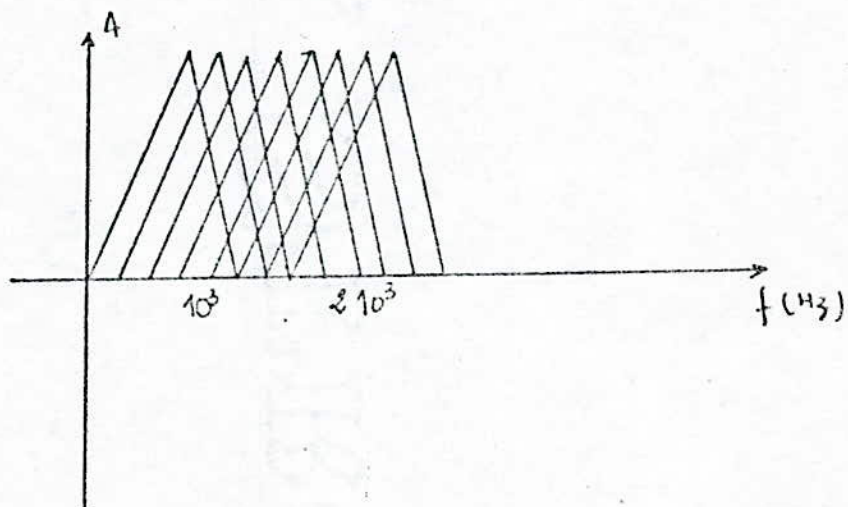
La notation avec l'indice  $P$  est utilisée pour souligner que la séquence  $C_P(n)$  n'est pas égale au cepstre mais reliée à ce dernier par  $C_P(n) = \sum_{-\infty}^{\infty} C(m + \tau \cdot N)$

Cepstre obtenu à partir d'une échelle de fréquences "MEL".

L'analyse précédente s'inspire de ce qui a été dit auparavant. Elle en diffère au niveau de la transformée de FOURIER du signal, celle-ci n'est plus linéaire en fréquence mais constituée d'une séquence de points qui sont les réponses en énergie d'un banc de filtres, dont la répartition suit l'échelle des fréquences dites "MEL".

Cette répartition, liée à des critères perceptifs jusqu'à 1000 Hz et logarithmique au delà.

Il faut utiliser un banc de filtres de  $N$  filtres triangulaires répartis comme il est indiqué par la figure 1.



Si  $N$  filtres sont exploités, les coefficients cepstraux sont obtenus par la transformation en cosinus suivante analogue à l'équation (4)

$$(5) \quad C_{mel}(i) = \sum_{k=1}^N L \omega_j E(k) \cdot \cos(i k \pi / N)$$

$1 \leq i \leq M$   
 $1 \leq k \leq N$

Où  $M$  est le nombre de coefficients du cepstre "MEL" que l'on retient et où les  $E(k)$  sont les réponses en énergie.

**CHAPITRE III**

**LES VOCODEURS A TRAVERS L'HISTOIRE**

- C H A P I T R E    I I I    -

- I - Introduction :

C'est à près de deux siècles que remontent les premières machines parlantes. Malgré leurs imperfections, elles en disaient assez pour qu'on ne puisse leur refuser le don de la parole.

Malheureusement, il ne reste de cette machine parlante ( l'une des toutes premières qui aient été réalisées ) qu'une description fort imprécise devant l'académie des sciences.

La machine à parler du Baron Wolfgang Von Kempelen, construite en 1791, a le grand mérite d'exister encore. Elle est exposée au Deutsches Museum de Munich. C'est un véritable " instrument de parole ", dont Von Kempelen était virtuose. Cette machine composée d'un soufflet, d'une " bouche " dont la variation du volume est effectuée par la main gauche pour l'émission des voyelles, de narines et de sifflets commandés par des leviers actionnés par la main droite. Von Kempelen avait tout de suite compris toute l'importance des liaisons entre les sons : mais dans sa simulation des organes phonatoires, la langue, les lèvres et les dents avaient été omises.

Bien d'autres chercheurs, tentèrent, après lui de créer des mécanismes imitant la voix. De Joseph Faber qui présenta en 1835 une machine qui chante, à Riesz qui, en 1937, conçut le dernier des synthétiseurs mécaniques. Une foule d'inventeurs marqua, avec plus ou moins de bonheur, l'histoire de la parole artificielle.

En 1939, Homer Dudley, des laboratoires Bell, publiait un article qui allait avoir un grand retentissement : " the vocoder " ( abréviation de Voice Coder ).

L'appareil réalisait l'analyse et la synthèse instantanées de la parole, pour améliorer les transmissions téléphoniques en réduisant les redondances sonores d'un message parlé.

Dudley reprenait dans cet appareil des principes établis trois années auparavant lorsqu'il avait présenté un synthétiseur de parole à commande manuelle : le Voder ( Voice demonstrator ).

Les principes fondamentaux n'ont guère subi de révolutions depuis le début de ces études. En revanche, c'est essentiellement la technologie qui a évolué. Il y a 20 ans le traitement de la parole s'effectuait encore sur des matériels analogiques, depuis le début des années 60 les chercheurs utilisent des calculateurs numériques. Le développement des circuits logiques en composants discrets, puis la réalisation des C.I ont permis d'obtenir dans un volume réduit des dispositifs d'une complexité et d'une puissance de calcul considérables. Les synthétiseurs de parole eux même sont maintenant fabriqués en technique numérique sous formes de " puce " de quelques mm<sup>2</sup>. Pour plus d'illustration, on donnera dans ce chapitre quelques réalisations modernes de type carte de reconnaissance, élaborée par VECSYS en collaboration avec le LUMSI qui figurent parmi les leaders de ce nouveau savoir-faire.

Ainsi VECSYS propose des applications dans le domaine de :

- La synthèse à partir du texte ----- IC0 85
- Restitution de parole composée ----- LPC, MIC, DELTA
- Reconnaissance des mots enchainés ----- RME 186
- ----- voisés ----- RMI 88, RMI 50
- Terminal vocal autonome ----- TLV 02 - TLV 50

Les fiches techniques de toutes ces cartes sont données en Annexe.

- II - Généralités sur les Vocodeurs :

1. Principe de base :

Les Vocodeurs permettent la numérisation du signal de parole avec des débits binaires inférieurs à 5000 bit/s.  
Pour obtenir une diminution de débit aussi importante par rapport au procédé de numérisation classique, le MIC (64 Kbit/s) utilisé en transmission téléphonique, il n'est plus possible de coder et de numériser l'onde temporelle. Il faut quitter le domaine temporel pour le domaine fréquentiel dans lequel on peut définir des paramètres qui sont suffisants pour représenter la parole et à partir desquels on peut synthétiser un signal de parole intelligible et présentant un certain agrément d'écoute. Ces paramètres décrivent à chaque instant le spectre instantané de fréquences et ils sont étroitement liés au processus de production du signal de la parole dans les organes phonatoires.

Les Vocodeurs, exploitant les propriétés intrinsèques du signal de parole, sont inaptes à traiter des signaux dont les propriétés spectrales et temporelles sont différentes.

Un des paramètres décrit la fréquence fondamentale ( fréquence de vibration des cordes vocales ), souvent appelée  $F_0$  ou " pitch " qui définit la structure fine du spectre instantané de fréquence. Quand il y a vibration de cordes vocales ( sons voisés ) nous sommes en présence d'un spectre de raies formé de la fréquence fondamentale et de ses harmoniques ( fig. 1 ).

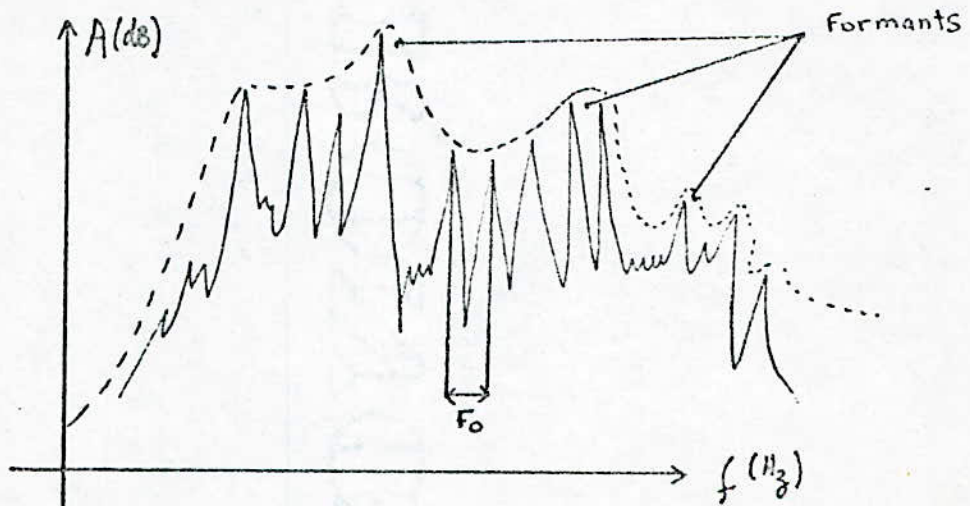


fig. 1 Spectre instantané d'une voyelle " O ".

En l'absence de vibration des cordes vocales ( sons non voisés ), il s'agit d'un spectre continu ou spectre de bruit.

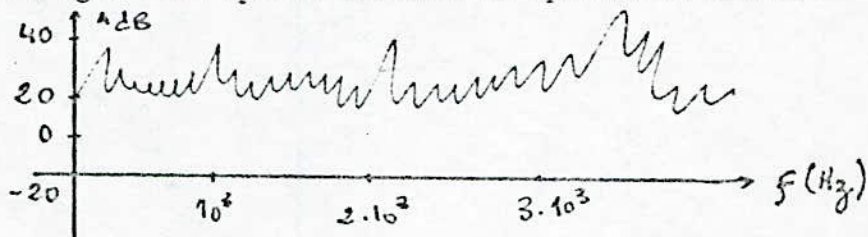


fig. 2 : spectre d'une consonne non voisée : " S ".

Les autres paramètres caractérisent l'évolution de l'enveloppe spectrale et permettent de situer les formants liés aux fréquences de résonance du conduit vocal.

Le schéma de principe général d'un Vocodeur est donné (figure III.)

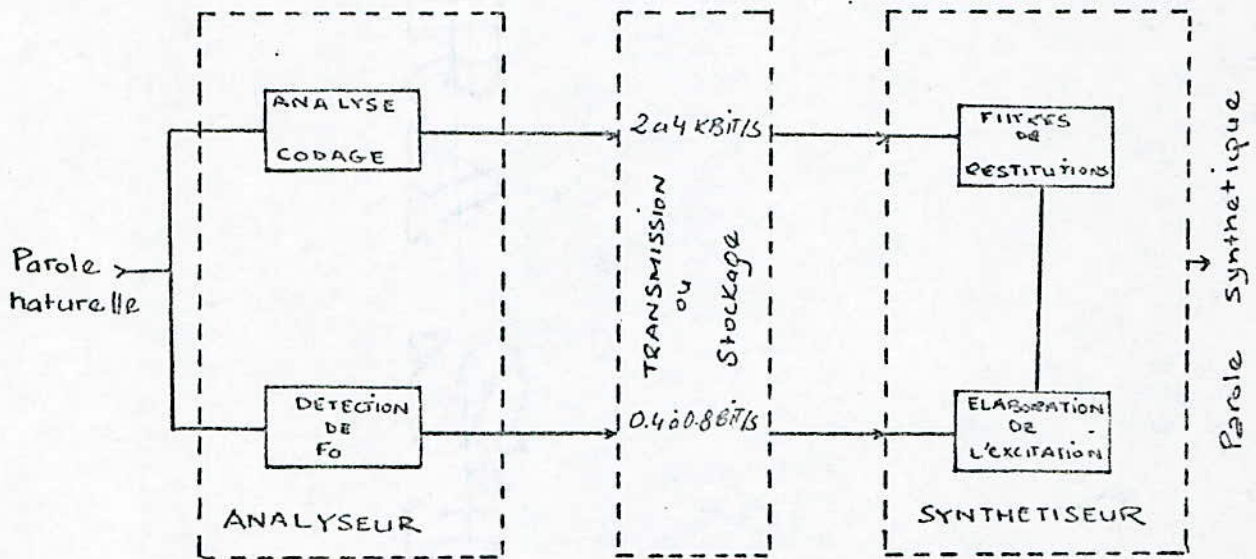
A l'analyse on distingue deux fonctions caractéristiques :

- Détection, mesure et codage de  $F_0$  (caractérisant la structure fine du spectre.).
- Analyse et codage de l'enveloppe spectrale.

A la synthèse on trouve les deux fonctions correspondantes :

- Elaboration d'un signal d'excitation (reproduisant la structure fine du spectre).
- Restitution de l'enveloppe spectrale.

- Schéma de principe général d'un Vocodeur - (fig 3)



Les principaux types de Vocodeurs sont les Vocodeurs à canaux, les Vocodeurs à formants et les Vocodeurs à prédiction linéaire.

## 2. Les principaux types de Vocodeurs :

### 2.1 Le Vocodeur à formants :

#### 2.1.1 Principes :

Au lieu de transmettre des paramètres définissant l'enveloppe d'un spectre instantané, sous forme d'amplitudes se rapportant à des plages de fréquences contigües, réparties sur tout le spectre comme dans le Vocodeur à canaux, on transmet ici des paramètres porteurs d'informations précises concernant l'emplacement, la largeur et l'amplitude des seuls formants. Dans l'hypothèse où l'on sait effectuer la détermination précise et continue de ces paramètres, par une analyse en temps réel, on améliore la qualité du signal transmis (à débit égal) ou on diminue le débit nécessaire.

Mais la détection automatique des formants est un problème très difficile à résoudre. À notre connaissance il n'existe pas actuellement un véritable Vocodeur à formants qui fonctionne d'une manière opérationnelle en temps réel et avec un débit nettement plus faible que celui des Vocodeurs à canaux (ce que le choix de paramètres plus pertinents pourrait faire espérer).

Par contre la technique des synthétiseurs à formants est bien maîtrisée. Les synthétiseurs à formants délivrent une parole de bonne qualité au point de vue agrément d'écoute à condition bien sûr qu'ils soient alimentés par des paramètres décrivant correctement l'évolution des formants. Pour qu'il en soit ainsi une intervention humaine est nécessaire au niveau de la correction des paramètres issus d'une analyse automatique.

### 2.1.2 - Les méthodes d'analyse :

Dans les Vocodeurs à formants le problème de l'analyse de la fréquence fondamentale se pose évidemment de la même manière que dans les Vocodeurs à canaux.

En ce qui concerne l'analyse des formants deux techniques élémentaires sont souvent citées :

- La 1ère consiste à rechercher les 3 premiers maximums du spectre instantané à l'aide d'un banc de filtres. C'est donc une analyse du style de celle du Vocodeur à canaux. Pour avoir une bonne résolution dans la détermination des fréquences des formants, il faut un grand nombre de filtres passe-bande contigus ( entre 30 et 50 ).

Cette analyse en est d'autant plus lourde à réaliser. De plus l'opération reste très difficile à effectuer. De nombreux maximums parasites <sup>peuvent apparaître</sup> dans le spectre. De plus le 1er max peut appartenir à  $F_0$  et non au premier formant.

- La deuxième technique consiste à séparer par filtrage les trois domaines fréquentiels des trois premiers formants. Dans chaque zone on mesure la densité moyenne des passages par zéro et l'énergie, on en déduit la fréquence et l'amplitude du formant correspondant.

Cette méthode ne peut pas donner entièrement satisfaction. Les zones fréquentielles sont déterminées statistiquement, mais les zones d'évolution des formants sont largement variables en fonction des locuteurs et plusieurs formants peuvent apparaître dans une même zone. On signale qu'il est possible de calculer avec précision les fréquences amplitudes et bandes passantes des formants, à partir des coefficients de prédiction linéaire. Mais pour obtenir une précision suffisante de la localisation des formants il faut utiliser une fenêtre temporelle très large, ce qui ne permet pas de suivre les évolutions rapides dans les zones de transition.

### 2.1.3 Les synthétiseurs à formants :

Comme dans le Vocodeur à canaux la synthèse est effectuée à l'aide de filtres passe-bande ( ou résonateur ) attaqués par un signal d'excitation modulé en amplitude. Ici, bien sûr, les filtres sont en nombre réduit puisqu'il s'agit de modéliser les seuls formants. Ces filtres sont variables en ce sens que leurs fréquences centrales ( et quelques fois leur coefficient de surtension ) sont commandées par les paramètres issus de l'analyse de formants. De nombreuses versions ont été étudiées.

On trouve en général une source périodique et une source de bruit qui peuvent être appliquées simultanément avec des pondérations différentes :

On peut trouver par exemple, 3 ou 4 circuits résonnants et un circuit antirésonnant.

Pour les sons voisés, 2 circuits résonnants et ~~pour~~ pour ~~non voisés~~ non voisés. On trouve deux types de structures.

- Dans la structure " // " les filtres sont tous alimentés par le même signal, qu'ils renforcent dans des bandes différentes et leur sorties sont additionnées.

\* " // "  $\equiv$  parallèle



- Dans la structure "serie" ou "cascade" les filtres sont placés en cascade : le filtre F 2 ne peut renforcer que le signal délivré par le filtre F 1, les gains des filtres sont multiplicatifs. En choisissant judicieusement la forme des courbes de réponse et le gain relatif des filtres, on peut satisfaire à une relation statistique entre les fréquences des formants et leurs amplitudes dans la parole. (voir fig 4 et fig 5)

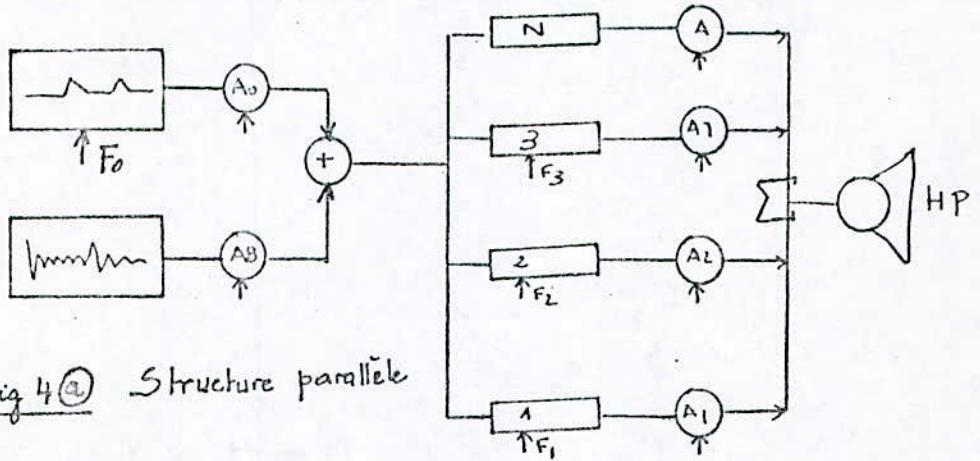
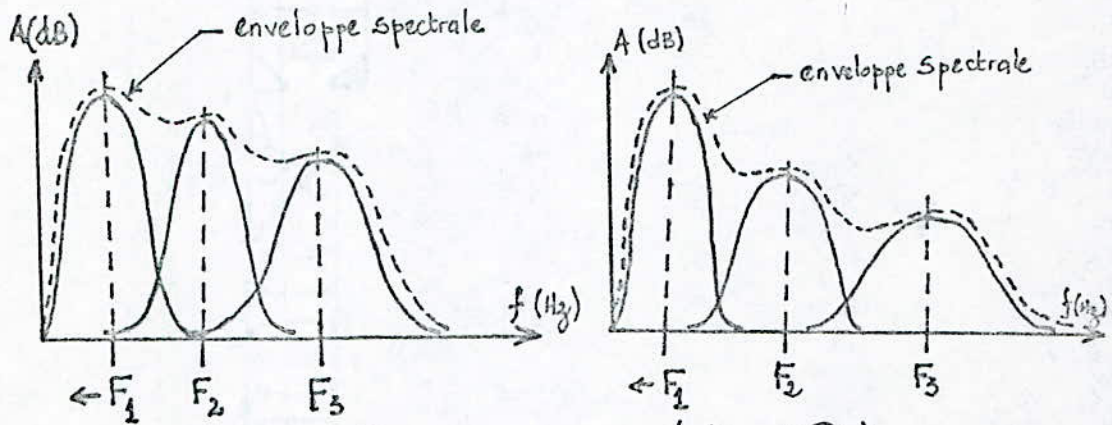


fig 4(a) Structure parallèle



( fig 4(b) ) ( fig 4(c) )  
Types de reponses : (Structure parallèle)

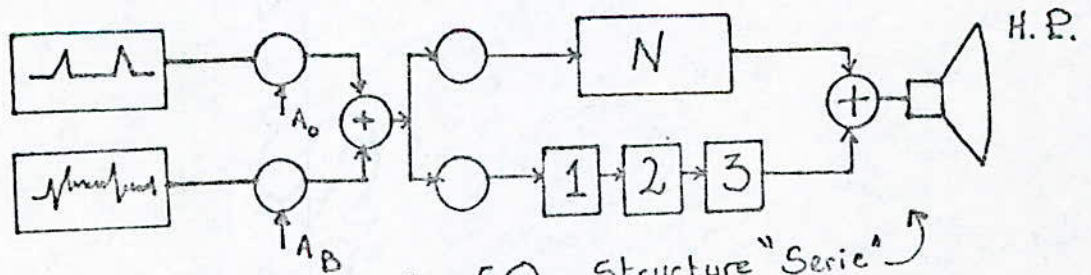


fig 5(a) : Structure "Serie"

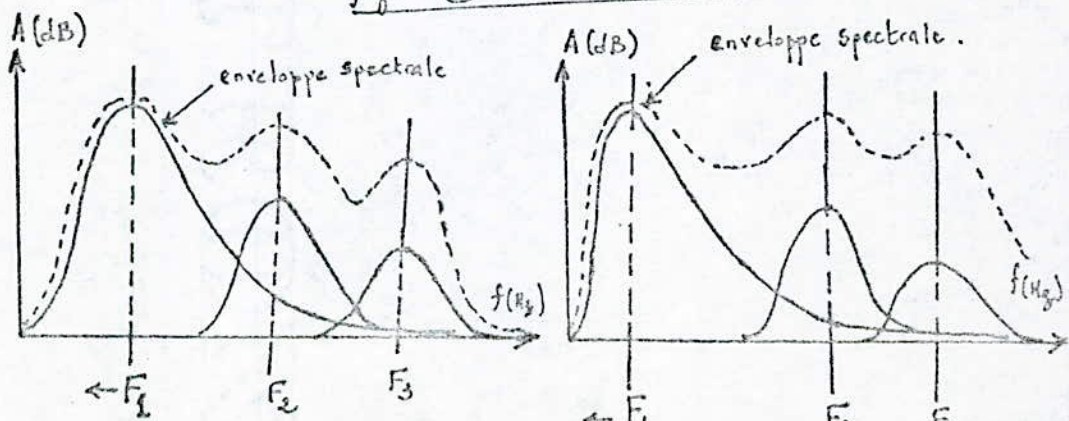


fig 5(b) , 5(c) - Reponses Types.

## 2.2 Le Vocodeur à prédiction linéaire :

### 2.2.1 - Principe :

Dans ce type de Vocodeur le principe est le suivant :  
On assimile le conduit vocal comme étant un filtre récursif dont on doit calculer tous les paramètres qui le caractérisent. Ces paramètres étant les coefficients de prédictions ainsi que la fréquence fondamentale  $F_0$  et ainsi on reconstituera le signal de parole à la sortie du filtre.

En effet, la fonction de transfert du conduit vocal, caractéristique du son émis, peut être représentée par celle d'un filtre numérique dont les paramètres sont calculés toutes les 10 ms. Le nombre de paramètres nécessaire pour une bonne représentation de ce dernier est voisin de 10 : mais cette analyse permet une réduction importante du débit d'informations transmises puisqu'il suffit de ne stocker ( ou de ne transmettre ) que 10 coefficients toutes les 10 ms pour représenter le signal vocal avec une bonne approximation.

Le filtre est plus simple à réaliser sous forme numérique que ceux du Vocodeur à canaux. L'opération peut être faite soit en logiciel ( dans un micro-processeur ), soit en matériel : circuit intégré unique permettant de réaliser cette opération en temps réel.

### 2.2.2 - Analyse :

L'analyseur de ce type de Vocodeur comporte trois blocs de traitement suivants :

#### - Prétraitement :

Consiste à choisir l'intervalle d'analyse, la pondération par une fenêtre et à effectuer la préaccentuation.

- Calcul des paramètres de la fonction de transfert : ces paramètres sont les gains  $G$ , les coefficients  $A(x)$ .

- *Decision* voisée/non voisée, consiste à déterminer si le signal est sourd ou voisé et à donner dans le dernier cas la période de mélodie.

2.2.3 - La synthèse :

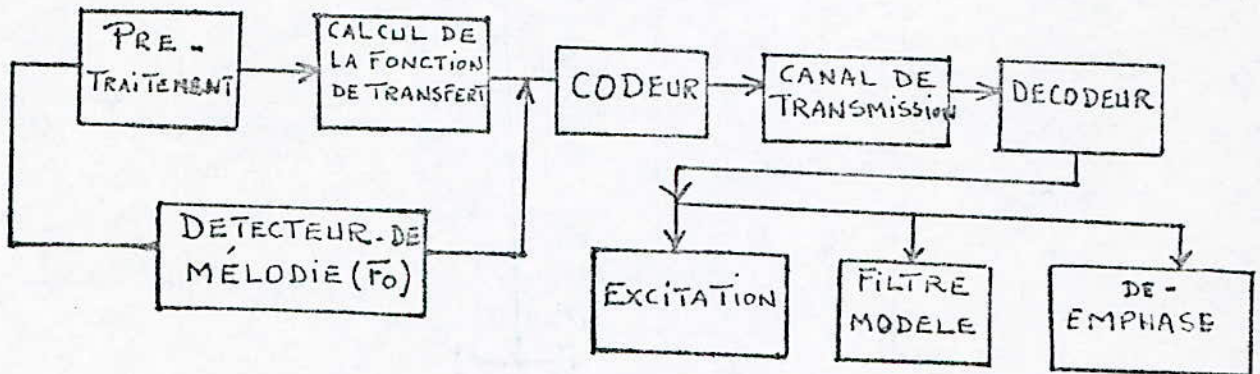
Le signal de la parole sera synthétisé de la manière suivante .  
grâce à une commande binaire et la nature du son selon qui  
soit voisé ou non voisé, le signal d'excitation est produit  
soit par un générateur d'impulsions ou bien un générateur de  
bruit blanc.

Les valeurs des coefficients de prédictions calculés au cours  
de l'analyse permettant de reconstituer le signal, prédit  $X(t)$   
à partir des échantillons  $X(n)$

$$X(n) = \sum_{k=1}^p a_k \cdot x(n-k)$$

La valeur de  $X(n)$  est combinée avec le signal d'excitation  
 $e(n)$  ; lorsque le dernier est présent, on obtient ainsi les  
échantillons du signal de parole à l'instant  $n.T$   
Enfin, un filtre passe-bas permettant la constitution du signal  
original.

fig 6 - Synoptique d'un Vocodeur à prédiction linéaire :



2 . 3 Le Vocodeur à canaux :

2.3.1 Principe :

Dans le Vocodeur à canaux la fonction d'analyse et de codage  
de l'enveloppe spectrale est effectuée à l'aide de canaux dont  
le nombre peut ~~varier~~, suivant les réalisations, entre 10 et 20.  
La figure 7 donne le schéma de principe d'un Vocodeur à  
canaux.

( La fig. 7 est en page suivante )

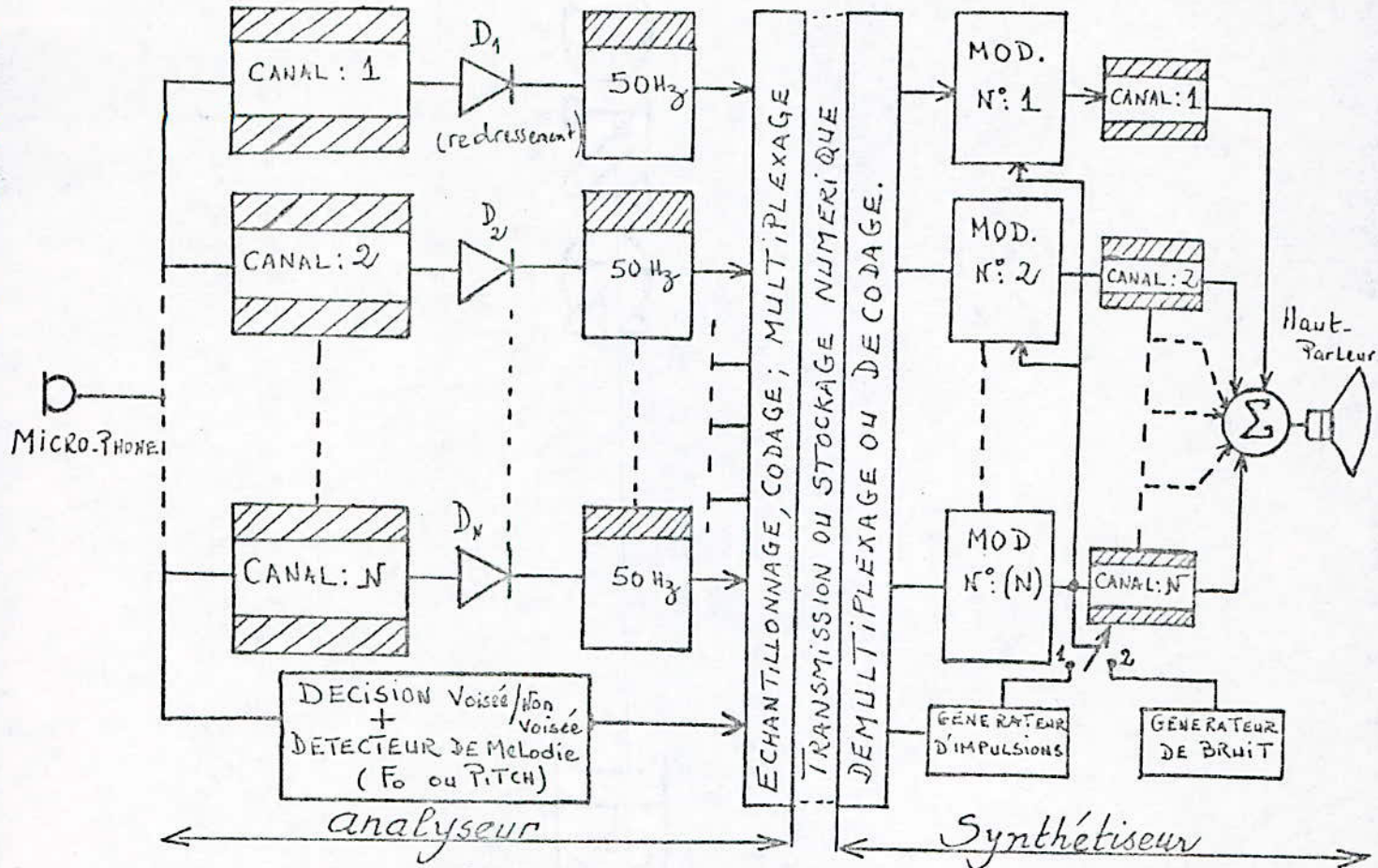


Fig. 7 ; Schéma de Principe d'un Vocodeur à Canaux.

### 2.3.2 L'Analyse :

Chaque canal traite une bande de fréquence déterminée.

Concrètement, l'analyse se fait à l'aide d'un banc de filtres passe-bande contigus.

Chaque filtre est suivi d'un étage de sélection lui-même suivi d'un étage de filtrage passe-bas qui a pour fonction de suivre les variations de l'énergie dans la bande filtrée et de rejeter les composantes spectrales qui tombent dans la bande. Les fréquences de coupures des filtres passe-bas sont de l'ordre de 20 à 50 KHz. Les étages passe-bas sont suivis d'un étage quantificateur. Pour faire cette analyse on fait les hypothèses suivantes :

- L'enveloppe d'un spectre instantané peut être modélisée par 10 à 20 plages fréquentielles à l'intérieur desquelles elle est constante.
- L'amplitude de l'enveloppe dans chaque plage peut être quantifiée avec un nombre faible de bits.
- La fréquence de variation de l'énergie dans chaque plage est inférieure à la fréquence de coupure des filtres passe-bas.

Sous réserve que les largeurs des canaux soient judicieusement choisies, l'expérience montre que ces hypothèses sont assez raisonnables.

Les données quantifiées des canaux sont multiplexées avec celle du détecteur de fréquence fondamentale pour former une trame dont la cadence de renouvellement doit être compatible avec la fréquence de coupure des filtres passe-bas, c'est à dire supérieur au double de cette fréquence.

2.3.3 La Synthèse :

La synthèse est effectuée à l'aide d'un banc de filtres passe-bande identiques ou semblables à ceux de l'analyse. Ces filtres sont excités par un signal d'excitation élaboré à partir des données concernant la détection  $F_0$ , le signal d'excitation doit avoir un spectre plat dans la bande de fréquence occupée par les filtres passe bande. Pour les sons voisés c'est un spectre de raies harmoniques de  $F_0$ . Pour les sons non voisés c'est un spectre de bruit. L'entrée de chaque filtre passe bande de synthèse est modulée en amplitude en fonction de l'énergie mesurée derrière le filtre d'analyse correspondant, le signal de sortie est obtenu par addition des sorties des filtres de synthèse.

## CHAPITRE IV

### METHODES GLOBALES POUR LA RECONNAISSANCE ET LA SYNTHÈSE DE LA PAROLE

#### PREAMBULE :

Le travail qui nous a été proposé ne concerne pas l'étude et la réalisation de la partie reconnaissance et celle de synthèse d'un vocodeur à canaux, notre objectif étant l'analyseur de celui-ci. Mais au cours de cette réalisation il s'avère indispensable de connaître les différentes méthodes qui permettent la reconnaissance et la synthèse du signal de la parole, afin de saisir *du* mieux, les raisons pour lesquelles, *on associe* tel carte ou système de reconnaissance, avec tel autre synthétiseur.

Nous précisons donc que notre intention n'est pas de définir toutes les méthodes employées tant elles sont nombreuses, mais d'expliquer au mieux l'essentiel de ces deux problèmes encore délicats et complexes aujourd'hui.

#### I - LA RECONNAISSANCE DE LA PAROLE :

##### I.1 - DEFINITION :

La reconnaissance de la parole est la faculté pour une machine de reconnaître, puis de réagir à la voix d'un système, qu'il soit monolocuteur ou multilocuteur. Le but que se fixe un système de reconnaissance est de soulager l'utilisateur du problème de transcodage, lorsqu'il vient faire exécuter un ordre par celle-ci. Le succès de tels dispositifs est sans conteste lié à leurs performances, leurs capacités de concurrencer, les autres dispositifs de communication dont nous disposons actuellement tels que, les claviers, les dispositifs graphiques "HOMMES MACHINES", les consoles de visualisation etc....

Ils devront apporter une simplification considérable au travail de l'opérateur et lui permettre dans le cas d'une erreur, une correction rapide et facile.

##### I.2 - LE PROBLEME TEL QU'IL SE POSE :

Le problème de la reconnaissance vocale est des plus épineux, c'est en fait le plus délicat comparé à ceux de l'analyse et de la synthèse. C'est pourquoi, les chercheurs n'ont pu le résoudre que pour des cas qui sont particuliers. Les méthodes mathématiques, et les différents algorithmes, se heurtent aux aléas posés par la pratique, car possédant peu de connaissances scientifiques sur les processus physiologiques <sup>ou</sup> psychologique mis en œuvre pour la compréhension du message vocal.

En termes clairs, il s'agirait simplement de pouvoir donner à une machine une oreille et une partie du cerveau associée, pour qu'elle puisse reconnaître et comprendre le message. Hélas ce n'est pas chose aisée car malgré les progrès prodigieux des composants dits de la 5ème génération soumis à des contraintes financières *trop fortes* pour être vulgarisés, en fait le but n'est pas ~~encore~~ près d'être atteint. Certes on n'arrête pas le progrès ou la recherche, qui ont pris aujourd'hui une nouvelle orientation nommée "l'intelligence artificielle". Elle constitue à elle seule une science, mais est encore à l'état embryonnaire.

### 1.3 - TECHNIQUES DE RECONNAISSANCES :

Au niveau des techniques de reconnaissances, on trouve deux types de méthodes complémentaires.

La méthode de reconnaissance dite globale, qui utilise des dispositifs de reconnaissance acoustique, dits globaux, ces derniers mémorisent et comparent entre eux des mots codés sous forme de paramètres acoustiques.

La méthode de reconnaissance dite analytique qui utilise des systèmes de reconnaissance phonétique, analytiques. Ceux-ci ont pour but d'extraire de la parole prononcée tous les phonèmes, et comparent entre elles des chaînes phonétiques.

En fait, il est difficile de préférer une méthode, un procédé par rapport à l'autre, tant les avantages de l'une des méthodes, de l'un des procédés pallient aux inconvénients de l'autre. Les systèmes globaux ont pour avantages le fait d'être quasiment indépendants de la langue. En contre partie ils sont limités par la capacité mémoire nécessaire aux codages des mots, et par le fait qu'ils sont le plus souvent monolocuteurs.

Les systèmes analytiques permettent de traiter théoriquement des vocabulaires importants (1000 mots ou plus), mais ils sont dépendants des caractéristiques phonétiques d'une langue. De plus la conception et l'utilisation de tels systèmes est souvent lourde, donc très onéreuse. Dans tout système de reconnaissance automatique de la parole, quelque soit la technique employée on rencontre les éléments suivants.

- Un analyseur acoustique destiné à extraire du signal des informations pertinentes, et à réduire le débit à moins de 5000 bits/sec.

- Un processeur de calcul, responsable de la reconnaissance proprement dite. Le plus souvent les opérations sont très répétitives, des architectures spécialisées sont alors requises (processeurs dynamiques: ex: TMS 320).

- Une memoire de tres grande capacite pour conserver les informations necessaires au systeme par exemple : pour un vocabulaire de 100 mots codes a 4800 bits /sec , il faut un espace memoire de 32 Koctets correspondants a 256 Kbits .

#### 1.4 - APPRENTISSAGE DES SYSTEMES :

On appelle apprentissage ,la procedure que doit suivre un utilisateur ,avant de pouvoir utiliser un systeme de reconnaissance .L'apprentissage differe selon que le systeme est monolocuteur ou multilocuteur .

- Pour un systeme multilocuteur ,aucun apprentissage n'est demande a l'utilisateur .Le vocabulaire reconnu par le systeme ,est immediatement memorise par ce dernier .

- Pour un systeme monolocuteur de type global ,le locuteur doit prononcer au mois une fois ,tous les mots de l'application durant l'apprentissage .Ceci est a la fois tres simple pour un vocabulaire tres reduit ,et devient vite tres eprouvant pour un grand vocabulaire.

- Pour un systeme monolocuteur de type analytique il suffit de prononcer quelques mots ou phrases de l'application .

#### 1.5 - LES PROBLEMES LIES AUX DEUX TYPES D'APPROCHES :

Dans le probleme de la reconnaissance nous distinguons maintenant clairement deux types d'approches :

##### a - L'approche global :

Elle consiste a considerer le mot ou le groupe de mots, comme une entite que l'on cherche a reconnaitre sans la scinder en ses constituants phonetiques .

##### b - L'approche analytique :

Celle-ci consiste a segmenter le message et identifier ainsi ses constituants (phonemes) .

Le premier probleme qui se pose aux deux types d'approches est la reconnaissance de la parole continue :

\* Si l'on a choisi le processus global ,donc reconnaissance par entites isolees (mots ou groupe de mots) comment alors creer ,l'articulation de ces dernieres apres les avoir reconnues pour donner l'effet de continuité ? Cela se fait naturellement chez l'homme mais comment l'adapter a une machine ? (effet bien connu des discontinuités de la parole chez les robots parlants).



\* Si l'on choisit le deuxième procédé le problème est le même car, comment faudrait-il co-articuler les phonèmes entre eux ?

On comprend bien, qu'à cause de la continuité du signal vocal les difficultés surgissent, la principale étant la détermination des frontières entre unités linguistiques. D'autres difficultés liées à la segmentation et à la co-articulation, des mots ou phonèmes, suscitent l'intérêt des chercheurs, nous ne citerons que les principales.

\* La segmentation va-t-elle précéder la phase de reconnaissance ou sera-t-elle menée en parallèle ? En particulier comment sera déterminé le nombre de mots dans une phrase ?

\* L'importance des effets de co-articulation. Seront-ils introduits dans le reconnaiseur de façon algorithmique ?

\* La recherche de la suite de mots, réalisant la meilleure coïncidence avec la phrase à reconnaître.

### 1.6 - DESCRIPTION D'UN SYSTEME DE RECONNAISSANCE GLOBAL :

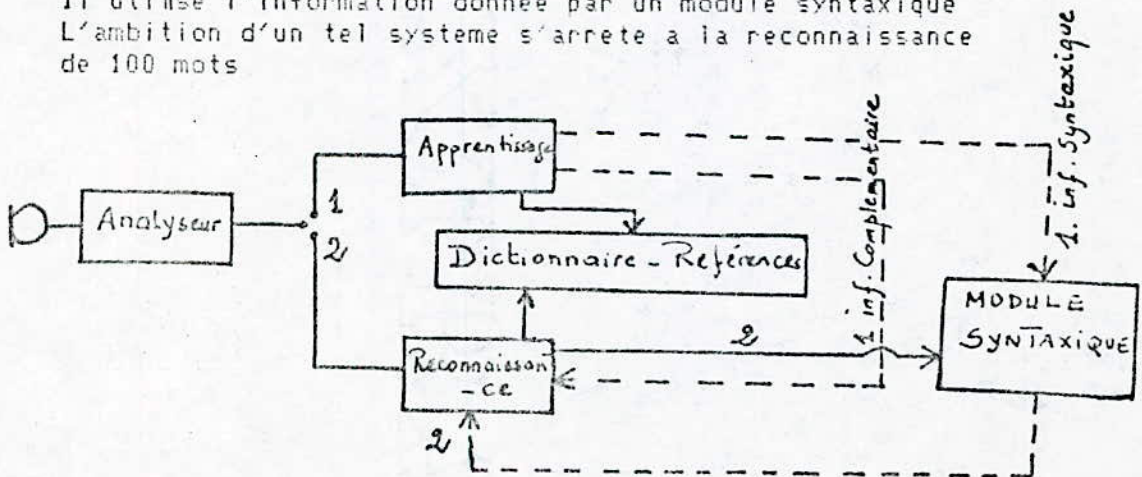
La figure ci-dessous résume l'organisation générale d'un système de reconnaissance global dont il faut distinguer deux étapes.

#### 1.6.1 - L'étape d'apprentissage :

Elle consiste à mémoriser les unités de références constituants ainsi un dictionnaire de mots de références. L'analyseur assure l'extraction des paramètres pertinents et réduit le débit pour la mémorisation. Selon l'application cette étape peut permettre l'introduction d'informations complémentaires telles que : règles de segmentations, de co-articulation etc.....

#### 1.6.2 - L'étape de reconnaissance :

Le reconnaiseur établit des distances entre références (segments de parole); il identifie la phrase inconnue. Il utilise l'information donnée par un module syntaxique. L'ambition d'un tel système s'arrête à la reconnaissance de 100 mots



( fig n°: 1)

1.7 - METHODES MATHÉMATIQUES :

Celles-ci sont tellement nombreuses qu'il est impossible de les citer toutes, l'objectif est d'en détailler une. Les principales méthodes sont :

- a - Etablissement d'une distance entre blocs de paramètres
- b - segmentation des mots dans une phrase
- c - Comparaison dynamique par alignement temporel
- d - Programmation dynamique

1.7.1 - Etablissement d'une distance entre blocs de paramètres :

Elle utilise les coefficients de l'analyse cepstrale faite au préalable dans l'analyseur. Dans cette approche un segment de parole est représenté par une succession d'échantillons porteurs de paramètres pertinents. Un échantillon est représenté par un vecteur. Et l'on entend par distance, la norme euclidienne entre deux vecteurs, pour apprécier la similarité entre les intervalles de paroles auxquels ils sont associés. Elle doit donc satisfaire :

- \*  $d(x,y) > 0$  quand  $x$  est différent de  $y$
- \*  $d(x,y) = 0$  quand  $x$  est égale à  $y$
- \*  $d(x,y) = d(y,x)$  (symétrie)
- \*  $d(x,y) < \text{ou} = d(x,z) + d(z,y)$

Cette méthode opère une comparaison entre deux spectres soit :

$S1(\text{EXP}(j\omega))$ , le premier et  $S2(\text{EXP}(j\omega))$  le second. La méthode consiste à évaluer la différence (erreur) ; entre les logarithmes des amplitudes de ces deux spectres et l'intégrer sur leur domaine d'existence. Soit  $D(\omega)$  la fonction erreur.

$$D(\omega) = \text{LOG} / S1(\text{EXP}(j\omega)) / - \text{LOG} / S2(\text{EXP}(j\omega)) /$$

Les distances sont alors définies par la formule :

$$(d_p)^p = \frac{1}{2n} \int_{-n}^{+n} |D(\omega)|^p \cdot d\omega$$

Où  $\omega$  est la fréquence réduite et où  $n$  représente la moitié de la fréquence d'échantillonnage.

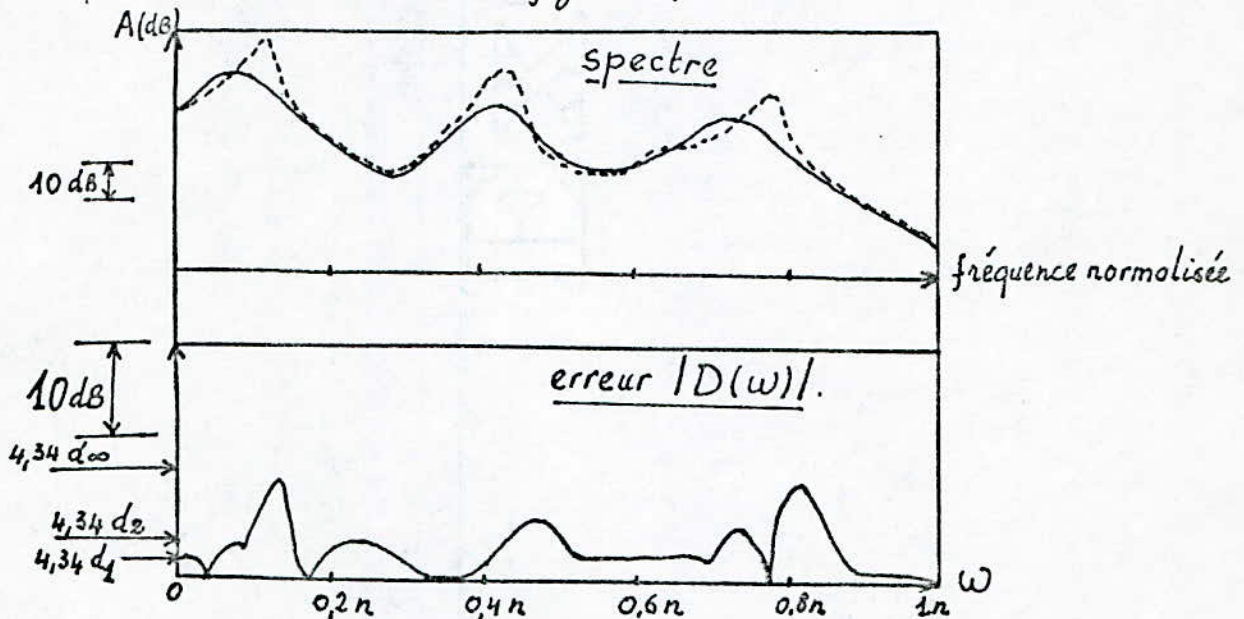
$$n = \frac{f_e}{2}$$

Pour  $p=1$  on definit :  $d_1 = \frac{1}{2n} \int_{-n}^{+n} |D(\omega)|. d\omega$

Pour  $p=2$  on definit :  $(d_2)^2 = \frac{1}{2n} \int_{-n}^{+n} |D(\omega)|^2. d\omega$

$d_1$  represente l'erreur moyenne, ( $d_2$ ) representant l'erreur quadratique moyenne,  $p$  designe alors en même temps le choix de la distance et l'ordre de l'erreur, d'autre part si on augmente  $p$  on diminue les erreurs importantes.

Exemples de comparaison obtenue avec la methode de prediction lineaire. (fig n°: 2)



La distance etablie precedement, doit etre conforme a la representation cepstrale on montre en fait si les  $C(k)$  sont les coefficients cepstraux du 1er spectre et  $C'(k)$  sont ceux du second, alors :

$$d = \sum_1^M |C(k) - C'(k)|$$

est equivalente a la distance precedente.

## II - SYNTHESE DE LA PAROLE :

### II.1 - DEFINITION :

Le probleme de la synthese est l'aptitude pour une machine de restituer la parole reconnue sous forme synthetique .

### II.2 - LES DIFFERENTES TECHNIQUES DE SYNTHESE :

Les methodes de synthese sont aussi nombreuses que celle de la reconnaissance .Dans leur multitude,nous avons choisi d'en citer les principales et d'en developper une particulierement .Celles-ci sont :

- a - Synthese par mots
- b - Synthese par diphones
- c - Synthese par regles
- d - Synthese a partir du texte par modification et concatenation de formes d'ondes

Cette derniere methode a ete proposee par :

CHARPENTIER.F et STELLA.M et recemment publiee par le C.N.E.T .

Elle consiste a faire une synthese a partir du texte ,en utilisant les diphones comme unites de parole .

Le dictionnaire de diphones est constitue des formes d'ondes ,elles memes munies d'un etiquetage des periodes de la frequence fondamentale  $F_0$  .Au cours de la synthese les formes d'ondes sont traitees par un systeme dit : systeme d'analyse-synthese permettant un controle independant de tous les parametres prosodiques ,tout en conservant un timbre de parole naturelle .

Ce systeme d'analyse synthese part d'une representation du signal de la parole,par sa transformee rapide de Fourier (F.F.T) a une cadence d'analyse pitch-synchrone. L'algorithme de synthese fonctionne par recouvrement et addition de petits signaux elementaires a court-terme , ce qui garantit une concatenation douce des formes d'ondes .La parole obtenue par cette methode possede une qualite plus naturelle que celle obtenus par prediction lineaire .L'approche des methodes par recouvrement et addition ( OVER LAP AND ADD, OLA par abreviation ) ,est preferable pour ces algorithmes plus rapides .

Les methodes OLA fonctionnent toutes par decomposition du signal en une suite de signaux elementaires a court terme (signaux CT) .

Ces signaux CT sont ensuite convenablement modifies et additionnes a une cadence differente .Celle-ci est variable et est nommee OLA synchrone (SOLA) ,qui permet de conserver une parole de bonne qualite ,en modifiant la vitesse d'articulation .

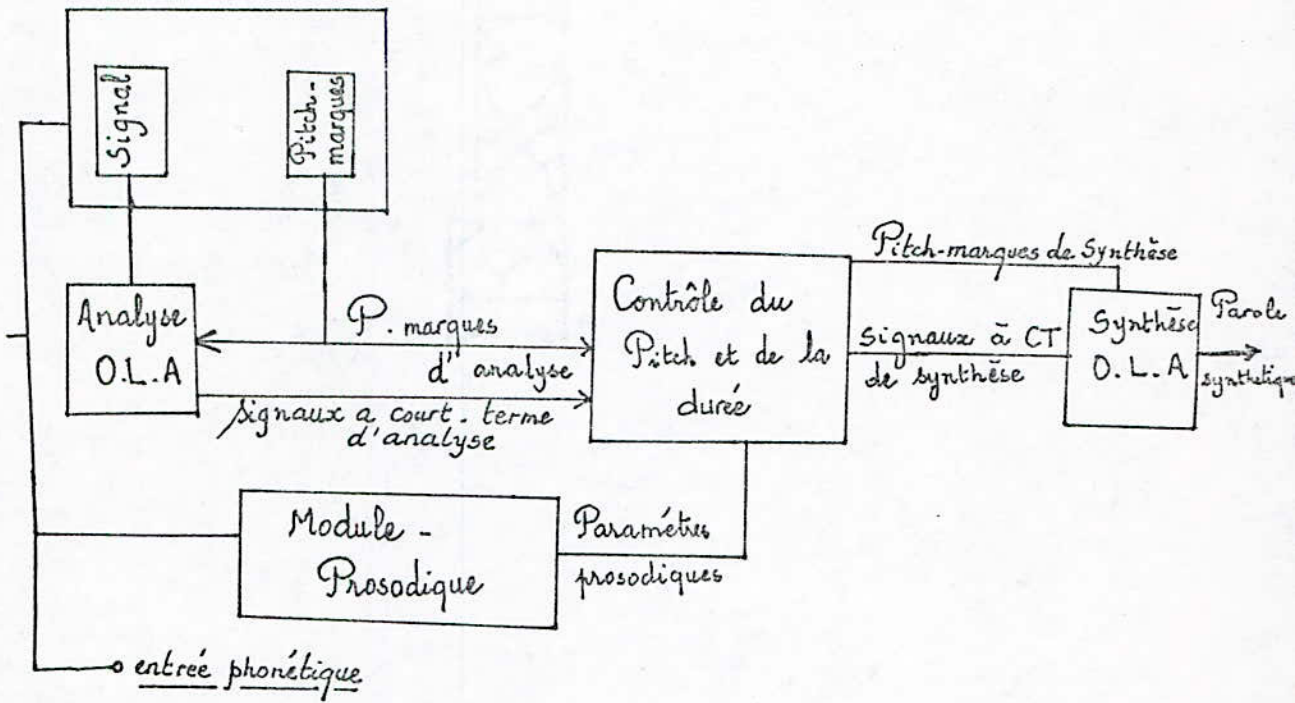


fig n°: 3 Schema d'organisation générale d'un système de synthèse par diphones

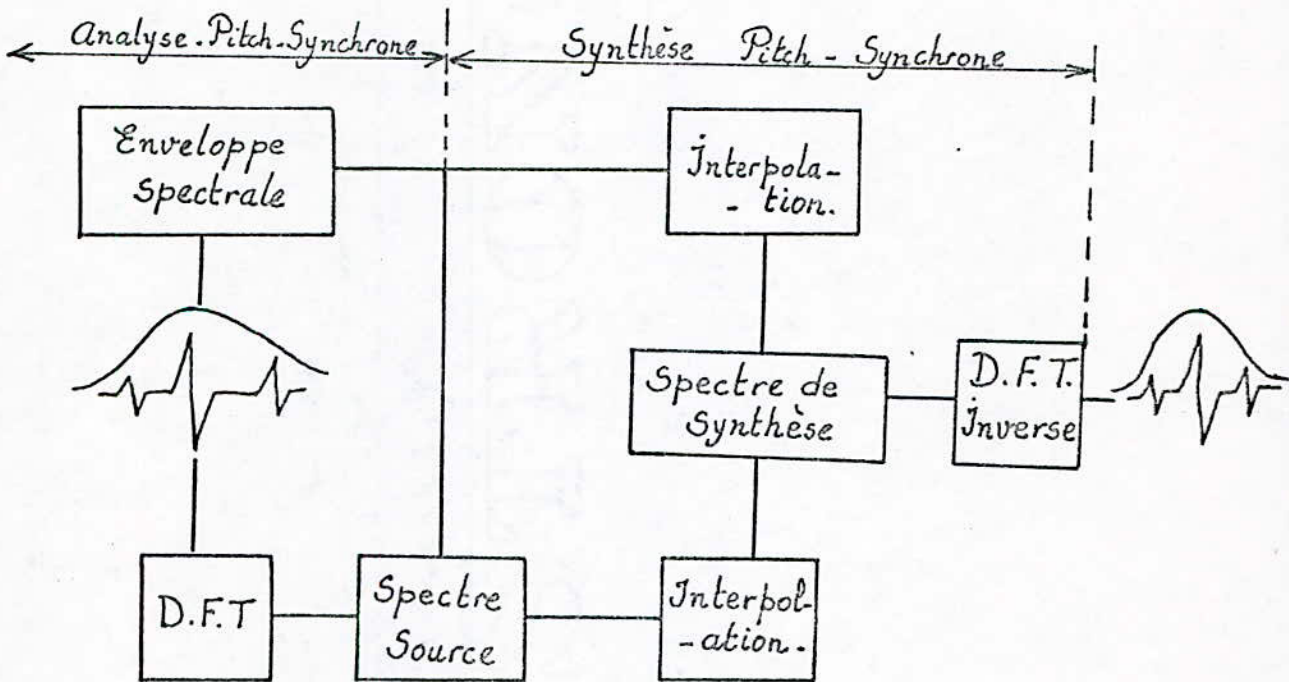


fig n° 4 : Description de L'algorithm: (P.S.O.L.A.)

L'intérêt de cette cadence non uniforme est la resynchronisation des signaux CT avec les périodes effectives du signal de synthèse. La méthode présentée par les chercheurs est la (PSOLA : PITCH SYNCHRONOUS OVER LAP ADD) qui généralise la SOLA en introduisant le pitch synchronisme de l'étape d'analyse. Le principe est de faire coïncider chaque signal CT avec une période de pitch, que ce soit à l'analyse ou à la synthèse. Une étape donc préliminaire est nécessaire afin de segmenter le signal en période de pitch consécutives. L'innovation apportée par la PSOLA consiste à intégrer dans le processus de recouvrement, et additionner un algorithme de modification du pitch et de l'enveloppe spectrale.

CHAPITRE V :

ETUDE DE LA REALISATION D'UN VOCODEUR A CANAUX :

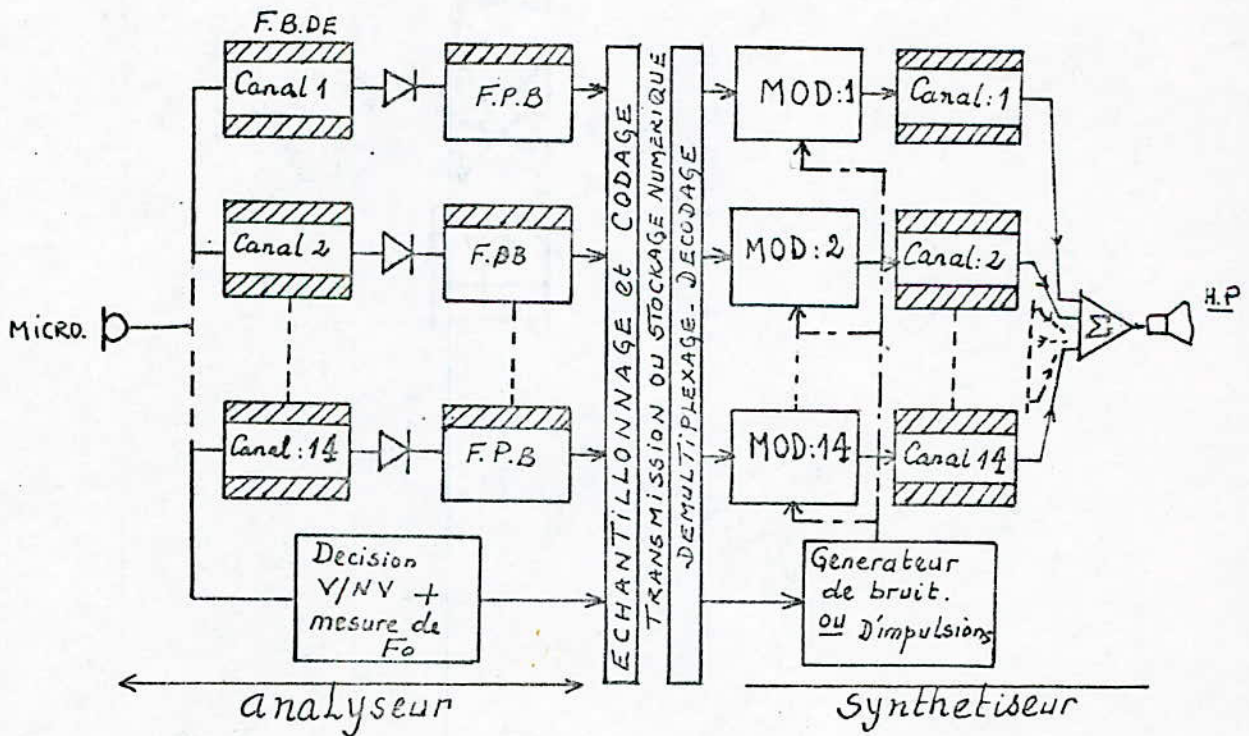
INTRODUCTION :

Ce chapitre aborde maintenant, l'étude pratique d'un vocodeur à canaux, ainsi que les différents schémas de sa réalisation. La première partie de ce travail consiste en l'étude détaillée d'un canal.

Nous y abordons le problème du filtrage avec la nécessité du choix de l'échelonnement des filtres, nous nous attachons à l'importance des normes et des critères qui ont permis leur conception, la détection et la numérisation y sont également étayées. Le rôle de chaque étage y est dégagé, afin d'expliquer la hiérarchie des étages constituant un canal donné. Ainsi numérisé, le signal est envoyé à une unité permettant le traitement et le stockage des données disponibles sur le convertisseur, pour cela nous avons pensé utiliser la carte du microprocesseur 6802 (KIT MOTOROLA) tout en la programmant de sorte qu'elle puisse déclencher les convertisseurs et stocker les données en mémoire.

La seconde partie de ce chapitre a été consacrée à l'étude et réalisation du détecteur de mélodie, qui est en fait un moyen de calcul de la fréquence fondamentale d'un son voisé. Nous ne pouvons ignorer la nécessité du bloc, détecteur de voisement qui, constitue un organe très important pour la décision voisée, non voisée et par manque de documentation

sur ce dernier, nous nous sommes limités simplement à son étude théorique.



synoptique du Vocodeur à 14 Canaux.

## 1 - ETUDE D'UN CANAL :

Introduction : Chaque canal de l'analyseur est constitué d'un filtre passe bande de fréquences déterminées priviliégiant certaines harmoniques et atténuant fortement d'autres, d'une détection, ou d'un redressement mono-alternance (on ne s'intéresse qu'à l'enveloppe des crêtes positives du signal) d'un filtre passe bas sévère qui sert à lisser le signal en supprimant les tensions résiduelles, d'un étage de conversion analogique/numérique ou se passe simultanément, l'échantillonnage, le blocage et le codage du signal.

L'ensemble des étages décrits dans cet ordre constitue un canal. Sont ainsi disposés en parallèle 14 canaux, dont les étages sont identiques aux différences près des bandes fréquentielles allouées à chacun d'entre eux. Cet ensemble forme le premier sous-bloc de l'analyseur, dans lequel sont extraits les formants et les paramètres spectraux de l'analyse.

### 1.1 - ETUDE D'UN CANAL :

#### 1.1.1 - Etude du filtrage :

Le filtrage est ici d'ordre actif, et sert à ne laisser passer que les fréquences comprises entre 200 Hz et 4300 Hz (voir plus loin l'échelonnement des filtres).

L'harmonique fondamentale est aiguillée vers le bloc détecteur de mélodie ou de pitch. Il nous est apparu ici nécessaire de clarifier ainsi l'étude du filtrage par un rappel de notions sur les filtres actifs.

#### 1.1.2 - Notions sur les filtres actifs :

Les filtres actifs sont constitués d'éléments passifs tels que : (R, C), et d'éléments actifs tels que : (A OP...). Le plus souvent on utilise des ampli-opérationnels pour des raisons de prix et d'encombrements. D'autre part ces filtres présentent un certain nombre d'avantages comme :

- a - Leur impédance d'entrée élevée
- b - Leur impédance de sortie qui souvent est faible, ce qui permet de faire des montages en cascade
- c - Offrent des possibilités de réglage du gain et de la fréquence d'accord par des moyens simples.

#### 1.1.3 - Principaux types de filtres :

Pour les besoins d'une comparaison des caractéristiques de ces filtres nous nous limitons qu'aux principales structures bien connues :

- a - La structure de Butterworth, bien connue pour la stabilité de son gain dans la bande passante.
- b - La structure de Tchebytchev.



a - Structure de Butterworth :

Elle offre l'avantage d'une reponse plate dans la bande passante ,et est definie par la fonction de Butterworth dont la formule est:

$$\left| \frac{v_s}{v_e} \right| = \left[ \frac{1}{1 + \left( \frac{\omega}{\omega_0} \right)^{2n}} \right]^{1/2}$$

$\omega_0$ : represente la pulsation de coupure  
 $n$  : represente l'ordre du filtre

La linearite de la reponse est suivie d'une chute assez reguliere du gain qui est d'environ : 6 dB/Octave  
 (Voir page suivante le trace des types de reponses de Butterworth pour differents ordres avec la table des coefficients pre-etablis)

b - Structure de Tchebytchev :

Les filtres actifs de Tchebytchev sont definis par la fonction de celui-ci qui est :

$$\left| \frac{v_s}{v_e} \right| = \left( \frac{1}{1 + \varepsilon^2 T_n^2 \left( \frac{\omega}{\omega_0} \right)} \right)^{1/2}$$

ou,  $\varepsilon$  represente un coefficient compris entre 0 et 1 et qui exprime l'ondulation  
 $n$  : represente l'ordre du filtre

$T_n \left( \frac{\omega}{\omega_0} \right)$  : represente le polynome d'ordre  $n$

En operant le changement de variable  $\omega/\omega_0 = x$  , la fonct. s'ecrit alors :

$$\left| \frac{v_s}{v_e} \right| = \left[ 1 + \varepsilon^2 T_n^2(x) \right]^{-1/2}$$

ou  $T_n(x)$ : represente le polynome transforme

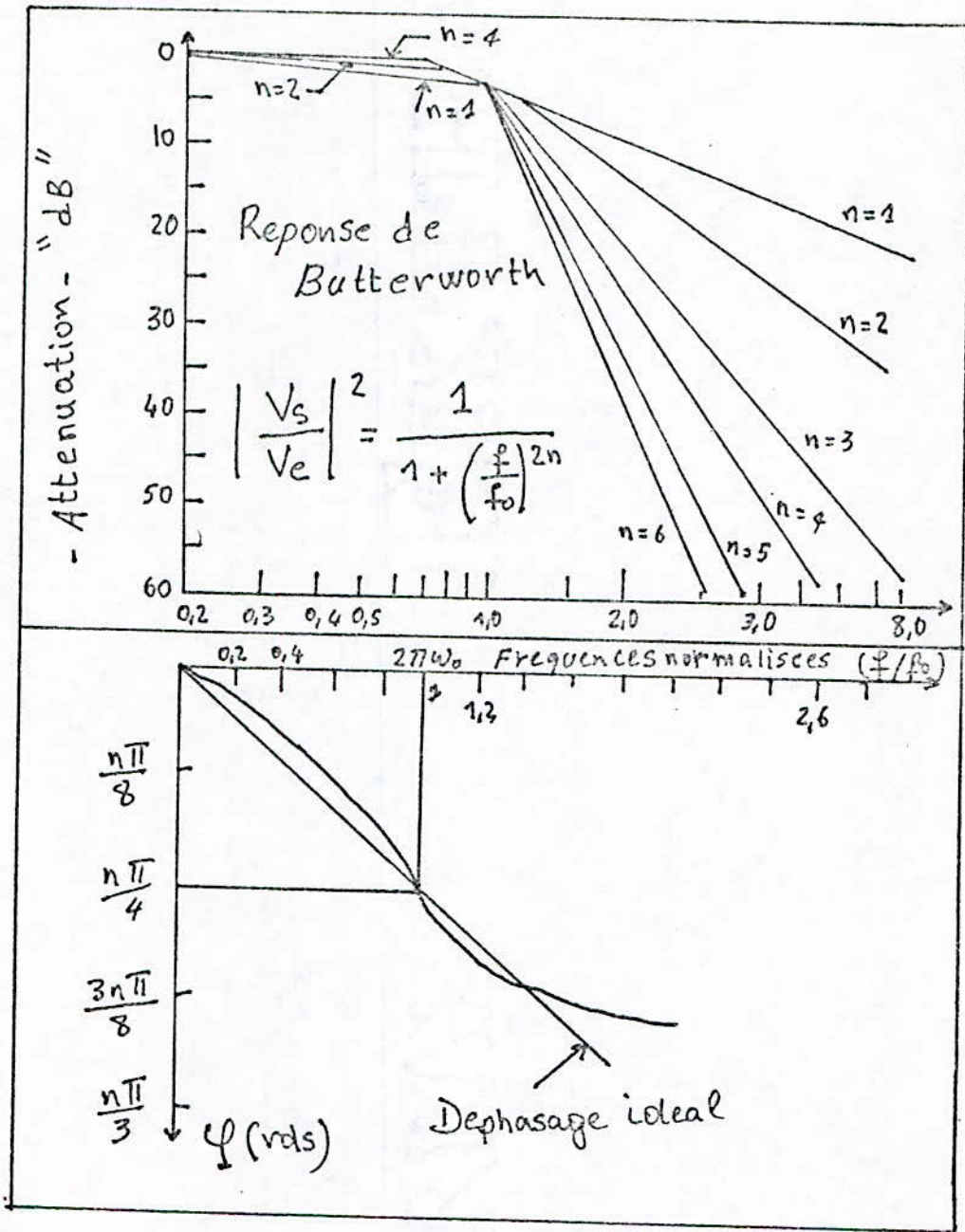
Nous rappelons que les polynomes de Tchebytchev s'ecrivent alors :

$$\begin{cases} T_n(x) = \cos (n \operatorname{ARCOS} (x)) & \text{pour } |x| < 1 \\ T_n(x) = \operatorname{CH} (n \operatorname{ARGCH} (x)) & \text{pour } |x| > 1 \end{cases}$$

ou CH :represente le cosinus hyperbolique et ARG designe l'argument hyperbolique .

Ces polynomes prennent une forme plus generale en les rearrangeant entre eux on arrive aux relations suivantes

$$\begin{cases} T_n(x) = 2x T_{n-1}(x) - T_{n-2}(x) & \text{pour } |x| < 1 \\ T_n(x) = 0.5 (T_{2n}(x) + 1) & \text{pour } |x| > 1 \end{cases}$$



Ordre 4		Cellule 1		Cellule 2	
ondulation	chute en dB à $2f_0$	$m_1$	$q_1$	$m_2$	$q_2$
0 dB	24	0,923	1,082	0,38	2,61

Ordre 2		Cellule 1	
ondulation	chute en dB à $2f_0$	$m$	$q$
0 dB	12	0,707	1,414

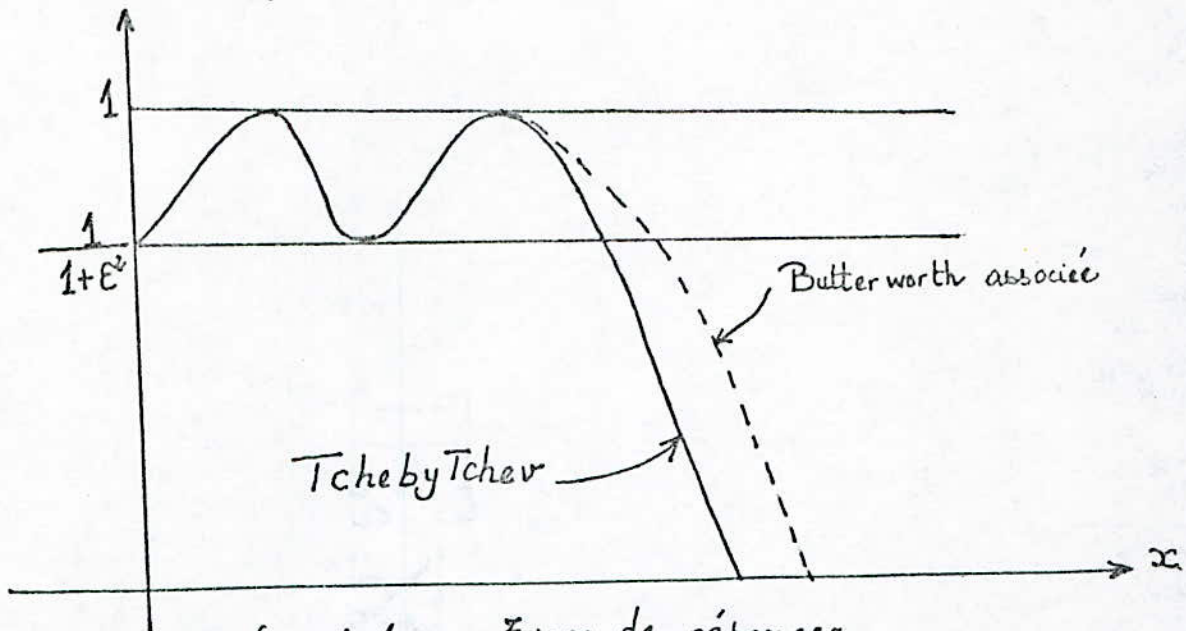


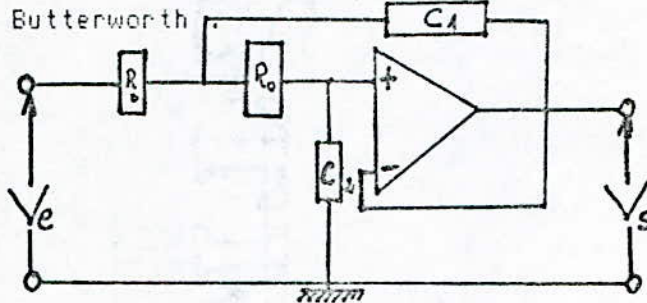
fig n°: 1 : Types de réponses

est plus forte → L'atténuation pour les fréquences élevées dans le cas, d'une structure de TchebyTchev que dans la structure de Butterworth du même ordre. Malheureusement les filtres de TchebyTchev ont pour désavantage, une très mauvaise réponse à un échelon de tension dans la bande passante. Ils sont surtout utilisés lorsque le signal utile a une fréquence fixe, stable, ce qui n'est pas le cas pour le signal de la parole stationnaire pendant quelques ms,

c'est pourquoi nous avons adopté la structure de Butterworth.

1.1.4 - Prototype passe-bas de Butterworth :

Le schéma suivant résume le principe d'un filtre passe bas de Butterworth.

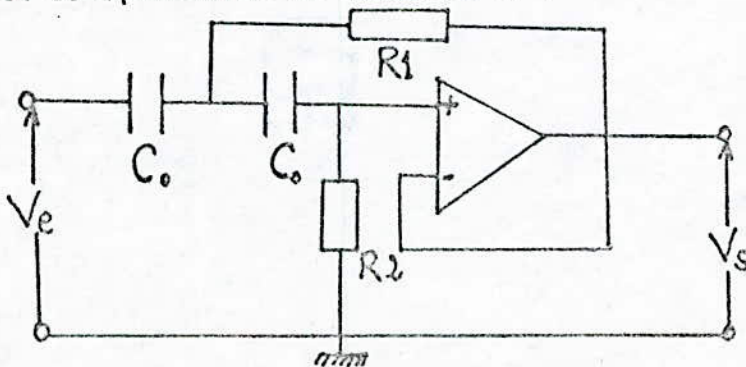


cellule d'ordre 2 dont la fonction de transfert est donnée par :

$$\left| \frac{V_s}{V_e} \right| = \left( \frac{1}{1 + \left(\frac{\omega}{\omega_0}\right)^4} \right)^{1/2}$$

1.1.5 - Prototype passe-haut de Butterworth :

Le schéma suivant résume le principe d'un filtre passe haut du 2ème ordre de Butterworth.



En fait une relation de biunivocite existe et permet le passage du filtre passe bas au filtre passe haut du meme ordre , il suffit pour cela de remplacer les resistances par des capacites et vice-versa . En effet la transposition de frequence  $p \rightarrow 1/p$  , permet d'ecrire si  $H(p)$  est la fonction de transfert du passe bas , la transposition etant operee on peut noter :

$$H(p) = H(1/p) \quad (\alpha)$$

Soit  $\omega_1$  la pulsation pour laquelle le prototype passe bas a pour fonction de transfert :  $|H(\omega_1)|$  pour module et  $\angle H(\omega_1)$  pour phase . L'expression  $(\alpha)$  s'ecrit alors :

$$|H(j\omega)| = |H_b(1/j\omega)| = |H_b(j\omega_1)|$$

Et par suite :  $1/j\omega = +$  ou  $- j\omega_1$  il vient alors  $\omega = +$  ou  $- 1/\omega_1$  .

A la frequence  $\omega_1 = 0$  il correspond  $\omega \rightarrow \infty$  et vice-versa, donc la frequence  $\omega$  ne peut correspondre qu'a celle du filtre passe haut associe .

### 1.1.6 - Filtres passe bande adopte :

En mettant en cascade un filtre passe haut , de frequence de coupure  $f_{c2}$  et un passe bas de frequence de coupure  $f_{c1}$  nous construisons ainsi un filtre passe bande dont les frequences admises sont comprises dans la bande passante  $B = f_{c2} - f_{c1}$  .

La structure adoptee est la suivante :

Le passe haut est constitue de deux cellules du 2<sup>eme</sup> ordre, pour ameliorer la raideur de l'attenuation du gain du simple au double lui conferant ainsi l'ordre 4 , en cascade avec un filtre passe bas

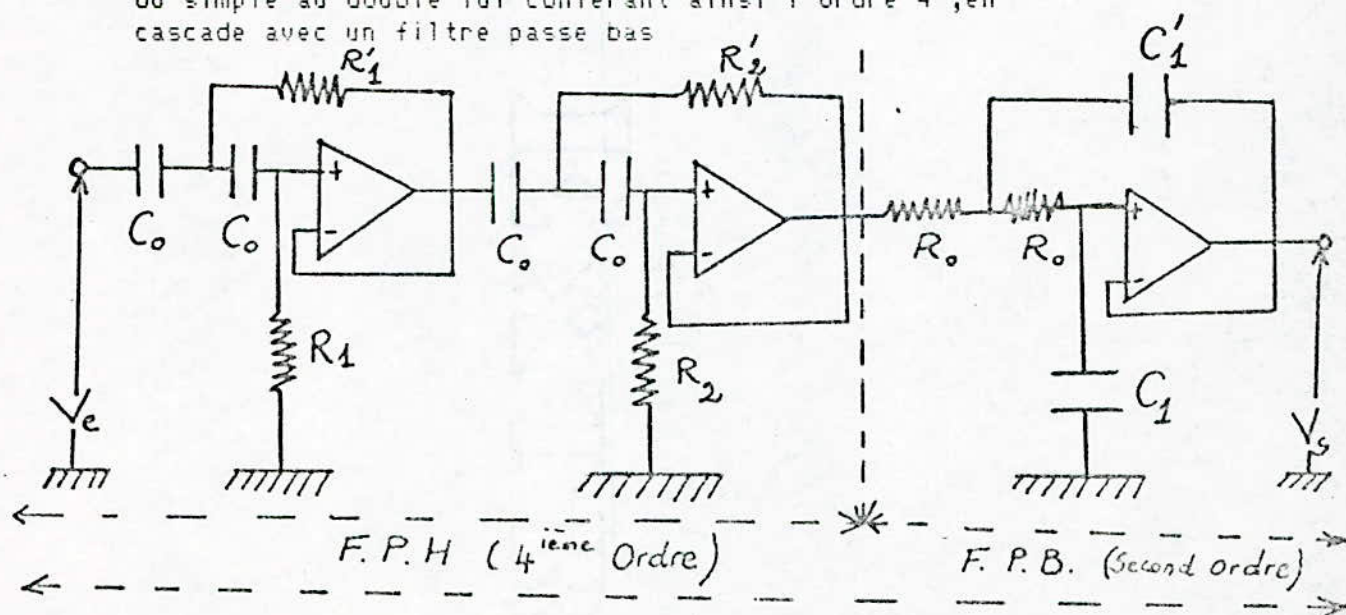


fig n°: 2 Structure Passe - Bande d'un canal

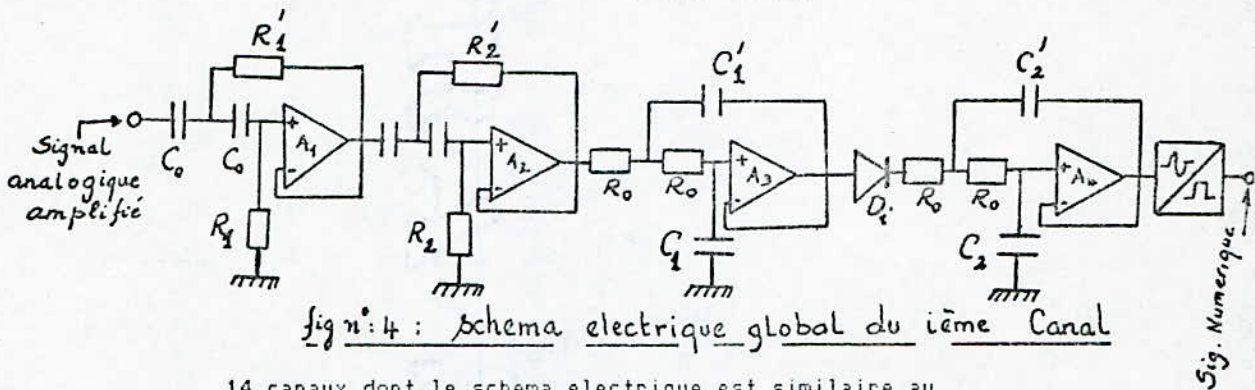
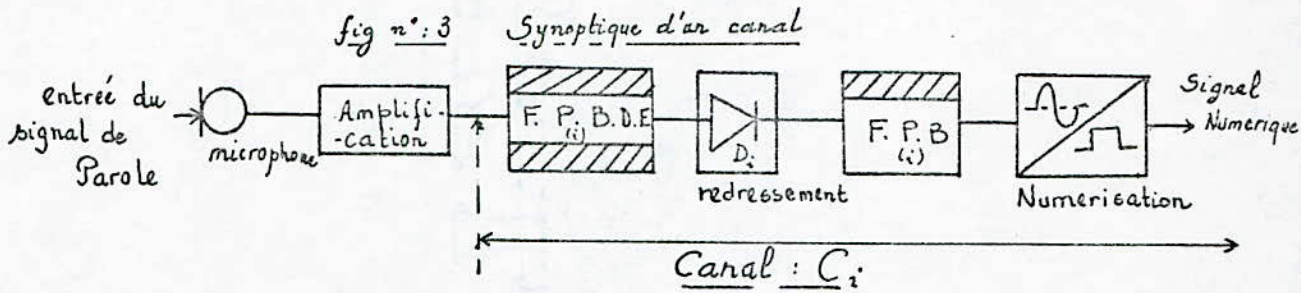
Le calcul des elements des filtres ,sont fait a partir des frequences de chaque canal .Le detail de ces calculs est donne au chapitre VI .

1.2 - DETECTION OU REDRESSEMENT :

Nous nous interessons uniquement a l'enveloppe de cretes positives du signal ,aussi l'etage de detection est constitue d'un montage a diode unique operant un redressement monoalternance .

1.3 - SCHEMA DE PRINCIPE GENERAL D'UN CANAL :

Pour l'instant un canal quelconque  $C_i$  peut etre schematise de la facon suivante :



14 canaux dont le schema electrique est similaire au precedent ,sont dispose en parallele pour couvrir la bande 200 - 4300 Hz du signal ,et assure ainsi une analyse spectrale sans pertes d'energie .Une perte en energie spectrale impliquerait automatiquement une perte en harmonique et donc en qualite ou intelligibilite du signal restitue .Pour pallier cet inconvenient ,il convient de choisir un echelonnement des filtres autre que lineaire ,sinon le probleme qui se poserait alors serait le suivant :

Supposons que nous ayons choisi un echelonnement lineaire,et fixons l'exemple a deux canaux par la figure suivante :

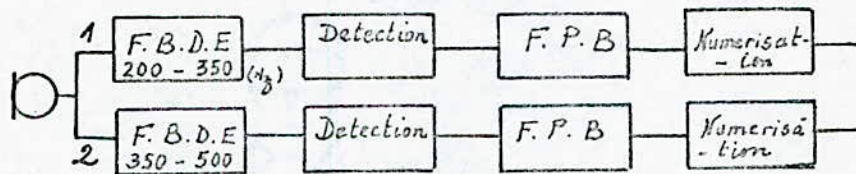
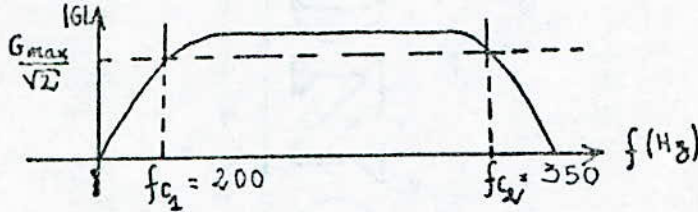
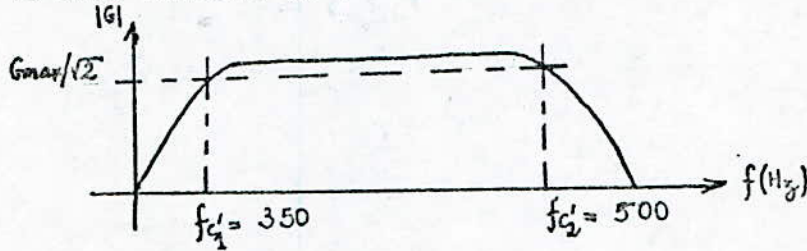


fig n°: 5 Problème de l'echelonnement Lineaire

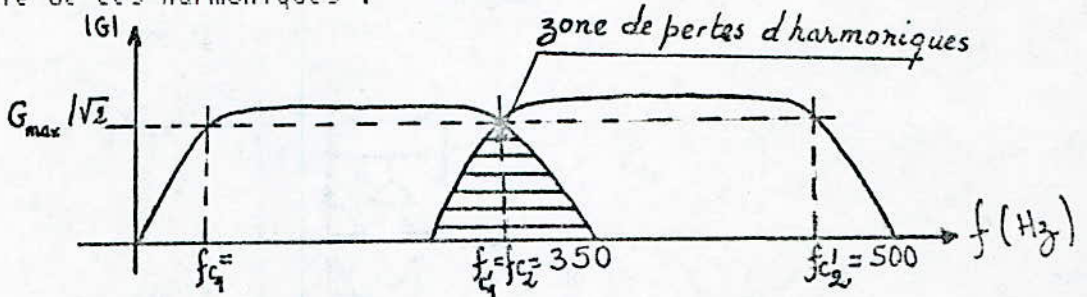
Le premier filtre passe bande travaille donc dans la zone lineaire de :



Le deuxieme filtre passe bande du canal 2 travaillerait dans la zone lineaire de :



En fait pour une portion du signal qui agirait sur les deux filtres simultanement, se trouverait amputee d'une partie de ces harmoniques .



Et l'intelligibilite et la qualite du son rendus s'en trouveraient affectes . C'est pourquoi, le probleme etant situe, d'autres echelonnements ont ete trouves tel que ceux logarithmiques afin d'apporter une solution . L'objet du prochain paragraphe est de montrer l'evolution de ces derniers et l'avantage d'en choisir un plutot qu'un autre.

#### 1.4 - CHOIX DE L'ECHELONNEMENT DES FILTRES :

Les filtres contenus dans chaque canal sont constitués ( voir étude précédente ) de filtres passe bande ( filtre passe haut et filtre passe bas en cascade ) et de filtres passe bas servant au lissage de la courbe qui est filtrée, ils présentent tous une réponse de Butterworth, les passes bandes répondant au 4<sup>ème</sup> ordre, les passes bas au second ordre, on trouvera dans le chapitre calculs, schémas, mesures les réponses de ces deux types de filtres.

L'échelonnement de ces filtres a partir de la largeur de la bande de fréquence choisie en général de 200 à 4300 Hz environ a subi une évolution permanente au cours du temps. Le 1<sup>er</sup> défini et essayé au C.N.E.T LANNION FRANCE était logarithmique. Les fréquences centrales voisines des filtres étaient définies par la relation suivante :

$$f(i+1) = 10^{\frac{1}{12}} f(i)$$

On alloua, alors à chaque canal une largeur ou une bande de fréquence, dont les fréquences extrêmes (de coupures) répondaient à la formule précédente, et on répartit la bande 200 - 4300 Hz ainsi sur 14 canaux. Le tableau qui est présenté sur la page suivante donne le détail de cette répartition.

Les essais qui furent effectués par le centre d'étude des télécommunications C.N.E.T firent ressortir les avantages de cet échelonnement à savoir :

a - Les coefficients de surtension étaient les mêmes, donc tous les filtres avaient les mêmes éléments résistifs.

b - La résolution de l'analyse dans le bas du spectre de fréquence était très fine, le 1<sup>er</sup> formant était localisé avec précision cependant un tel échelonnement présentait aussi des inconvénients et fut abandonné au profit d'un deuxième échelonnement créé en 1970. Les filtres en dessous de 1000 Hz qui recouvraient en gros le 1<sup>er</sup> formant furent jugés trop étroits. Dans le bas du spectre de fréquence on ne trouvait en moyenne, qu'une raie harmonique de la fréquence fondamentale tout les deux ou trois canaux pour les voies masculines, et tout les trois ou cinq canaux pour celles féminines.

Canal numero	frequence -centrale $F_0$ (Hz)	frequences (Hz) de coupures $\bar{a}$ - 6dB	$\Delta F = B \bar{a}$ (Hz) - 6dB	$\frac{F_0}{\Delta F} \equiv Q$
1	274	248 - 301	52	5,2
2	332	301 - 365	64	5,2
3	402	365 - 442	77	5,2
4	487	442 - 536	94	5,2
5	590	536 - 649	113	5,2
6	715	649 - 787	138	5,2
7	866	787 - 953	166	5,2
8	1049	953 - 1155	202	5,2
9	1271	1155 - 1399	244	5,2
10	1540	1399 - 1695	296	5,2
11	1866	1695 - 2054	359	5,2
12	2260	2054 - 2488	434	5,2
13	2739	2488 - 3014	526	5,2
14	3318	3014 - 3652	638	5,2

Tableau n° 1: Echelonnement n° 1



Le tableau 2 donne la repartition en fréquence de l'échelonnement 2

Canal $x^{\circ}$	fréquence: $F_0$ (Hz)	fréquences de coupures $f_c$ à (-6dB)	$\Delta F_{\bar{a}} - 6dB$	$F_0 / \Delta F$
1	335	250 - 450	200	1,7
2	541	450 - 650	200	2,7
3	743	650 - 850	200	3,7
4	945	850 - 1050	200	4,7
5	1168	1050 - 1300	250	4,7
6	1442	1300 - 1600	300	4,8
7	1744	1600 - 1900	300	5,8
8	2045	1900 - 2200	300	6,8
<del>9</del>	2345	2200 - 2500	300	<del>7,8</del>
10	2645	2500 - 2800	300	8,8
<del>11</del>	2946	2800 - 3100	300	9,8
12	3294	3100 - 3500	400	8,2
13	3695	3500 - 3900	400	9,2
14	4095	3900 - 4300	400	10,2

Tableau n° 2 - échelonnement n° 2

Cet échelonnement apporta pour gros avantages l'élargissement de la bande globale des fréquences, sans changer le nombre de canaux en portant ainsi l'augmentation à 650 Hz environ mais l'inconvénient qu'on lui reprochait, était celui d'avoir des bandes excessives pour les premiers canaux ce qui entraînait alors des coefficients de surtension et des pentes des filtres trop faibles.

Et toujours d'après la même source C.N.E.T à 670 Hz par exemple l'atténuation apportée par le 1er canal était de 21 dB, autrement dit si un maximum du spectre de fréquence tombait aux alentours de 700 Hz il était impossible de trouver un minimum correct autour de 300 Hz (limite se trouvant dans la bande du 1er canal) à une distance fréquentielle pourtant de 400 Hz, la résolution de l'analyse dans le bas du spectre saurait donc être insuffisante. Ces constatations ont amené les chercheurs à définir un 3ème échelonnement dit échelonnement III dans le détail est donné dans le tableau 3.

Cet échelonnement se devait de concilier en créant un compromis autant que possible entre les avantages des précédents, tout en évitant leur défaut.

Canal <sub>n°</sub>	$F_0$ (Hz)	$f_c: \begin{matrix} i=1,2 \\ \bar{a} - 6dB \end{matrix}$	$\Delta F_{\bar{a} - 6dB}$	$F_0/\Delta F$
1	296	250 - 350	100	2,95
2	410	350 - 480	130	3,15
3	554	480 - 640	160	3,45
4	729	640 - 830	190	3,85
5	934	830 - 1050	220	4,25
6	1168	1050 - 1300	250	4,65
7	1433	1300 - 1580	280	5,1
8	1728	1580 - 1890	310	5,6
9	2053	1890 - 2230	340	6,05
10	2408	2230 - 2600	370	6,5
11	2793	2600 - 3000	400	7
12	3200	3000 - 3430	430	7,45
13	3653	3430 - 3890	460	7,95
14	4128	3890 - 4380	490	8,5

Tableau n°3 : Echelonnement 3

L'échelonnement (3) fut créé en 1979 au C.N.E.T LANNION et fut introduit dans le vocodeur C.N.E.T de la même année. Il se caractérise par une progression arithmétique des écarts des fréquences  $F$ . Y'avait le même rapport largeur de bande globale sur le nombre de canaux que le précédent. L'atténuation dans le 1er filtre passe bande du 1er canal pour une fréquence valant le double de la fréquence centrale était de 31 dB soit une amélioration de 10 dB par rapport à l'échelonnement (2), d'autre part il a été observé pour celui-ci, une régularité dans la progression des rapports  $F_0/\Delta F$  et des coefficients de surtension.

1.4.1 - Test de comparaison et choix de l'échelonnement :

Le dilemme pose par les différents échelonnements dans leurs avantages et inconvénients maintenant cernés a savoir :

a - L'élargissement de la bande globale sans changer le nombre de canaux .

b - La résolution de l'analyse spectrale plus ou moins fine afin de localiser le 1er formant avec précision , dans le bas du spectre .

c - L'évolution et les rapports entre les valeurs extrêmes des coefficients de surtension ont conduit les chercheurs du C.N.E.T a opérer des tests d'écoutes , ces tests comparatifs avaient pour but de dégager la meilleure écoute qu'offrait chacun d'entre eux en réalisant une étude statistique optimale sur les propriétés subjectives de l'oreille .

Ainsi on s'attela a faire prononcer un certain nombre de phrases en utilisant des locuteurs masculins et féminins et on enregistrait alors simplement l'indice préférentiel d'écoute des trois échelonnements .

Les résultats obtenus sont ceux donnés dans le tableau quatre

	globalement	Locuteurs masculins	Locuteurs féminins
Préférence : Ech. 3	82 %	95 %	69 %
Préférence : Ech 2	15 %	5 %	25 %
Pas de Préférence	3 %	0 %	6 %

Tableau : 4

\* RESULTATS DU TEST D'ECOUTE PORTANT SUR LES ECHELONNEMENTS 2 ET 3 \*

Il apparut alors nettement que le 3<sup>ème</sup> échelonnement était le meilleur . Cette préférence était beaucoup plus forte pour les locuteurs masculins pourtant les voies féminines se caractérisent par une concentration plus accentuée de l'énergie dans le bas du spectre . On conclut alors que ce n'était l'allure moyenne de l'enveloppe spectrale mais plutôt moyen des raies harmoniques de la fréquence fondamentale et l'échelonnement 3 favorisant d'avantage les voies masculines a cause de l'espacement moyen des 1<sup>ers</sup> canaux qui étaient compatibles avec l'espacement moyen des 1<sup>eres</sup> raies harmoniques du fondamentale ce qui n'était pas le cas avec l'échelonnement 2 .

Pour préciser cette idée, un trace des  $F = f(F_0)$  pour les différents échelonnements fut alors fait et pour mieux s'approcher des propriétés de l'oreille il fut adjoint sur le même trace un échelonnement comparatif, correspondant aux bandes critiques (Voir fig 6). D'après les courbes de la figure 6, on montra alors que pour se rapprocher encore plus des propriétés de l'oreille un 4ème échelonnement fut alors créé mais celui-ci demandait l'adjonction de deux canaux supplémentaires portant ainsi leur nombre à 16, pour répondre au mieux aux exigences de l'oreille, le tableau 5 en donne la répartition car l'échelonnement 3 ne répondait plus aux exigences de l'oreille en haut du spectre des fréquences aussi il subsista donc un doute dans son choix. Un examen plus approfondi sur l'échelonnement 2 fut alors repris pour déterminer le taux de corrélation que pouvait avoir ce dernier dans la détermination des formants.

( Pour plus d'illustration se référer au tableau 5 de la page suivante ).

On fit alors répéter à des cobayes les cinq phrases contenues dans le tableau 6 ci-dessous et on jugea une nouvelle fois du pourcentage des préférences de l'échelonnement, les résultats enregistrés dans le tableau 6 furent les suivants.

Préférence : Ech. 2	Phrase
37 %	"Le Charretier fouette son cheval"
21 %	"le sculpteur taille la Pierre "
18 %	"le Fantôme hante le château"
15 %	"le docteur ordonne un médicament"
9 %	"le crédit est consenti par la banque"

Tableau : 6

Tout se passe comme si les sons voisés tel que "CH" et "S" qui ont un maximum d'énergie dans le haut du spectre. Pour la phrase " Le charretier ....." prononcée par un locuteur féminin 45 % des préférences vont à l'échelonnement 2, 37 % vont à l'échelonnement 3, pour 18 % de jugements indéterminés.

Il ressort de cette étude que l'insuffisance de l'échelonnement 2, n'a pas été certainement déterminée.

Devant ces résultats, et tout en sachant que le signal de la parole n'est pas entièrement localisé dans le bas du spectre, ni totalement dans le haut du spectre et dans le cadre de notre réalisation, circonscrite surtout aux voyelles orales, que les 1ers formants sont localisés en moyenne autour de (1000 Hz et plus); notre choix s'est fixé sur l'échelonnement 2 pour l'avantage de sa résolution dans le haut du spectre.

dont on sait que

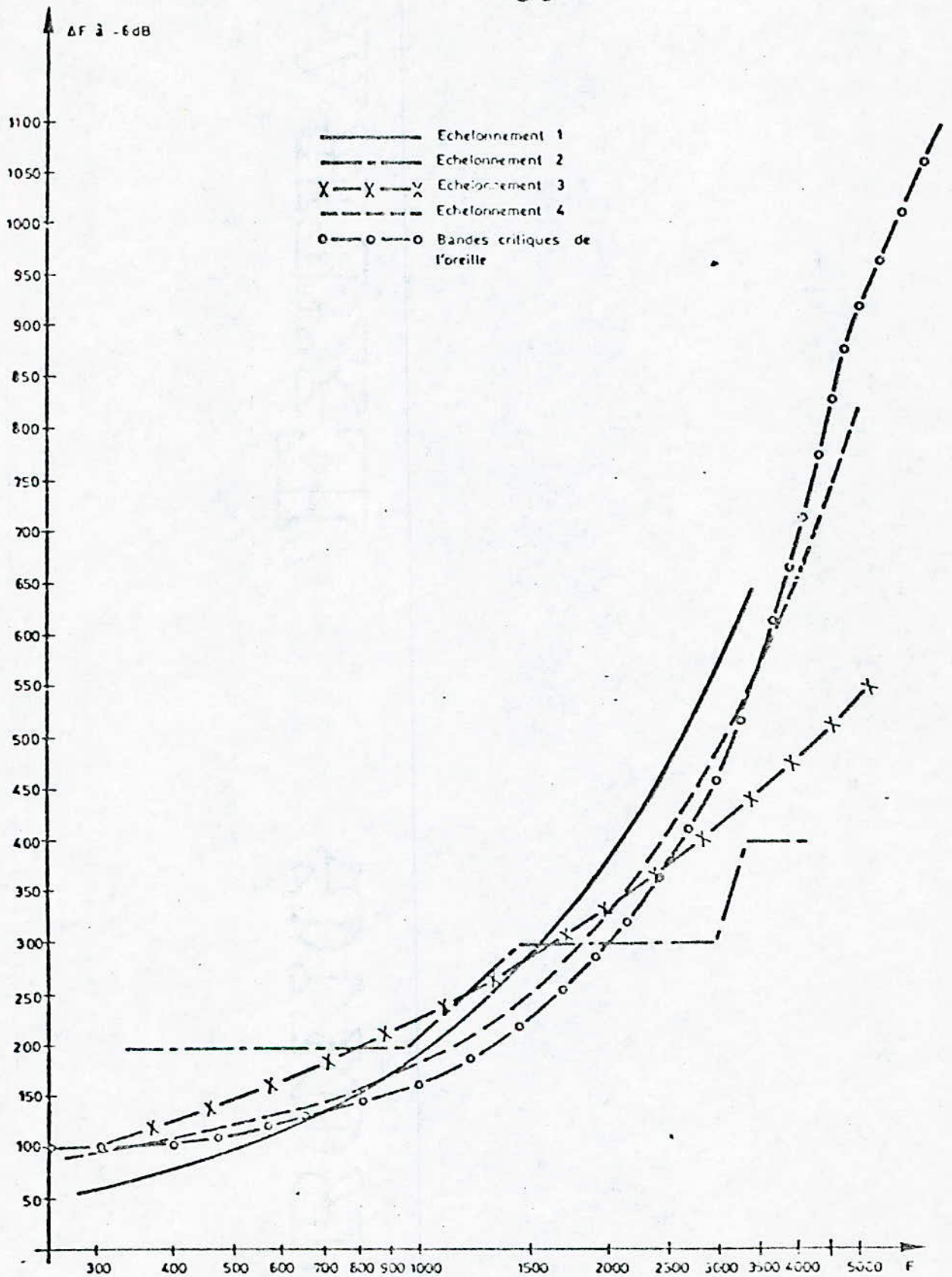


FIGURE N° 6 : Evolution des  $\Delta F$  en fonction de  $F$  pour 4 échelonnements de filtres vocodeur et pour l'échelonnement des bandes critiques de l'oreille.

$\Delta F$  variations as a function of  $F$  for the 4 different versions of the vocoder channels and for the channels with the critical band characteristics of the ear.

Canal Numero	$F_0$ (Hz)	$F_c \bar{a} - 6 \text{ dB}$ (Hz)	$\Delta F$ (Hz) $\bar{a}$ - 6 dB	$F_0/\Delta F$
1	261	220 - 310	90	2,90
2	359	310 - 415	105	3,42
3	471	415 - 535	120	3,93
4	599	535 - 670	135	4,43
5	741	670 - 820	150	4,94
6	901	820 - 990	170	5,30
7	1083	990 - 1185	195	5,55
8	1290	1185 - 1405	220	5,87
9	1527	1405 - 1660	255	6
10	1804	1660 - 1960	300	6
11	2130	1960 - 2315	355	6
12	2516	2315 - 2735	420	6
13	2972	2735 - 3230	495	6
14	3510	3230 - 3815	585	6
15	4146	3815 - 4505	690	6
16	4896	4505 - 5320	815	6

Tableau 5 : Echelonnement (4)

### 1.5 - NUMERISATION DU SIGNAL :

La numerisation du signal est realisee par convertisseur analogique numerique (C.A.N) plus precisement celui de la firme NATIONAL (ADC 0804) .Le detail concernant son schema interne et son fonctionnement sera donne plus loin .

Cette partie de numerisation necessite aussi un echantillonneur bloqueur a cause des variations rapides du signal de la parole .

Le role de l'echantillonneur bloqueur est donc de garder en memoire analogique la valeur instantanee du signal, jusqu'a ce que le convertisseur puisse lui faire correspondre une valeur numerique donnee;N codee ,qui sera alors transmise par la suite a l' unite de traitement.

#### 1.5.1 - L'echantillonnage et le blocage :

##### a - L'echantillonnage :

C'est la premiere operation a effectuer lors d'une conversion numerique analogique comme precise ci-dessus , le role de l'echantillonneur est donc essentiellement de prelever a un instant donne t ,un echantillon donne du signal ,cette operation se repetant periodiquement ,cela a pour effet de transformer une fonction continue en une fonction discontinue par simple decoupage de la premiere .  
Toutefois l'operation d'echantillonnage exige la satisfaction de deux conditions ,qui introduisent deux parametres importants pour le choix de l'interval de temps separant deux echantillons successifs appelee "periode d'echantillonnage " .

Le premier etant le temps de conversion  $T_c$  du convertisseur qui doit rigoureusement correspondre au temps  $T_e$  separant deux echantillons .

Si la periode  $T_e$  d'echantillonnage est inferieur au  $T_c$  la conversion est alors erronee .D'autre part le signal de la parole presente une forme quelconque d'evolution dans le temps ,car compose d'un nombre plus ou moins grand de composantes sinusoidales .

le 2eme parametre important est donc la representation spectrale du signal ,dans ce cas la periode d'echantillonnage  $T_e$  est choisie de maniere a ce qu'elle puisse satisfaire le Theoreme de SHANNON (Voir chapitre II),de telle sorte que la frequence d'echantillonnage  $F_e = 1/T_e$  verifie la relation :

$$F_e \geq 2 F_{max}$$

b - Bloquage :

Le bloquage de l'échantillon consiste à le maintenir stable dans une mémoire pendant toute la durée de conversion  $T_c$  du C.A.N, autrement dit la tension à l'entrée du C.A.N doit rester constante pendant toute la durée de conversion ce qui conduit alors à l'utilisation d'un circuit de bloquage faisant suite à l'échantillonneur sinon la conversion est erronée.

Le circuit assurant la fonction d'échantillonnage et qui est associé au C.A.N (ADC 0804) est le LF 398

c - Etude de l'échantillonneur bloqueur LF 398 :

i - Schema interne (Voir fig n°7) :

Le LF 398 comprend :

\* Deux amplis A1 et A2 montés en suiveur dont les caractéristiques sont bien connues (impédance d'entrée très élevée et souvent infinie, et le gain a pour valeur 1).

\* A1 sert à découpler l'entrée analogique de la capacité C afin d'éviter les oscillations de la tension de charge pendant la période d'échantillonnage.

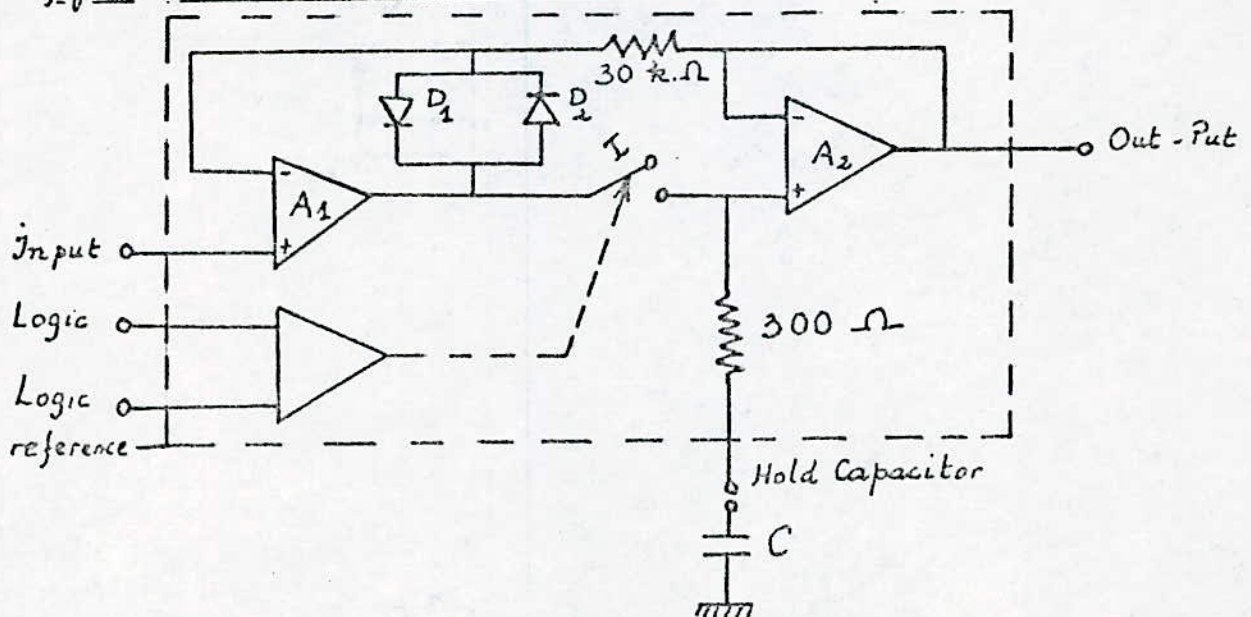
\* A2 sert à isoler la capacité C de la sortie afin de pouvoir éviter la décharge du condensateur pendant la période de bloquage (C voit une impédance très grande).

\* Deux diodes D1 et D2 servant à éviter la saturation de A1 en fixant le gain de celui-ci.

\* Un interrupteur I et sa logique de commande assurant ainsi l'échantillonnage.

\* Une résistance R1 limitant le courant qui passe entre A1 et A2 pendant la période de bloquage, la tension aux bornes de R1 est égale à la différence entre la tension d'entrée et celle de sortie.

fig n°7 \* SCHEMA FONCTIONNEL INTERNE DU LF 398 \*





ii - Fonctionnement en échantillonneur :

Sous l'action d'un signal de commande (niveau logique 1) l'interrupteur I se ferme ,mettant ainsi en liaison la sortie de A1 avec le condensateur ,celui-ci se charge grace au courant de sortie de l'ampli A1 tout en suivant l'évolution du signal d'entrée .

iii - Fonctionnement en bloqueur :

Des qu'un niveau "0" logique arrive l'interrupteur I s'ouvre et permet ainsi l'isolement de la capacité C du signal d'entrée .  
La tension de charge de C ne décroît pas car celle ci se trouve entre l'ouverture de I et l'impédance très élevée de A2 ,la valeur est donc maintenue ou bloquée .

NB : (Pour les caractéristiques techniques de ce composant se référer à l'annexe ) .

d - La conversion analogique numérique :

Pour notre réalisation le composant qui assure cette tâche est l'ADC 0804 ,pour cela nous avons jugé utile de ~~lais~~ réserver une partie à son fonctionnement .

i - Schema interne :

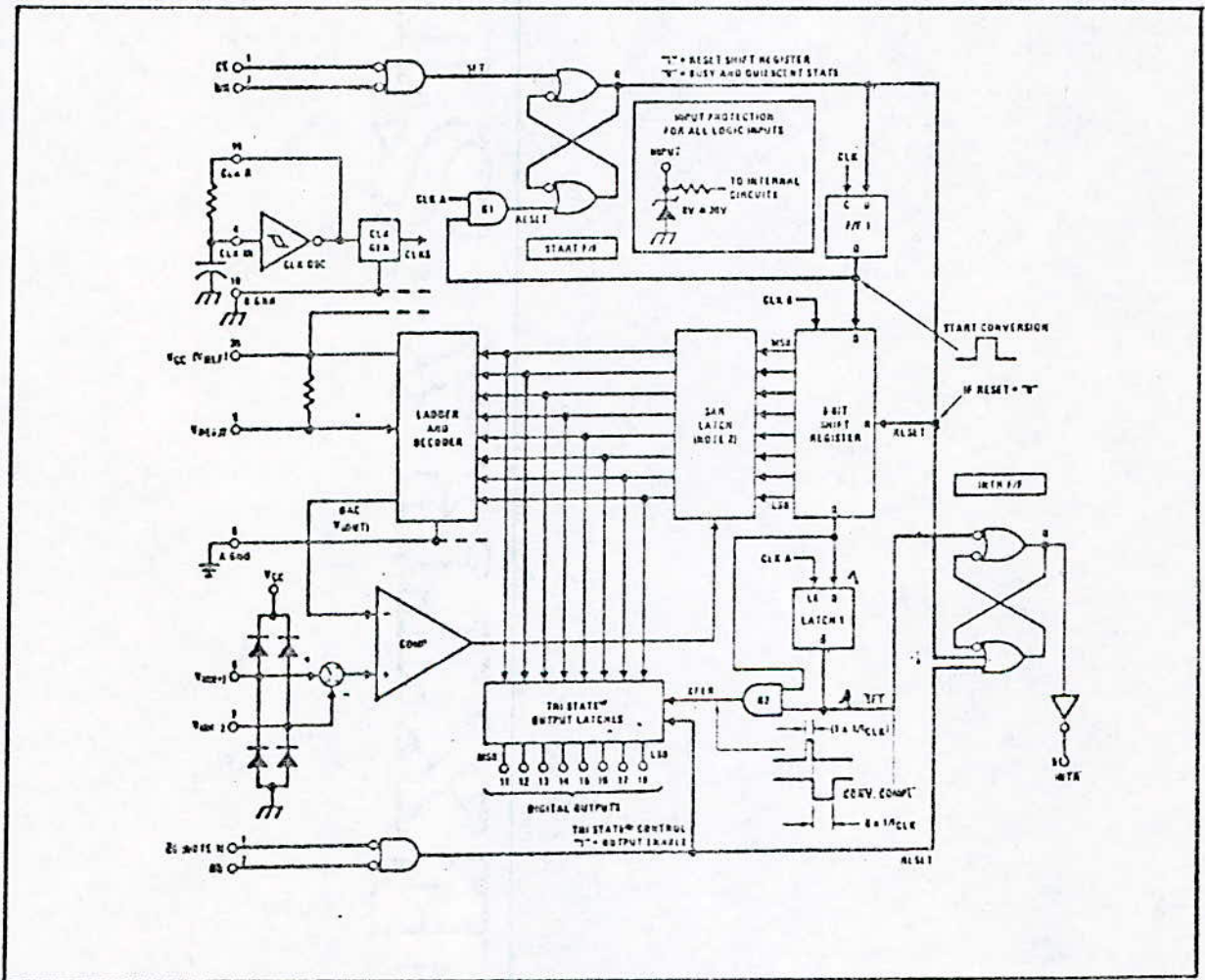
Se reporter à la fig n°: 8.  
On trouvera en annexe une notice détaillée donnant les principales caractéristiques du C.A.N ainsi que son brochage .

ii - Fonctionnement :

Pour étudier le fonctionnement de ce composant il faut présenter les trois commandes suivantes :

- 
- \* CS : Signal actif à l'état bas ,valide les deux signaux de lecture et d'écriture RD et WR et sélectionnant ainsi le C.A.N .
- 
- \* WR : Le passage de l'état ~~bas~~ <sup>haut</sup> à l'état ~~haut~~ <sup>bas</sup> de ce signal permet le lancement de la conversion .
- 
- \* RD : La présence d'un état bas sur cette entrée permet la lecture de l'information binaire ,disponible sur le bus de données du C.A.N en sortie .

En appliquant une tension  $V_{in}$  à l'entrée du C.A.N et des l'arrivée d'un niveau bas sur CS ~~et~~ WR entame la mise à



Schema Interne de l'ADC 0804

"1" de la sortie de la bascule RS START F/F ce qui implique alors : l'initialisation du registre a decalage "8 bits shift register".

La mise a "1" de la sortie INTR indiquant ainsi qu'une conversion est en cours d'execution.

L'horloge CLK de la bascule D'F/F1 transfere alors le "1" logique a la sortie Q de cette derniere .Ce qui a pour effet de mettre un zero au niveau de la bascule START F/F au top d'horloge CLKA eliminant ainsi une seconde initialisation du registre pendant la conversion.

La conversion a lieu ainsi :

A l'activation de l'horloge CLKB du "shift register" le "1" logique present a l'entree du registre est transmis au SAR LATCH (successive approximation register latch). Le C.N.A convertit cette valeur numerique en une grandeur analogique ( tension ) qui sera comparee a  $V_{in}$  , selon le resultat de cette comparaison le MSB du SAR sera lailsee a "1" ou a "0" .

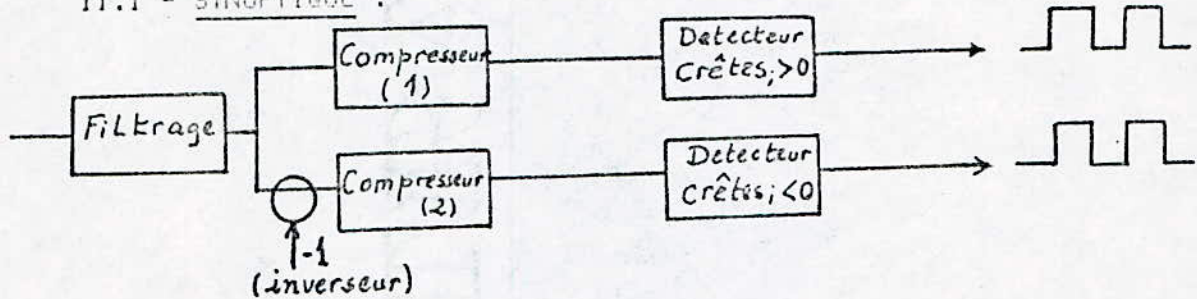
A la prochaine activation de l'horloge CLKB, il y'a decalage du "1" present dans le registre a decalage puis une nouvelle comparaison est lancee .Apres huit comparaisons "decalage successif du 1 sur les sorties du SAR" On obtient la valeur numerique N (suite de 1 ou de 0) correspondant a  $V_{in}$  .

Arrive en fin de conversion (sortie LSB du registre),le LSB est transmis a la bascule D'LATCH1 permet de mettre a zero la bascule INTRF/F ce qui a pour effet de mettre un zero sur INTR ,indiquant que celle ci est passee de l'etat haut a l'etat bas,et que la conversion est donc terminee .

Le signal XFER (sortie de la porte AND G2) est le resultat de la combinaison de la sortie D'LATCH1 et la sortie Q du shift register,permettant le stockage et le transfert de la conversion dans le registre "TRI STATE OUTPUT LATCHES " une mise a l'etat bas simultane de RD et CS permet la lecture de la valeur numerique N sur 8 bits disponibles maintenant dans le registre "OUTPUT LATCHES" une remise a zero (RAZ) de la bascule INTRF/F donc passage de INTR de l'etat bas a l'etat haut et une initialisation du registre a decalage ,ces parametres indiquant qu'un nouveau cycle de conversion est en cours d'execution .

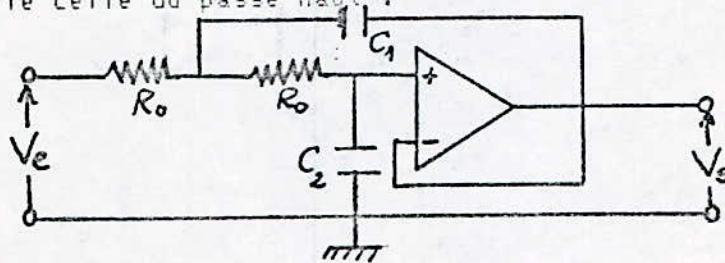
II - ETUDE ET REALISATION DU DETECTEUR DE MELODIE :

II.1 - SYNOPTIQUE :



II.2 - FILTRAGE :

C'est la structure adoptée dans la partie analyseur par canaux, le filtrage est donc constitué de deux filtres passe haut et un passe bas le tout dispose en cascade. Le calcul de ces filtres se fait comme précédemment. On part de la structure passe bas pour déterminer par symétrie celle du passe haut.



II.2.1 - Calcul du filtre passe bas d'ordre 2 (Butterworth) :

On a :  $\omega_0 = 2 F_0$

En premier lieu dans le cas d'un passe bas, on fixe la valeur de la résistance  $R_0$ , ceci permet d'avoir la capacité intermédiaire :

$$C_0 = \frac{1}{R_0 \omega_0}$$

Ceci permet de déterminer  $C_1$  et  $C_2$  par les expressions suivantes :

$$C_1 = m C_0$$

$$C_2 = q C_0$$

$m, q$  étant les coefficients précalculés de la structure Butterworth correspondant à l'ordre choisi. Ainsi dans le cas du second ordre :

$$m = 0.707 \quad \text{et} \quad q = 1.414$$

II.2.2 - Détermination du passe haut d'ordre 2 :

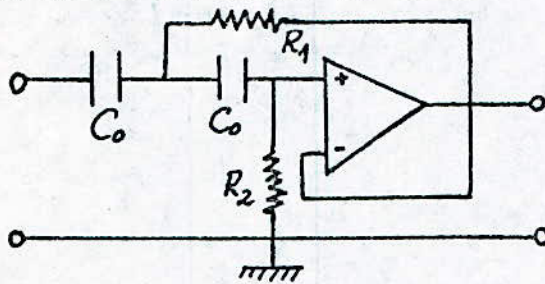
pour la détermination du filtre passe haut, c'est la valeur de  $C_0$  qui est fixée, et la valeur intermédiaire prise en compte est :

$$R_0 = \frac{1}{C_0 \omega_0}$$

Et pour déterminer les résistances  $R_1$  et  $R_2$  on divise respectivement  $R_0$  par les coefficients  $m$  et  $q$ .  
Donc les valeurs de  $R_1$  et  $R_2$  sont :

$$R_1 = \frac{R_0}{m} \qquad R_2 = \frac{R_0}{q}$$

La structure est donnée par le schéma suivant :



Cette méthode est utilisée pour déterminer les composants  $R, C$  du filtrage passe bande dans le détecteur de mélodie la structure du filtre passe bande est donnée par la figure ci-dessous.

*figure identique à celle décrite dans  
L'Etude d'un canal.*

*Se reporter à la figure précédente  
( § I.1.5 )*

Les valeurs de ces composants sont données au chapitre VI

11.3 - ETUDE DU BLOC COMPRESSEUR :

11.3.1 - Introduction :

Le signal filtre subit une compression de l'enveloppe des crêtes maximums, dans le cas où l'énergie du signal varie rapidement il n'est pas aisé de détecter les crêtes significatives. La chaîne de traitement suit deux branches distinctes mais identiques ; l'une pour la détection des crêtes positives et l'autre pour les crêtes négatives. Les deux compresseurs utilisés se caractérisent par un temps de réaction instantané à la montée et un temps de réaction relativement long à la descente. Cet effet est obtenu à l'aide d'une mémoire de crêtes, qui commande la compression. Cette valeur est mémorisée à peu près linéairement après un certain retard.

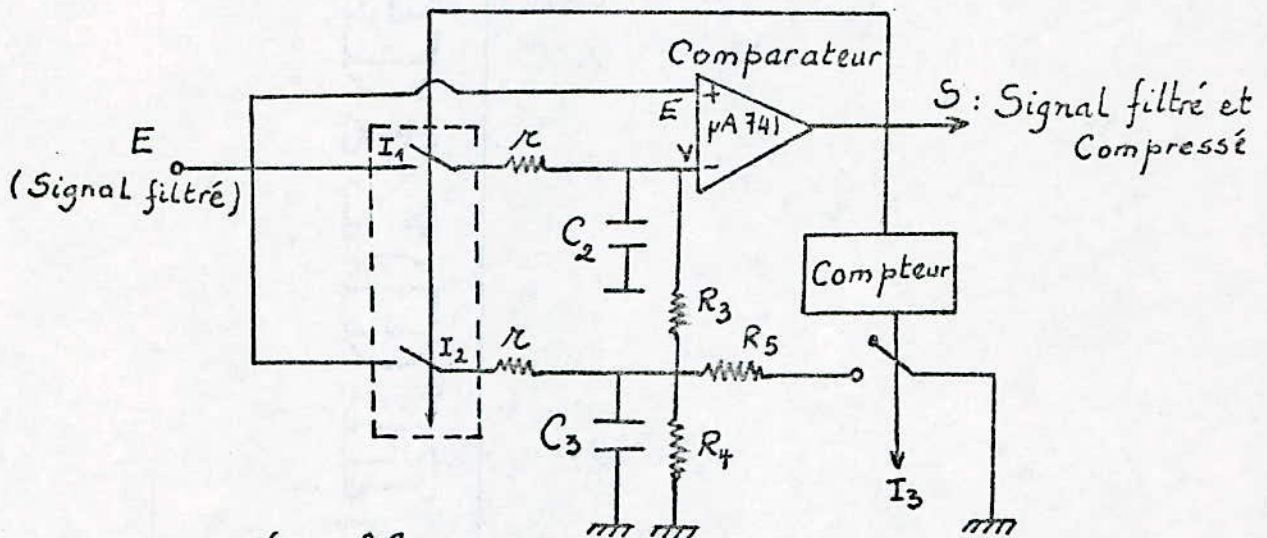


fig n° 9 : schema fonctionnel du Compresseur

11.3.2 - Etude du fonctionnement :

Le fonctionnement de ce bloc est le suivant :

- \* - Si la tension d'entrée  $E > V$ , la sortie du comparateur entraîne la fermeture des interrupteurs I1 et I2 de sorte que les capacités sont instantanément chargées. Soit  $W = V = E$ .
- \*\* - Si la tension  $E < V$ , la sortie du comparateur tombe à zéro nous avons alors une fonction de perte de mémoire pendant ce temps les interrupteurs I1 et I2 sont ouverts et la fuite du circuit de mémoire de crête se fait suivant un profil minimisant la distorsion du signal compressé ..

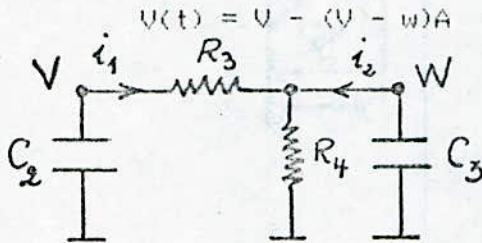
$$-V(t) = \frac{(V - W)t}{R_3 C_2} + Cte$$

Au début de la décharge on a : à  $t = 0$  on a  $V(0) = V$

d'où :

$$-V(t) = \frac{(V - W)T}{R_3 C_2} - V$$

D'où alors en posant  $A = T / R_3 C_2$  on a :



D'autre part :

$$V = R_3 i_1 + R_4 (i_1 + i_2) = (R_3 + R_4) i_1 + R_4 i_2$$

or  $i_1 = \frac{V - W}{R_3}$  et  $i_2 = -C_3 \frac{dW}{dt}$

$\Rightarrow V = \left( \frac{R_3 + R_4}{R_3} \right) (V - W) - R_4 C_3 (dW/dt)$ , en arrangeant cette equation on a :

$$\left[ \frac{V}{R_4 C_3} - \left( \frac{R_3 + R_4}{R_3} \right) \left( \frac{V - W}{R_4 C_3} \right) \right] dt = -dW$$

d'où  $W(t) = W - \left[ V - \left( \frac{R_3 + R_4}{R_3} \right) (V - W) \right] \frac{T}{R_4 C_3}$  resultat obtenu, en prenant les mêmes conditions initiales que précédemment

Il vient alors :  $W(t) = W - (W(R_3 + R_4) - VR_4) \frac{T}{R_3 R_4 C_3}$

Posons :  $\begin{cases} B = (R_3 + R_4) \\ C = R_4 \\ X = \frac{T}{R_3 R_4 C_3} \end{cases} \Rightarrow \underline{W(t) = W - (BW - CV)X}$

Dans le cas d'une chute de l'énergie du signal si la tension E reste inférieure à V au delà de 12 ms le décompteur entraîne la fermeture I3 et les constantes B, C, X deviennent alors :

$$\begin{cases} B = R_3 + (R_4 \parallel R_5) \\ C = R_4 \parallel R_5 \\ X = \frac{T}{R_3 \cdot (R_4 \parallel R_5) \cdot C_3} \end{cases}$$

Ce qui provoque une perte de mémoire plus rapide afin de mieux suivre les crêtes .

\*\*\* - Le compteur genere une impulsion toute les 12 ms, celle-ci commande la fermeture de l'interrupteur I3 et permet ainsi la décharge rapide du circuit mémoire de crete

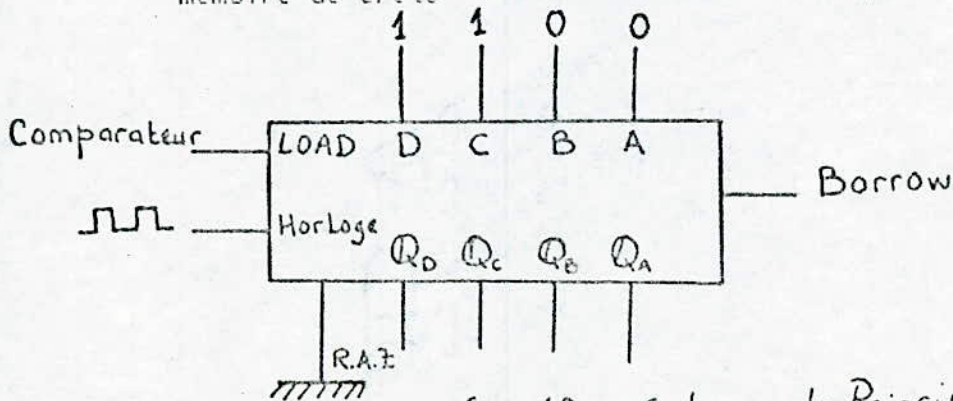


fig 10 : Schema de Principe du décompteur

Le compteur precedent est un compteur hexadecimal de tupe programmable (SN 74 LS 193) la sortie du comparateur attaque l'entree LOAD du compteur. Ainsi le debut du comptage a lieu lorsque LOAD est au niveau zero ,les entrees D,C,B,A sont fixees en logique T.T.L a 1100 (12 ms) La periode de l'horloge a ete fixee a 1 ms celle-ci est genere par un timer (NE 555 monte en astable) de periode

$$T = ( RA + 2 RB ) C \text{ LOG}(2)$$

Le decomptage demarre a partir du moment ou l'entree LOAD passe de "1" a "0" c'est a dire des que la tension E devient inferieure a  $V$  .

La sortie BORROW se met a l'etat "1" des le debut du decomptage .Au moment ou les sorties QD,QC,QB,QA atteignent la valeur affichees sur les entrees D,C,B,A la sortie BORROW se remet a zero ,formant ainsi un creneau de duree egale a celle d'une impulsion d'horloge. On place un inverseur sur la sortie BORROW du decompteur de sorte qu'elle reste a l'etat zero pendant 12 ms et juste au moment ou les 12 ms sont comptees elle se remet a l'etat "1".Ce passage de l'etat "0" a l'etat "1" toute les 12 ms de la sortie BORROW qui est reliee d'autre part a l'interrupteur I3 va permettre l'ouverture de ce dernier.

La fermeture →



II.4 - ETUDE DU BLOC DETECTEUR DE CRETES :

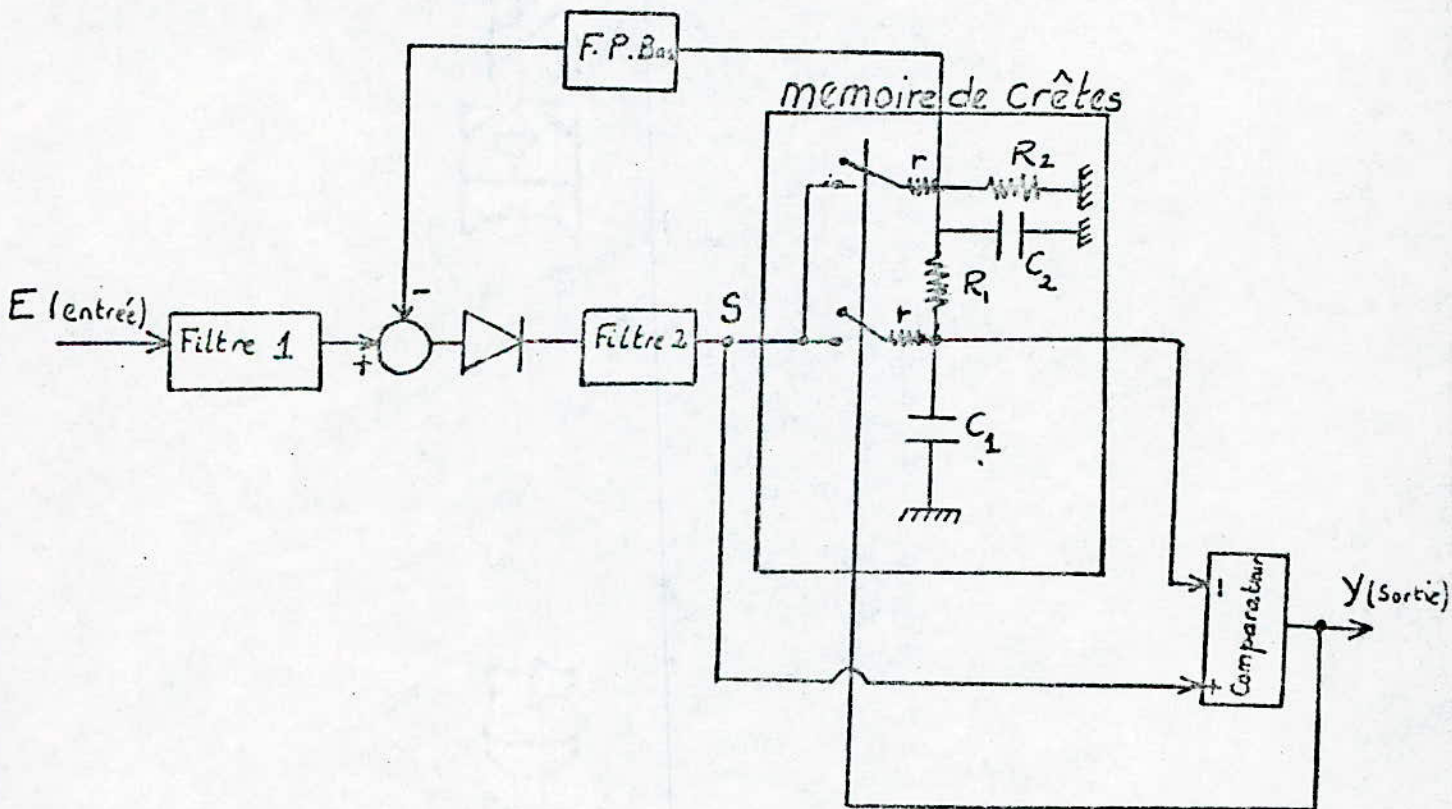
II.4.1 - Introduction :

Le schéma fonctionnel du détecteur de mélodie est représenté par la figure ci-dessous .

*( figure identique à celle décrite au paragraphe II.1)  
(s'y référer).*

Le signal est au préalable filtré afin de faciliter la détection des crêtes significatives .Ce préfiltrage comporte un passe haut de fréquence de coupure  $F_{c1}=70$  Hz et un passe bas de fréquence de coupure  $F_{c2}=400$  Hz. La chaîne de traitement suit deux branches distinctes , l'une pour la détection des crêtes positives , l'autre pour celles négatives .

fig:11\* SYNOPTIQUE DU DETECTEUR DE CRETES ( VERSION NUMERIQUE ) \*



11.4.2 - Fonctionnement du detecteur de cretes :

Le signal filtre subit une compression des cretes maximum, car la detection des cretes significatives n'est pas chose facile, dans le cas d'une evolution rapide de l'energie du signal. Les compresseurs se caracterisent par un temps de reaction instantane a la montee, et un temps de reaction relativement lent a la descente. Cet effet est obtenu a l'aide d'une memoire de cretes qui commande la compression. Le fonctionnement de la memoire des cretes est donnee en fin de paragraphe.

Le travail du detecteur de cretes consiste essentiellement en un repereage des cretes qui marquent la periodicite de  $F_0$ , en emettant periodiquement des paquets d'impulsions en nombre variable, mais dont la derniere impulsion coincide avec le point extremum de la crete positive et inversement avec les cretes negatives. Par consequent les deux detecteurs de cretes delivrent en parallele des suites de creneaux. La valeur de  $T_0$  du fondamentale est defini comme l'interval de temps le plus court separant deux creneaux et satisfaisant aux deux conditions suivantes :

\* La polarite la plus favorable ait ete reperee (Celle dont le nombre de cretes detectees est superieur au second) .

\*\* Au moins une crete de polarite <sup>contraire</sup> fait ete detectee

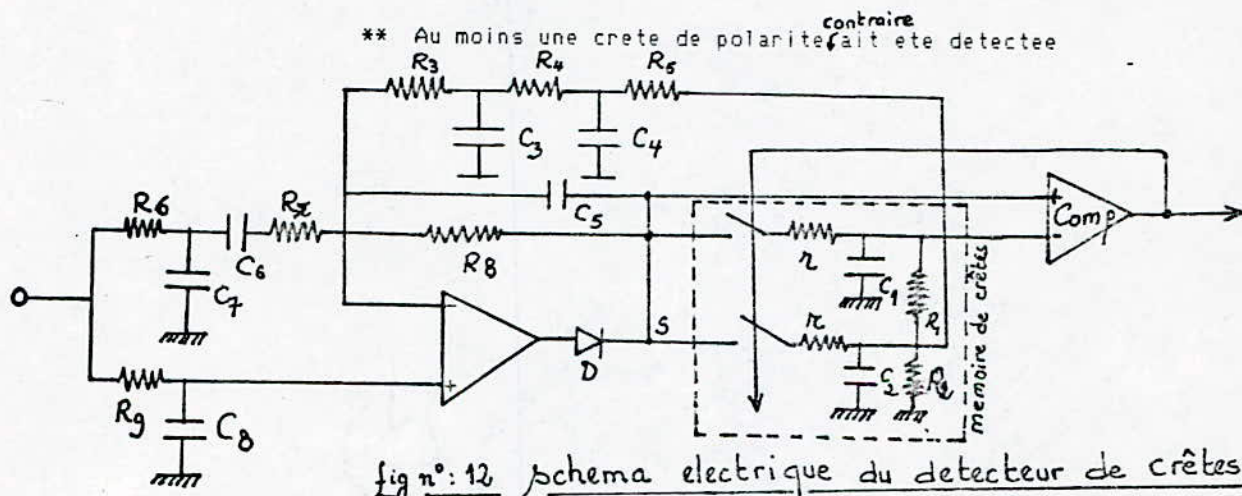


fig n°: 12 schema electrique du detecteur de cretes

- Fonctionnement de la memoire de cretes :

En schematisant le fonctionnement de la memoire de crete d'un detecteur donne, en correlation avec ces autres etages, on obtiendrait la figure ci-dessous :

\* Choisissons le cas positif par exemple :

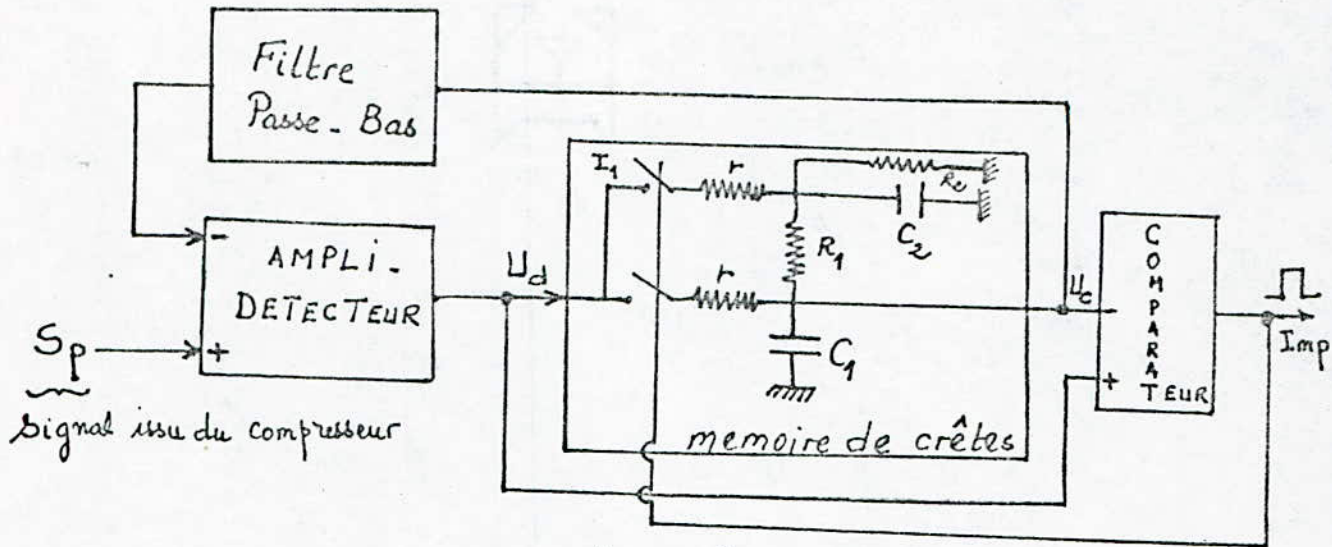


fig: 13 Schema fonctionnel.

Le signal  $S_p$ , attaque l'entrée positive de l'ampli, détecteur simple alternance dont le signal de sortie  $U_d$ , est connecté à un circuit mémoire de crêtes similaire à celui du compresseur.

La mémoire de crêtes délivre un signal d'enveloppe de crêtes  $U_c$ .

Les signaux  $U_d$  et  $U_c$  sont connectés aux entrées "+" et "-" du comparateur.

Lorsque la sortie (IMP) du comparateur est positive donc  $U_d >$  ou  $= U_c$ , les interrupteurs  $I_1$  et  $I_2$  se ferment les condensateurs  $C_1$  et  $C_2$  se chargent à travers  $(r)$ .

Des que  $U_d <$  ou  $= U_c$  les condensateurs  $C_1$  et  $C_2$  se déchargent à travers respectivement  $R_1$  et  $R_2$ .

L'équation de décharge est du 2ème ordre et est donnée par:

$$\frac{d^2 U_c}{dt^2} (\theta_1 \cdot \theta_2) + \frac{dU_c}{dt} (\theta_1 + \theta_2 + \theta_3) + U_c = 0$$

La solution d'une telle équation est de la forme:

$$U_c(t) = K_1 e^{x_1 t} + K_2 e^{x_2 t}$$

$x_1$  et  $x_2$  étant les racines réelles de l'équation

$$\theta_1 \theta_2 x^2 + (\theta_1 + \theta_2 + \theta_3) x + 1 = 0$$

$$\text{avec } \theta_1 = R_1 C_1 \quad \theta_2 = R_2 C_2 \quad \theta_3 = R_2 C_1$$

Les conditions initiales à l'instant  $t=0$  de l'ouverture de  $I_1$  et  $I_2$  sont  $U_{c1}(0) = U_{c2}(0) = U_{cmax}$  et  $U'_c(0) = 0$

Elles permettent de déterminer:

$$U_c(t) = \frac{x_2}{x_2 - x_1} U_{cmax} e^{x_1 t} + \frac{x_1}{x_1 - x_2} U_{cmax} e^{x_2 t}$$

La decharge resultante est donc la somme de deux decharges exponentielles .

Le signal  $U_c$  est retarde dans un filtre passe bas et est applique a l'entree "-" de l'ampli detecteur dans le but d'affaiblir le signal  $U_d$  apres la detection d'une crete significative. Cet affaiblissement va en diminuant au rythme de la decroissance de  $U_c$  jusqu'a la detection de la prochaine crete significative. Le filtre permet aussi l'affaiblissement des composantes H.F introduites par la fin de la decharge des capacites .

Après avoir étudié en détails chaque composante de ce bloc, nous donnons ici un schéma global du détecteur de mélodie (Fo). (Se reporter au Chap. VI) fig n° 14

Les tests sur ce bloc se sont faits au fur et à mesure de l'étude en conséquence de quoi on trouvera des traces de courbes de réponses de chaque étage formant ce bloc dans le chapitre VI. Les résultats y sont étayés en fonction des contraintes pratiques et des difficultés rencontrées.

Les tests se sont faits sur une chaîne de traitement nous avons choisi celle qui détecte les crêtes positives, la seconde étant identique à l'inversion pres.

### III - STOCKAGE DES DONNEES ET INTERFACAGE :

#### III.1 - INTRODUCTION :

La complexité du signal de la parole a fait en sorte que l'utilisation d'une carte de traitement est apparue nécessaire, tel qu'il a été posé, le problème résidait au niveau de la numérisation et par suite le stockage des données disponibles sur le convertisseur. Comme il a été signalé auparavant, le C.A.N a un mode de sélection (CS), un mode de début de conversion ( $\overline{WR}$ ) et enfin un mode d'exploitation des données ( $\overline{RD}$ ) ; la nécessité de présence de ces trois signaux nous ont aiguillé vers l'accès au mode d'interfacage assuré par le kit MOTOROLA 6802.

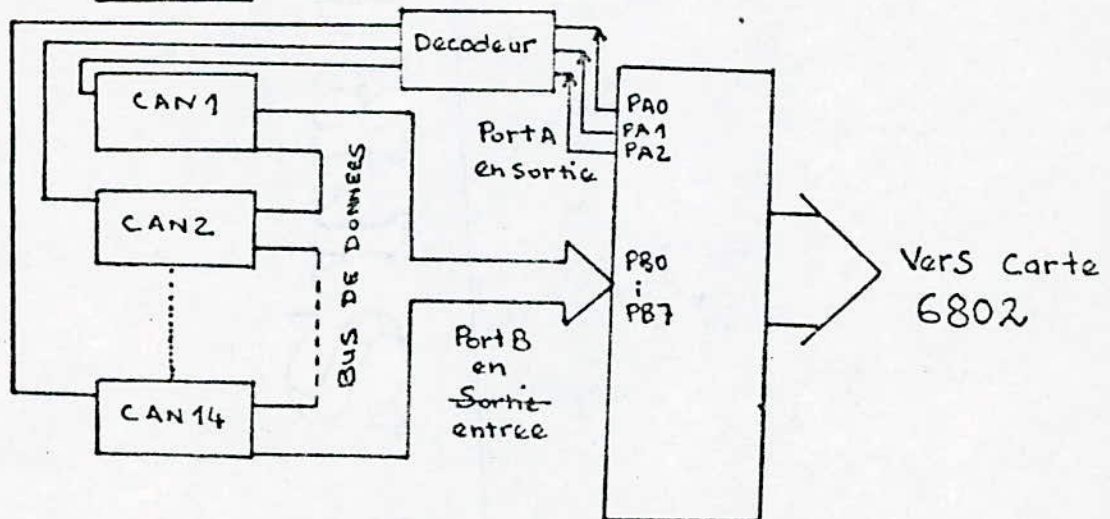
#### III.2 - LE PIA 6821 :

##### III.2.1 - Présentation :

C'est un composant de type boîtier d'entrée sortie qui est appelé encore adaptateur d'interface parallèle (PIA 6821 de MOTOROLA).

Le PIA a une interface bus style MOTOROLA, deux ports d'entrée sortie à huit bits, avec deux lignes d'asservissement chacun et deux lignes de demandes d'interruption. (Le schéma fonctionnel ainsi que le brochage de ce composant seront donnés en annexe).

##### III.2.2 - Synoptique :



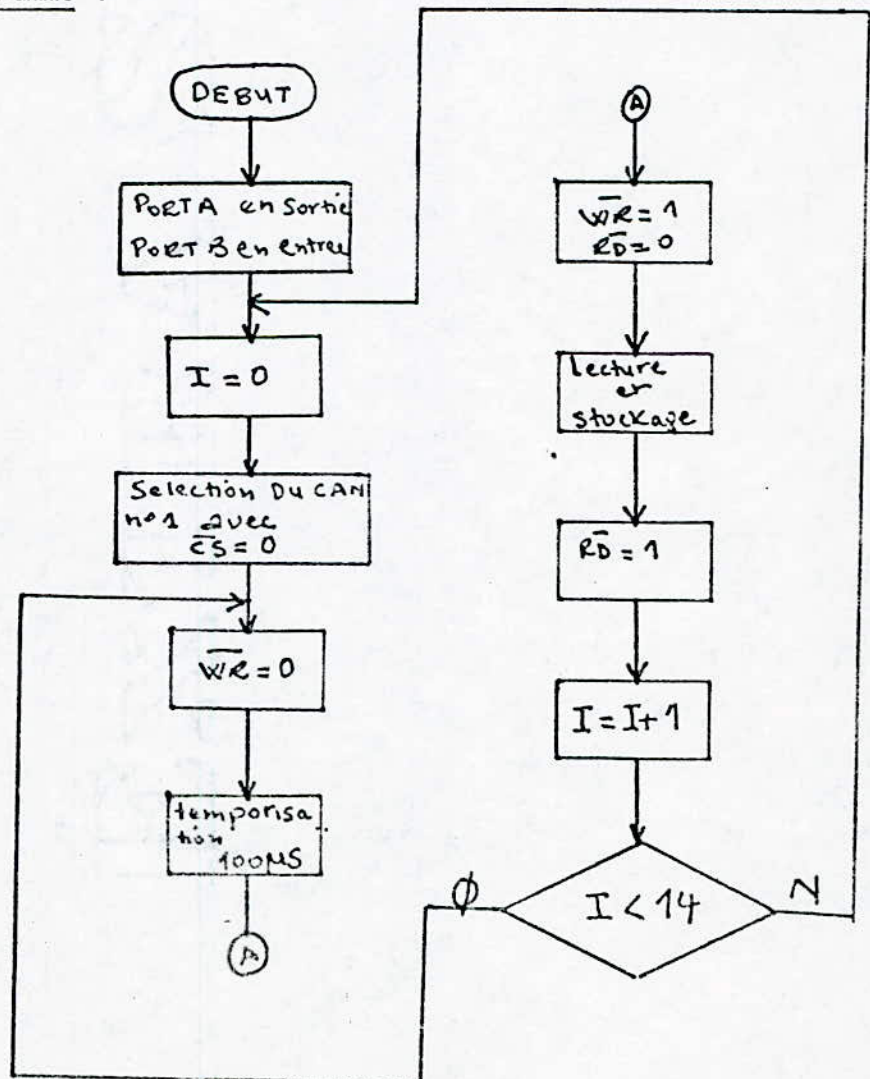
a - description :

le vocodeur que nous sommes entrain d'étudier se compose de 14 canaux ,dont rappelons le chaque canal comprend les elements suivants :

- \* Un filtre passe bande
- \* Redresseur simple alternance
- \* Un echantillonneur bloqueur
- \* Un convertisseur

Le but de cet interfacage est donc de selectionner une fois chacun des convertisseurs et stocker leurs donnees en RAM .Cette operation est possible en utilisant un decodeur qui lui meme commande par le MPU 6802 selection -ne un convertisseur parmi les 14 ,pour cela nous avons pense utilise le 74 LS 138 qui est un decodeur un parmi huit .Afin de mettre en execution tout ce qu'on vient de de dire ,nous avons etabli un programme qui se charge de selectionner les convertisseurs et de transferer leurs donnees en RAM .

b - Organigramme :



c - Programme en assembleur :

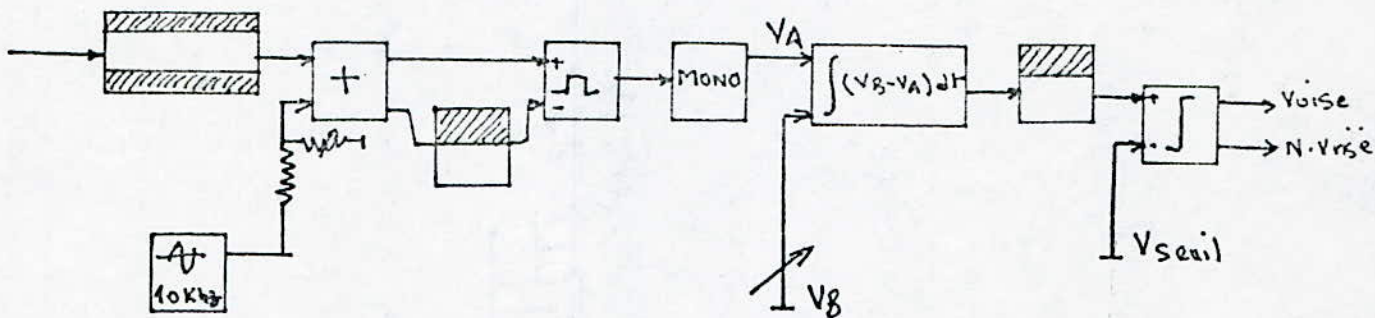
	CLR E481	7F E4 81
	LDA #FF	86 FF
	STAA E480	97 E4 80
	CLR E483	7F E4 83
	CLRA	4F
	STAA E482	97 E4 82
	LDA #00000100	86 04
RS	CLR E480	4F E4 80
LP	LDA #00110100	86 34
	STAA E481	97 E4 81
ST	LDA #10	86 10
	DECA	4A
	BNE ST	26 FC
	LDA #00111100	86 3C
	STAA E481	97 E4 81
	LDA #00110100	86 34
	STAA E483	97 E4 83
	LDX # 0000	CE 00
	LDA E482	86 E4 82
	STAA 0,X	97 00
	INX	08
	LDA #00111100	86 3C
	STAA E483	97 E4 83
	INC E480	7C E4 80
	LDA E480	96 E4 80
	CMPE #0E	81 0E
	BEQ RS	27 CE

#### IV - LE DETECTEUR DE VOISEMENT :

##### IV.1 - INTRODUCTION :

Comme il a été signalé dans l'introduction, le détecteur de voisement, représente un organe très important dans un vocodeur à canaux. Pour cette raison, malgré les difficultés et le manque de documentation concernant ce bloc, nous décrivons le fonctionnement de la version analogique de ce dernier.

##### IV.2 - SYNOPTIQUE :



##### IV.3 - FONCTIONNEMENT DU DETECTEUR DE VOISEMENT :

Comme le montre le synoptique ci-dessus, le signal subit d'abord un filtrage dont la fonction de transfert est schématisée sur la figure 2.

On lui additionne alors, à un niveau réglable, une fréquence pure de 10 KHz. L'ajustage de ce niveau détermine le seuil d'énergie en dessous duquel tous les sons sont reconnus comme non voisés. Nous trouvons ensuite un dispositif de détection des passages par zéro, ou plus précisément des passages par la valeur moyenne obtenue dans un filtre passa bas dont la fréquence de coupure est de 250 Hz. Ce bloc déclenche ensuite un monostable qui délivre des impulsions calibrées VA, de 0,8 ms, sur les passages par la moyenne d'un même sens.

Les impulsions VA attaquent un intégrateur qui délivre une tension :

$$V_S = k \int (V_B - V_A) \cdot dt$$

VB étant une tension réglable.

En l'absence d'impulsions VA, VS sera positive. VB est ajustée afin que VS devienne négative pour tout son non voisé, c'est à dire quand les passages par la moyenne sont nombreux.

VS subit un filtrage passe bas dans un circuit RC dont la fréquence de coupure est de 35 Hz, ceci a pour but de supprimer les rebondissements indésirables. Enfin un comparateur prend la décision voisement ou non selon le signe de VS.



IV.4 - DECISION :

La decision sur le voisement ou non du signal de la parole peut s'effectuer a la sortie de l'integrateur. Cependant ,cette sortie evolue en forme de dents de scie de longueur variable ,ce qui signifie qu'il peut exister des oscillations autour d'une valeur de seuil . Par consequent ,il est necessaire de filtrer le signal afin de proceder a un lissage de ce dernier ,et c'est la sortie de ce filtre qui est comparee a un seuil en dessous duquel le son est reconnu comme non voise .

CHAPITRE VI

CALCULS, TESTS ET REALISATION

I - CALCULS :

I.1 - COMPOSANTS CONSTITUANTS LE BAND ANALYSEUR :

CANAL 1 :

F1 = 250 Hz    F2 = 450 Hz

FILTRE PASSE HAUT

R1 = 3.3 K    R2 = 2.7 K    R3 = 8.2 K    R4 = 1.2 K

FILTRE PASSE BAS

C1 = 27 nF    C2 = 47 nF

CANAL 2 :

F1 = 450 Hz    F2 = 650 Hz

FILTRE PASSE HAUT

R1 = 1.8 K    R2 = 1.8 K    R3 = 4.7 K    R4 = 0.68 K

FILTRE PASSE BAS

C1 = 18 nF    C2 = 33 nF

CANAL 3 :

F1 = 650 Hz    F2 = 890 Hz

FILTRE PASSE HAUT

R1 = 1.2 K    R2 = 1 K    R3 = 3.3 K    R4 = 4.7 K

FILTRE PASSE BAS

C1 = 12 nF    C2 = 27 nF

CANAL 4 :

F1 = 890 Hz    F2 = 1050 Hz

FILTRE PASSE HAUT

R1 = 0.82 K    R2 = 0.82 K    R3 = 2.2 K    R4 = 0.33 K

FILTRE PASSE BAS

C1 = 10 nF    C2 = 22 nF

CANAL 5 :

F1 = 1050 Hz F2 = 1300 Hz

FILTRE PASSE HAUT

R1 = 0.68 K R2 = 0.68 K R3 = 1.8 K R4 = 0.27 K

FILTRE PASSE BAS

C1 = 8.2 nF C2 = 18 nF

CANAL 6 :

F1 = 1300 Hz F2 = 1600 Hz

FILTRE PASSE HAUT

R1 = 0.62 K R2 = 0.56 K R3 = 1.5 K R4 = 0.22 K

FILTRE PASSE BAS

C1 = 6.8 nF C2 = 12 nF

CANAL 7

F1 = 1600 Hz F2 = 1900 Hz

FILTRE PASSE HAUT

R1 = 0.47 K R2 = 0.39 K R3 = 1.2 K R4 = 0.18 K

FILTRE PASSE BAS

C1 = 5.6 nF C2 = 12 nF

CANAL 8 :

F1 = 1900 Hz F2 = 2200 Hz

FILTRE PASSE HAUT

R1 = 0.47 K R2 = 0.39 K R3 = 1 K R4 = 0.15 K

FILTRE PASSE BAS

C1 = 5.6 nF C2 = 10 nF

CANAL 9 :

F1 = 2200 Hz F2 = 2500 Hz

FILTRE PASSE HAUT

R1 = 0.39 K R2 = 0.33 K R3 = 0.82 K R4 = 0.12 K

FILTRE PASSE BAS

C1 = 4.7 nF C2 = 8.2 nF

CANAL 10 :

F1 = 2500 Hz F2 = 2800 Hz

FILTRE PASSE HAUT

R1 = 0.33 K R2 = 0.27 K R3 = 0.82 K R4 = 0.12 K

FILTRE PASSE BAS

C1 = 4.7 nF C2 = 8.2 nF

CANAL 11 :

F1 = 2800 Hz F2 = 3100 Hz

FILTRE PASSE HAUT

R1 = 0.27 K R2 = 0.22 K R3 = 0.68 K R4 = 0.10 K

FILTRE PASSE BAS

C1 = 3.7 nF C2 = 6.8 nF

CANAL 12 :

F1 = 3100 Hz F2 = 3500 Hz

FILTRE PASSE HAUT

R1 = 0.27 K R2 = 0.22 K R3 = 0.68 K R4 = 0.082 K

FILTRE PASSE BAS

C1 = 3.3 nF C2 = 6.8 nF

CANAL 13 :

F1 = 3500 Hz F2 = 3900 Hz

FILTRE PASSE HAUT

R1 = 0.22 K R2 = 0.18 K R3 = 0.56 K R4 = 0.082 K

FILTRE PASSE BAS

C1 = 2.7 nF C2 = 5.6 nF

CANAL 14 :

F1 = 3900 Hz F2 = 4300 Hz

FILTRE PASSE HAUT

R1 = 0.22 K   R2 = 0.18 K   R3 = 0.47 K   R4 = 0.068 K

FILTRE PASSE BAS

C1 = 2.7 nF   C2 = 5.6 nF

On notera que chaque canal comportera aussi 4 capacites de 220 nF et 2 resistances de 10 K et ce respectivement pour les filtres passe haut et les filtres passe bas .

L' etape de numerisation etant faite par un A D C 0804 et un echantillonneur bloqueur de reference L F 398 .

1.2 - COMPOSANTS CONSTITUANTS LE DETECTEUR DE MELODIE :

1.2.1 - FILTRAGE :

Ce bloc comprend les composants suivants :

R1 = 15 K R1' = 10 K R2 = 22 K R2' = 3.9 K  
C1 = 22 nF C1' = 47 nF C0 = 220 nF R0 = 12 K

LM 324

1.2.2 - COMPRESSEUR :

r = 10 K R3 = 36 K R4 = 47 K R5 = 10 K  
RA = 1 K RB = 6.8 K RL1 = RL2 = 0.18 K  
C2 = C3 = 220 nF C = 100 nF

CD 4016 , SN 74193 , SN 7404 , NE 555 , LM 741 , DIN 4001

1.2.3 - DETECTEUR DE CRETES :

R7 = 12 K R8 = 10 K R9 = 47 K R10 = 30 K R14 = 10 K  
R15 = 10 K R16 = 22 K R12 = 33 K R13 = 6.8 K r = 10 K  
RCH1 = 10 K

C4 = 33 nF C5 = 39 nF C6 = 220 nF C7 = 4.7 nF  
C8 = 220nF C9 = 220 nF C10 = 47 nF C11 = 47 nF

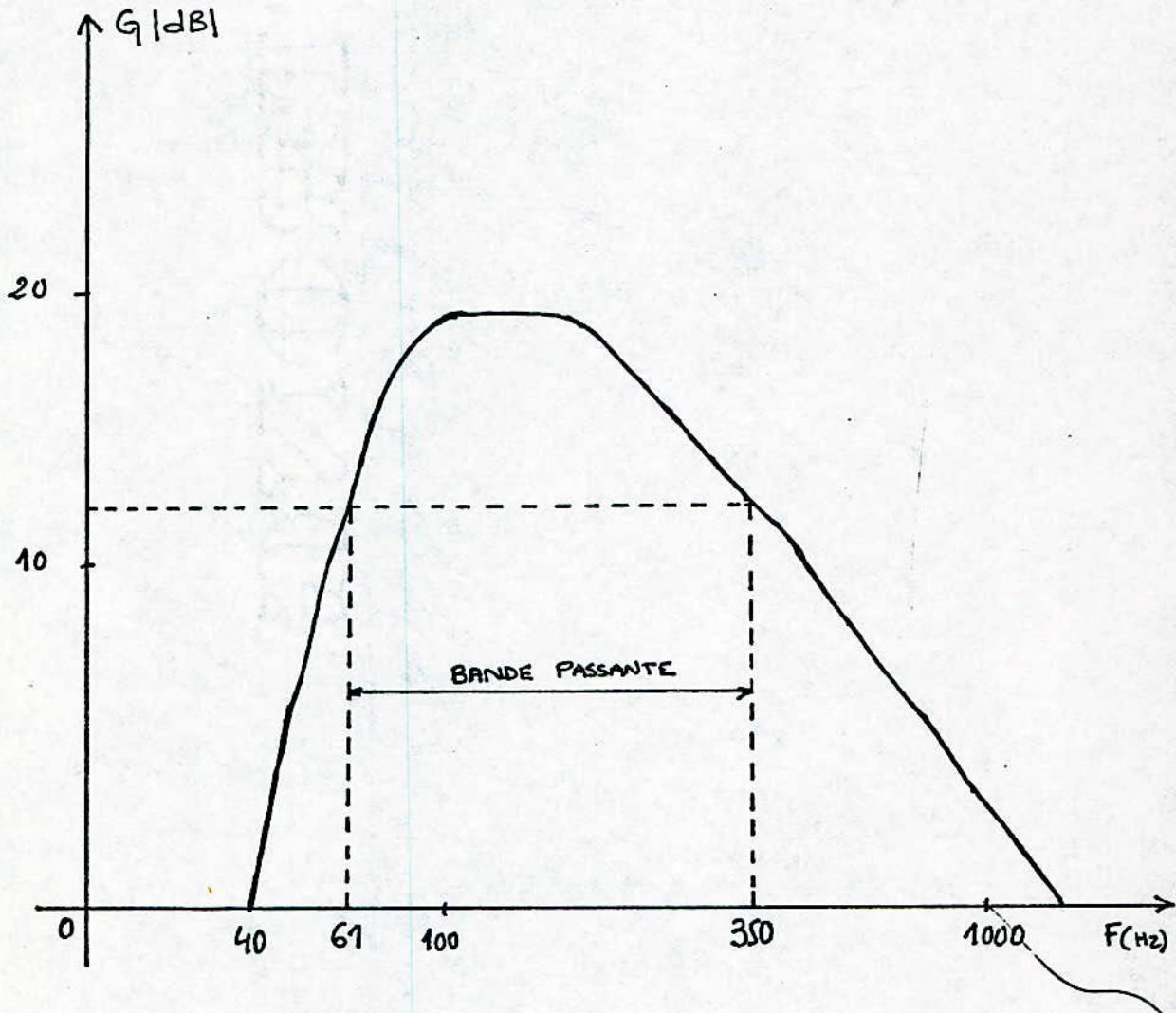
2 LM 741 , CD 4016

II - TESTS :

Les tests qui vont être énumérés dans ce chapitre, ont été établis sur le détecteur de mélodie, dont l'entrée de ce bloc a été générée par un G.B.F.

II.1 - FILTRAGE :

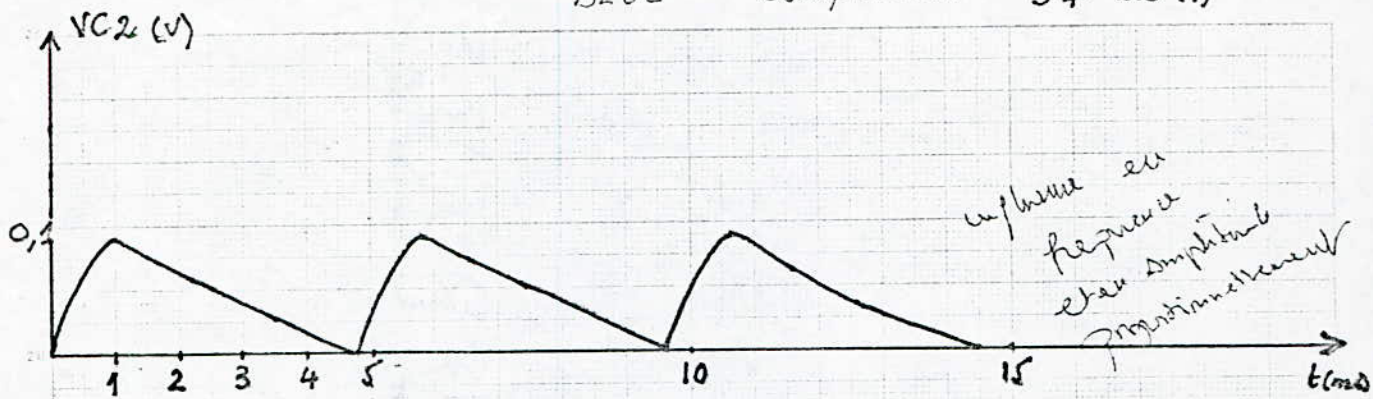
F (Hz)	Vs (v)	G (dB)
50	0.2	7.44
60	0.38	13.006
70	0.44	14.28
80	0.72	18.56
90	0.8	19.48
100	0.82	19.687
120	0.82	19.687
140	0.82	19.687
160	0.72	19.025
180	0.64	17.54
200	0.6	16.18
250	0.48	15.04
300	0.42	13.88
350	0.36	12.54
400	0.3	10.96
420	0.3	10.97
440	0.3	10.97
460	0.28	10.35
480	0.26	9.70
500	0.25	9.20
600	0.2	7.44
1100	0.085	0



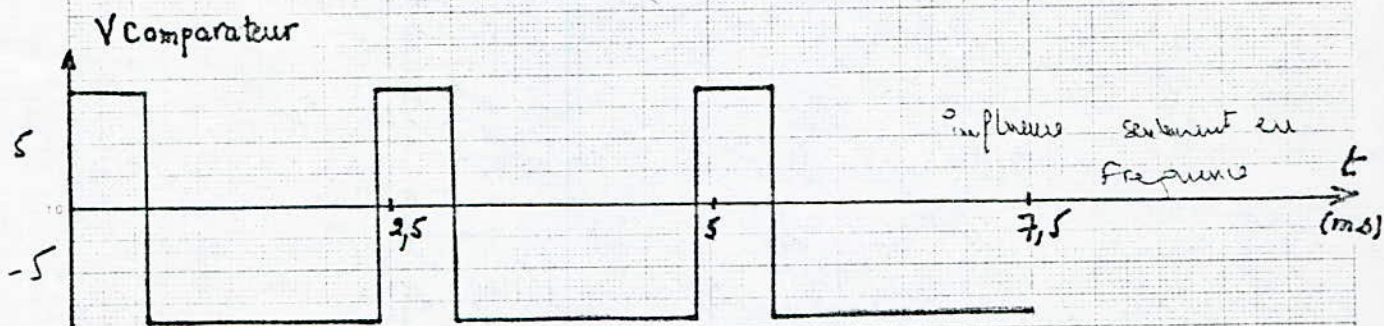
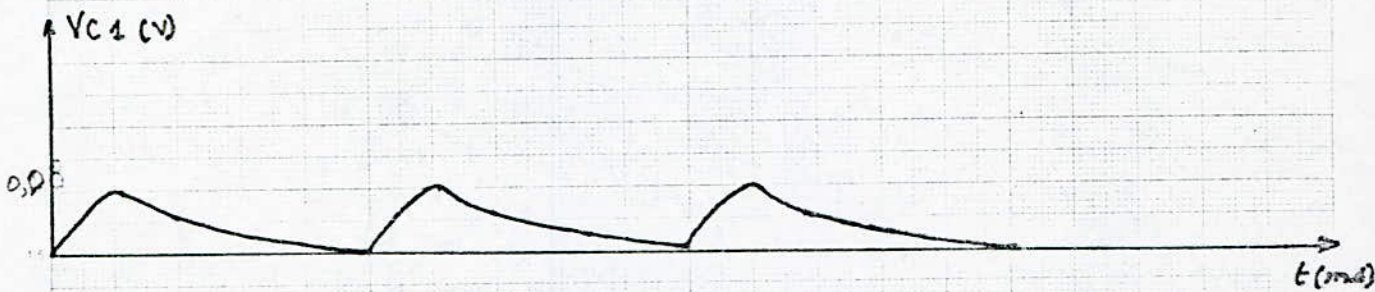
Courbe de Reponse du Filtre  
     passe bande (courbe n°: 1)



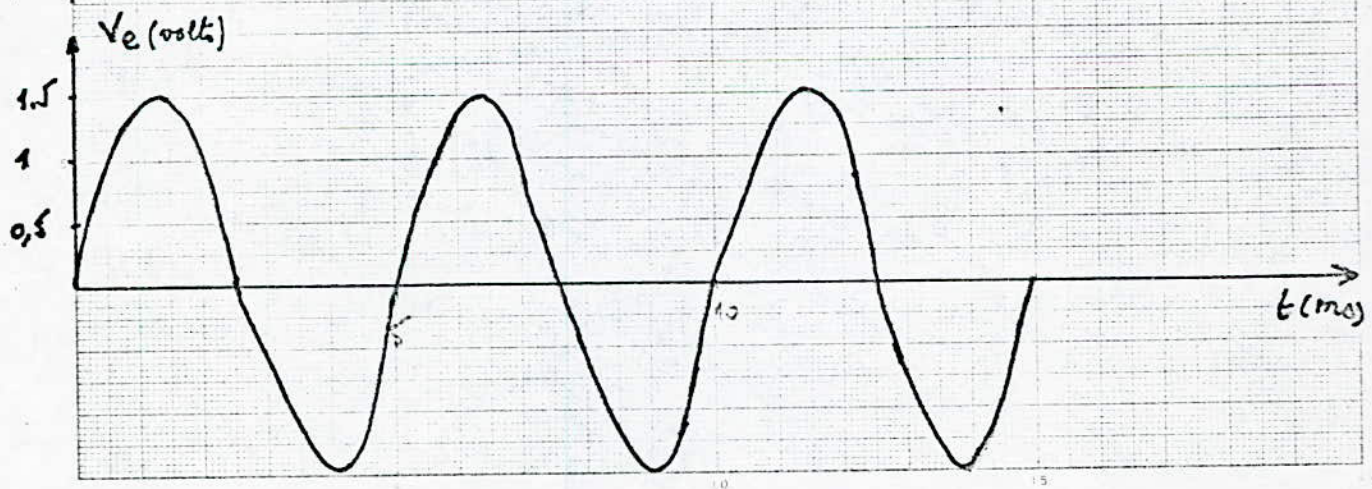
BLOC Compresseur 94 bis (1)



influence en  
fréquence  
et en amplitude  
proportionnellement

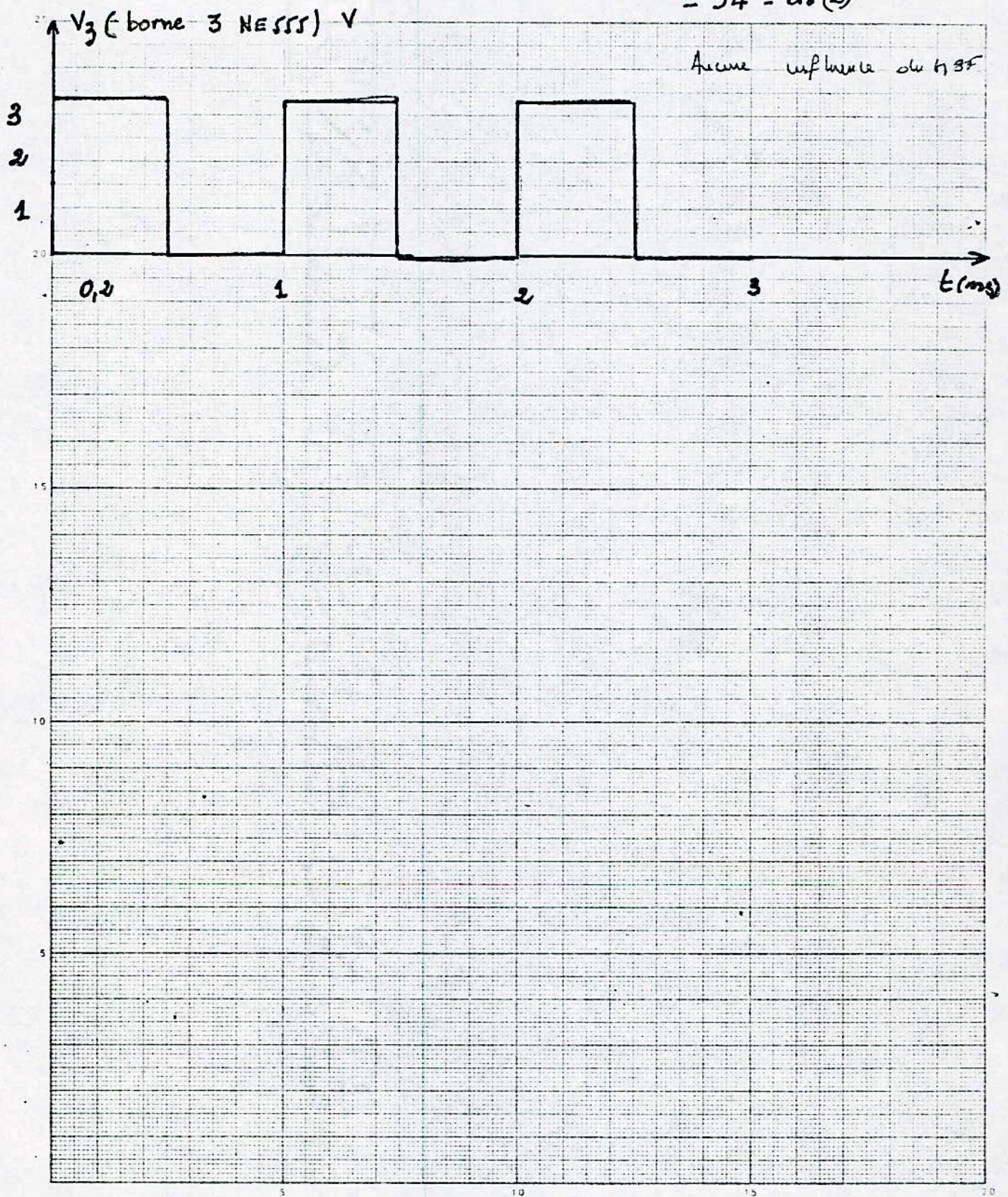


influence seulement en  
fréquence



Courbe n° 2:

- 94 - bis (2)



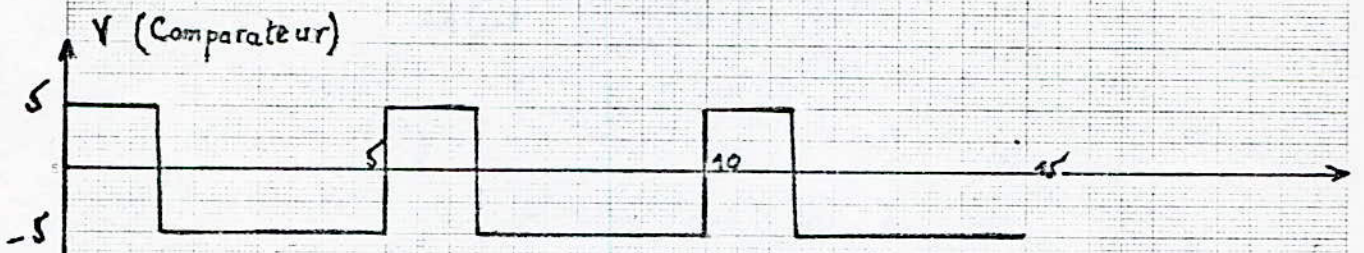
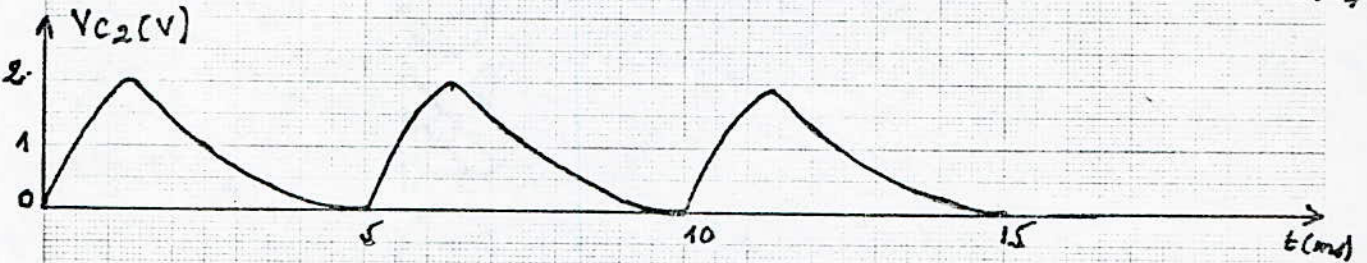
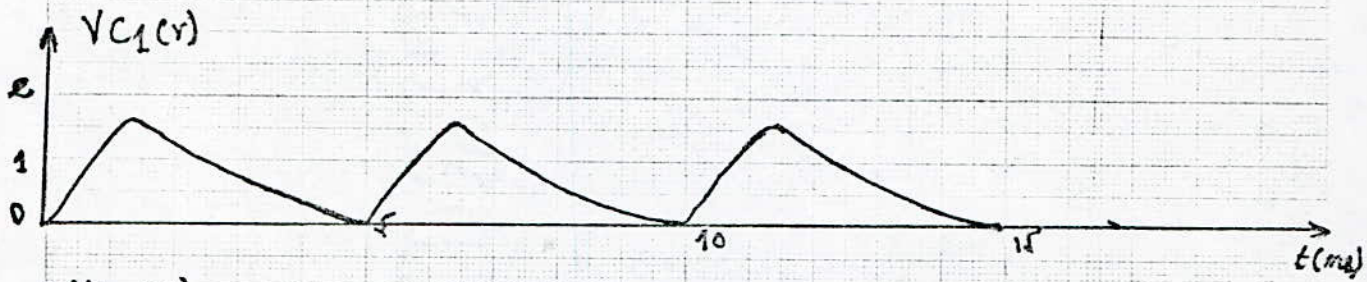
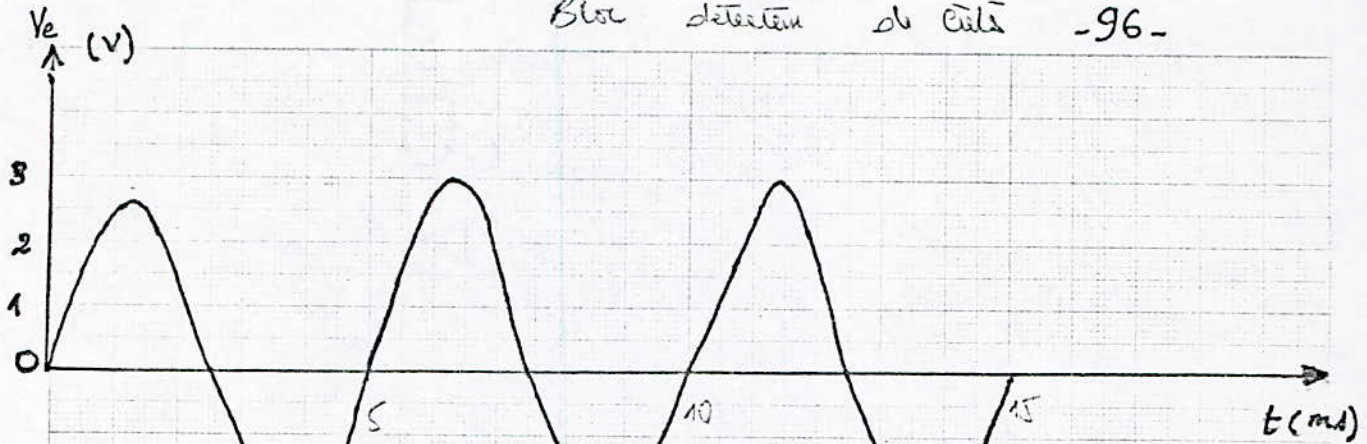
11.2 - COMPRESSEUR

Pour le test se referer a la courbe 2

11.3 - DETECTEUR DE CRETES

*pour les tests concernant ce bloc se reporter à la  
- Courbe n°:3*

Bloc détecteur de crête -96-



influence Proportionnelle de l'amplitude  
et de la fréquence.

III - REALISATION :

III.1 - SCHEMA ELECTRIQUE DU DETECTEUR DE MELODIE :

Voir page suivante

III.2 - CIRCUIT IMPRIME ET SCHEMA D'IMPLANTATION :

Ils seront classes juste apres le circuit electrique du detecteur de melodie .

III.3 - CIRCUIT ELECTRIQUE GLOBAL DU BANC ANALYSEUR :

Se referer en fin de volume .

STRONG

EXT

- 98 -

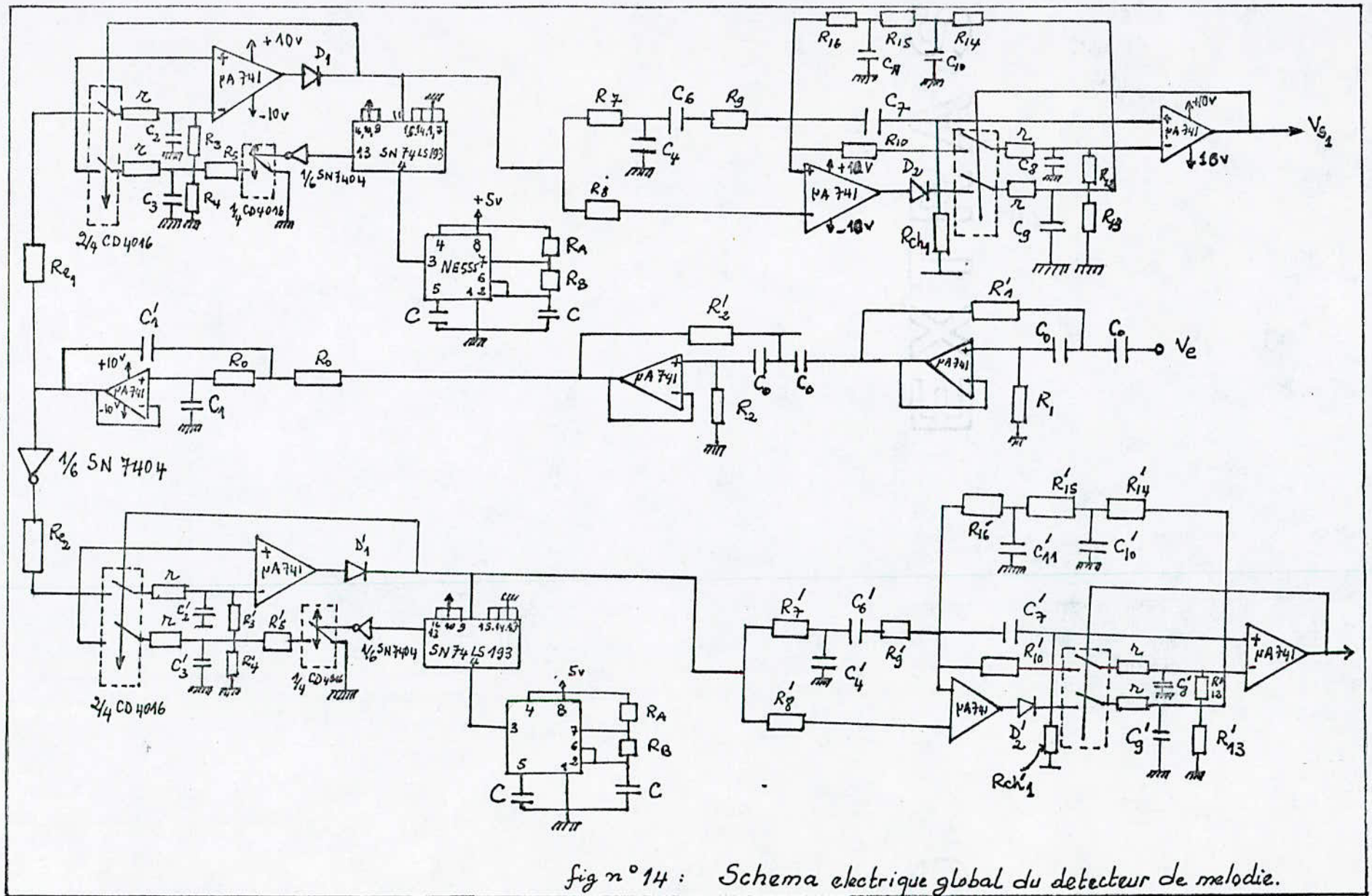
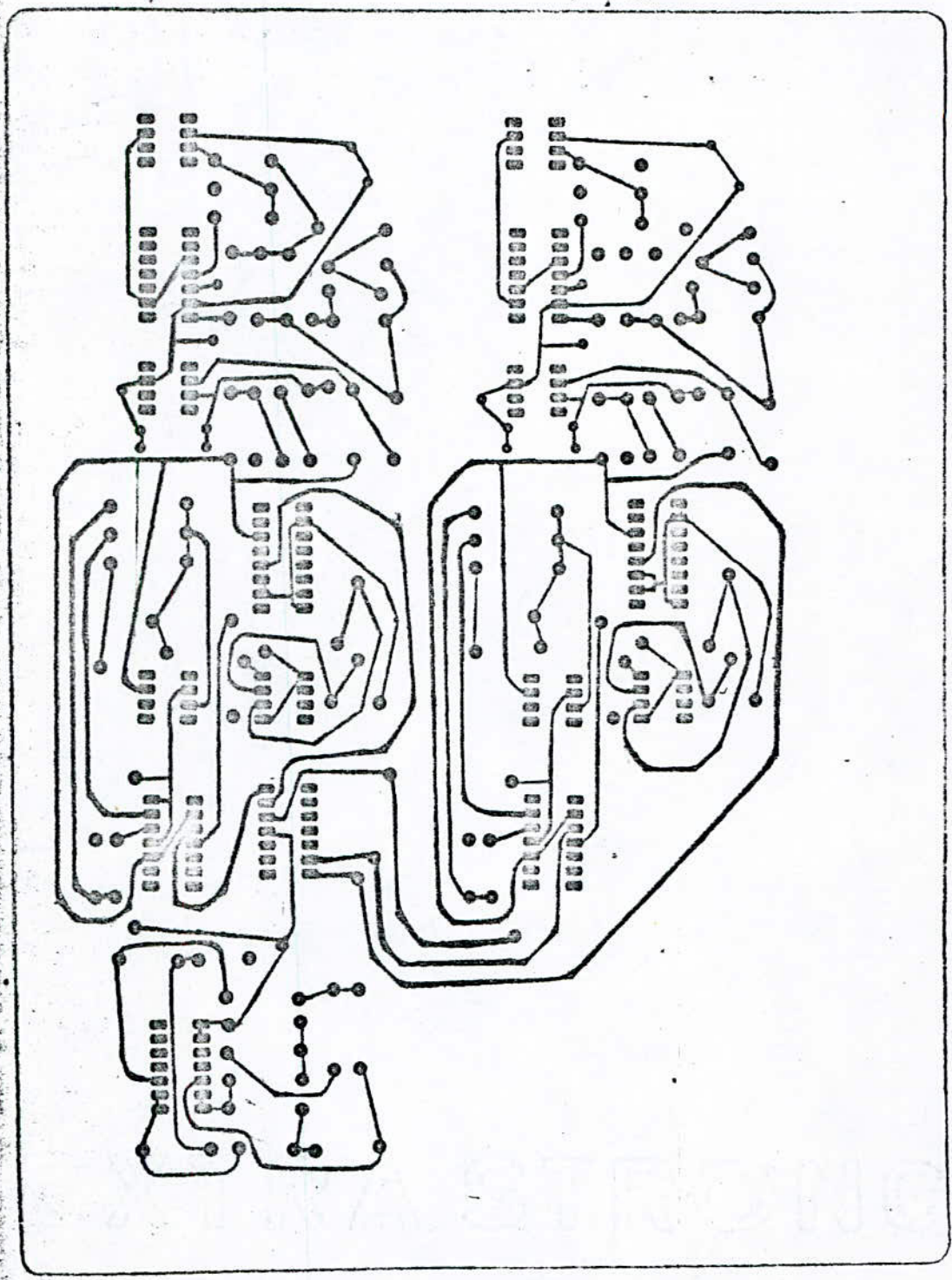
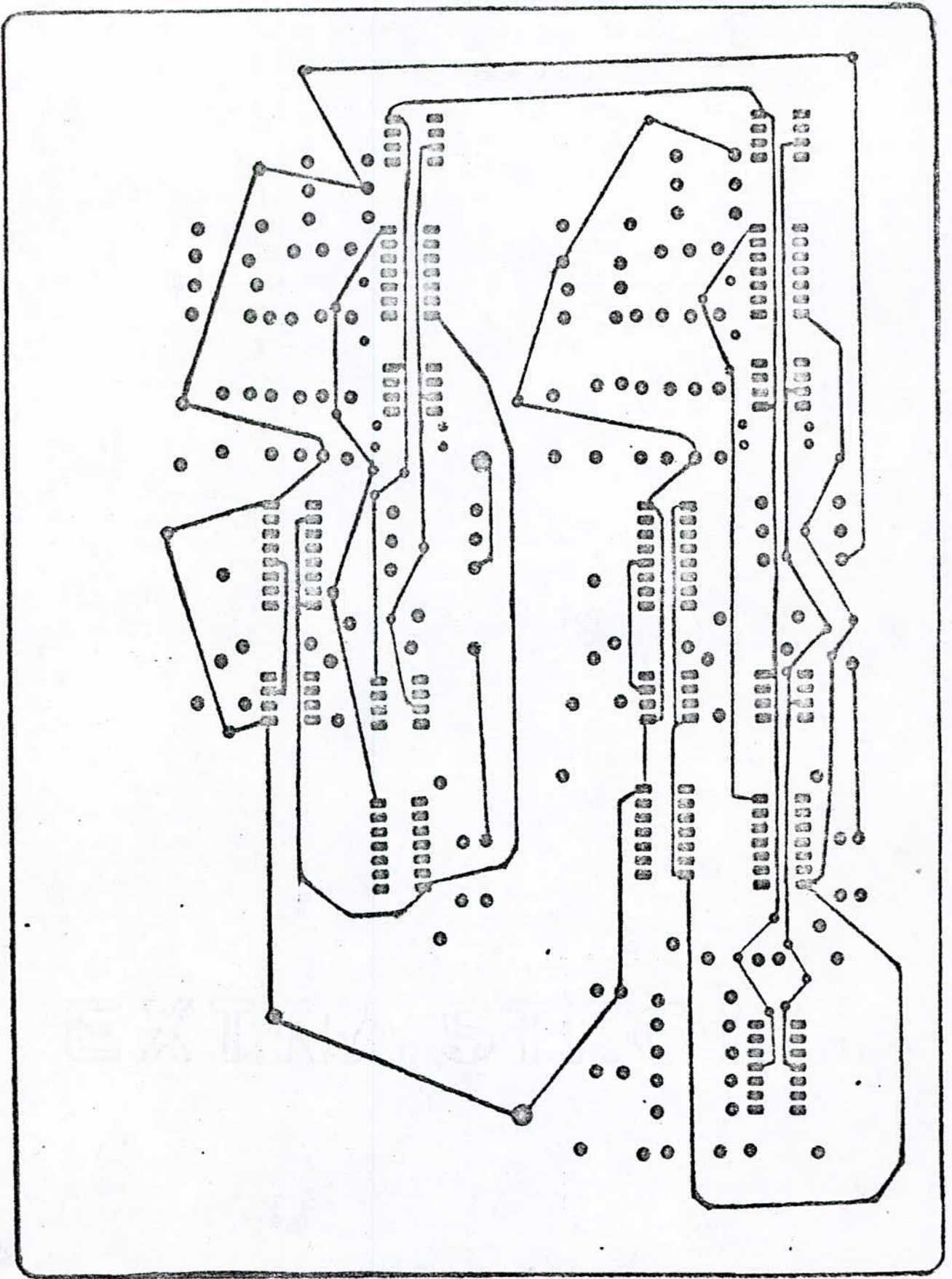


fig n° 14 : Schema electrique global du detecteur de melodie.



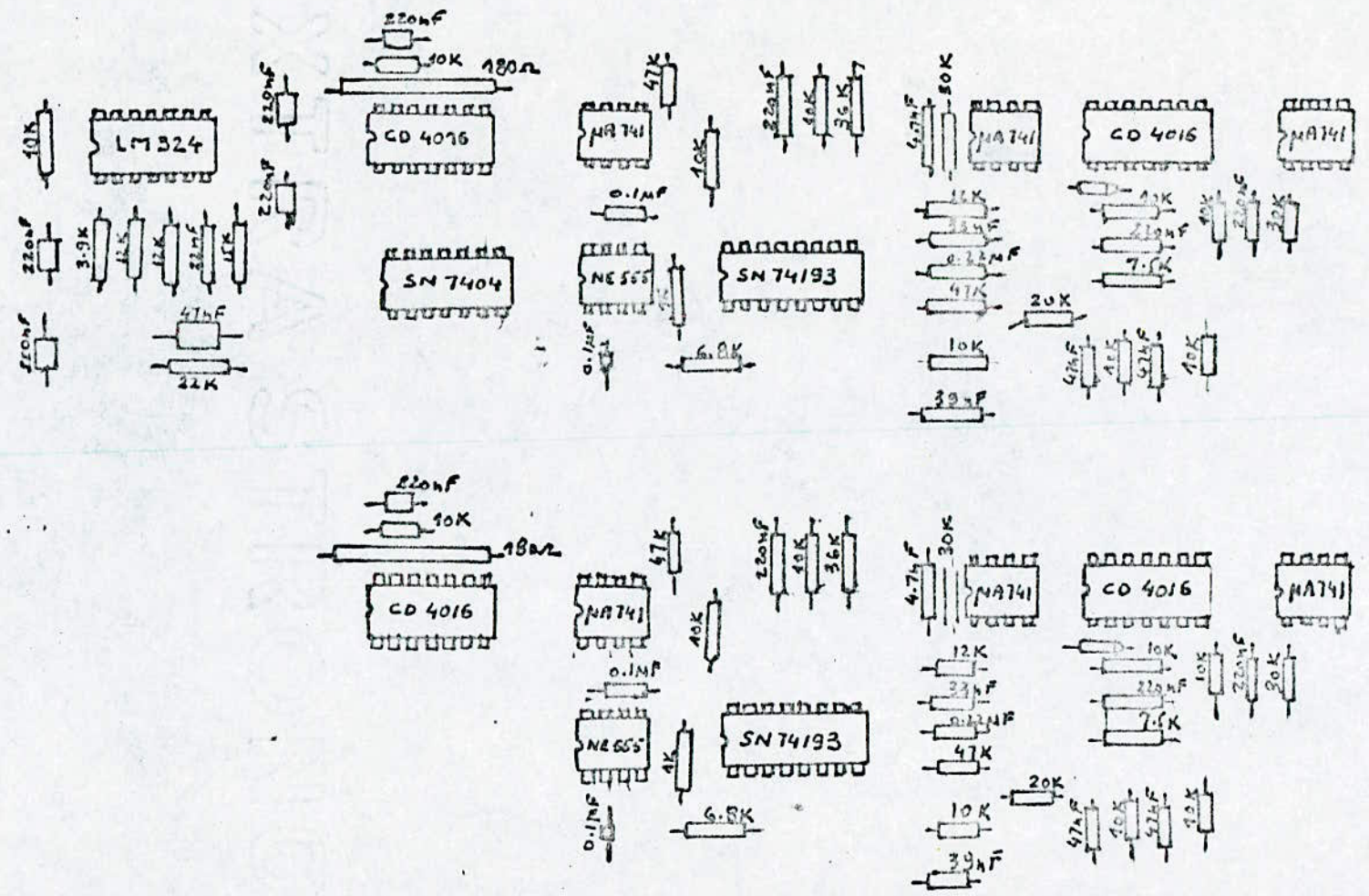
Circuit Imprime du detecteur de Melodie

Cote Composants



Circuit imprime du detecteur de Melodie  
2eme face





Schema D'implantation

CONCLUSIONS GENERALES :

En conclusion nous pouvons dire que ce travail nous a permis d'approfondir nos connaissances, tant en théorie qu'en pratique dans le domaine riche et vaste de l'analyse la reconnaissance, et la synthèse de la parole, dont on ne compte plus aujourd'hui les domaines d'applications.

Notons que le manque de composants, ne nous a pas permis de réaliser un vocodeur, mais cela ne nous a pas privé de réaliser un détecteur de pitch ou de mélodie qui permet de calculer la fréquence  $F_0$  d'un son voisé, de prévoir si le temps et le matériel inexistant en ce moment, nous le permettent, la réalisation du banc analyseur et la carte de mémorisation des données en vue d'une reconnaissance.

D'autre part le manque de documentation précise, concernant le détecteur de voisement, n'a pas rendu possible son édification.

Sur le plan des résultats, nous avons effectué les tests concernant la mélodie dont les courbes et les résultats sont décrits au chapitre VI.

Ils donnent une idée objective, des difficultés pratiques, qu'une telle réalisation est susceptible de rencontrer.

Nous espérons que cette modeste contribution, aura permis l'évolution de ce qui a été déjà élaboré, qu'elle aura aidé à poser les jalons qui permettront aux générations futures de réaliser intégralement le vocodeur à canaux.

Les perspectives d'avenir, dans le domaine du traitement automatique de la parole sont variées. Nous donnons en annexe deux (2), quelques exemples des progrès surprenants faits à ce jour.

L'intelligence artificielle, annonce l'aube d'une ère nouvelle dans le domaine, il semble que nous allons vers le robot " ANDROÏDE ". Est-il pour demain ou pour dans quelques années ? Il est certain qu'il ne courra pas les rues avant longtemps.

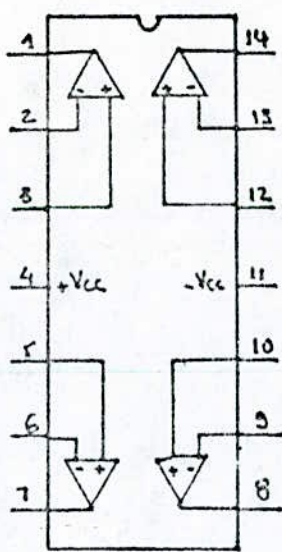
EXTRA STRONG

```
*****  
* ANNEXE 1 *  
*****
```

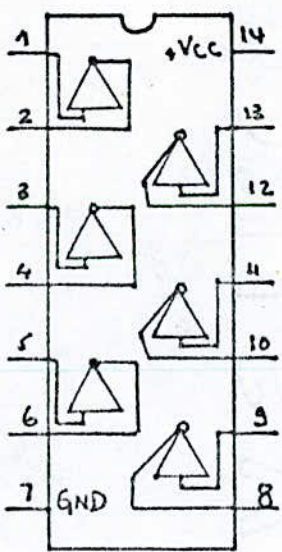
EXTRA STRONG

Circuits de la réalisation

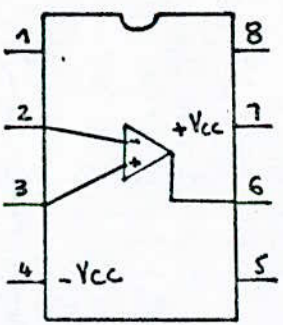
Avec leur notice de brochage



LM324



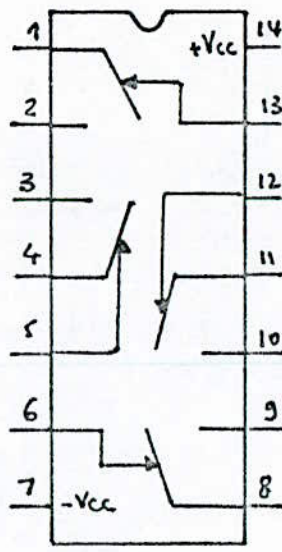
SN7404



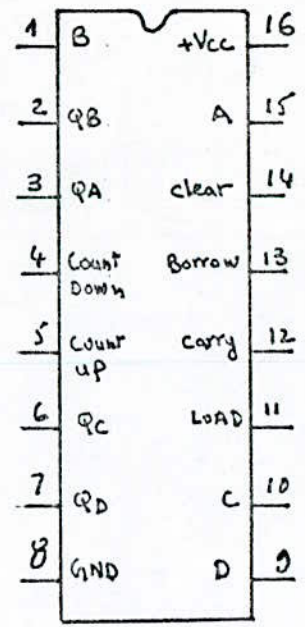
MA 741



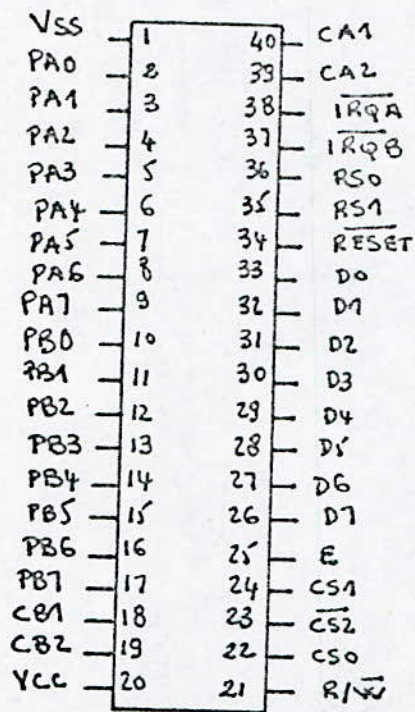
NE 555



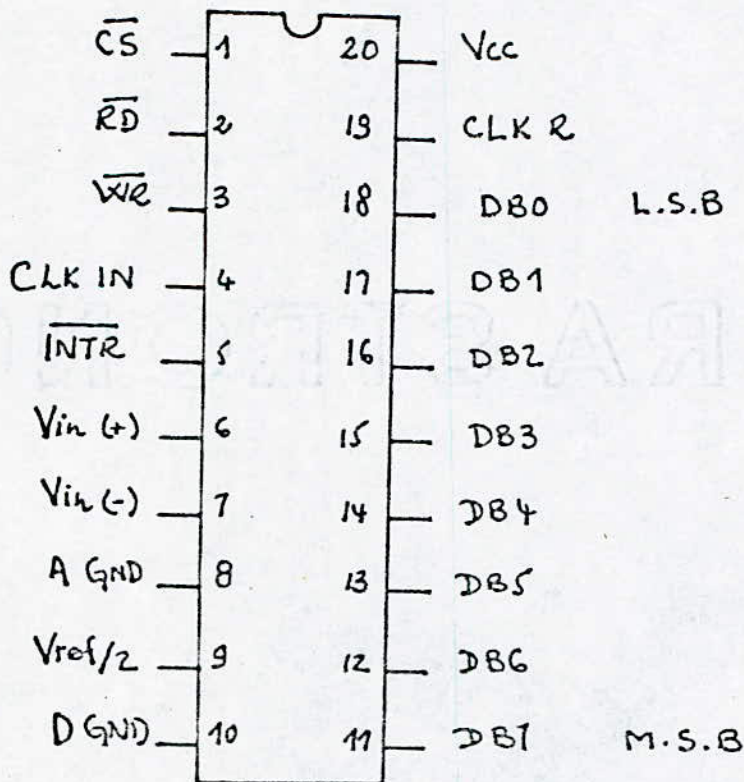
CD 4016



SN74193



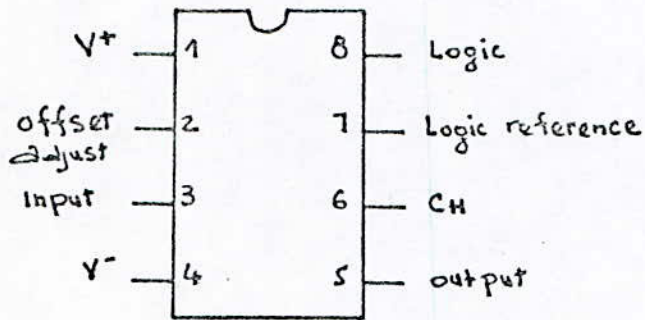
PIA 6821



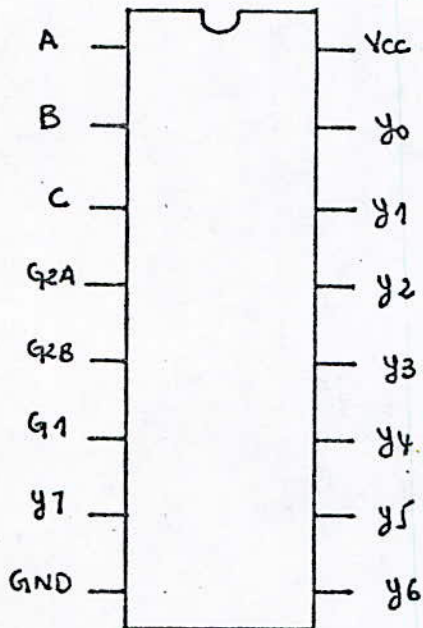
ADC 0804

Caracteristiques

1. Resolution : 8 bits
2. T. C : 100µs
3. Alimentation : 5v
4. erreur : ± 1L.S.B



LF 338



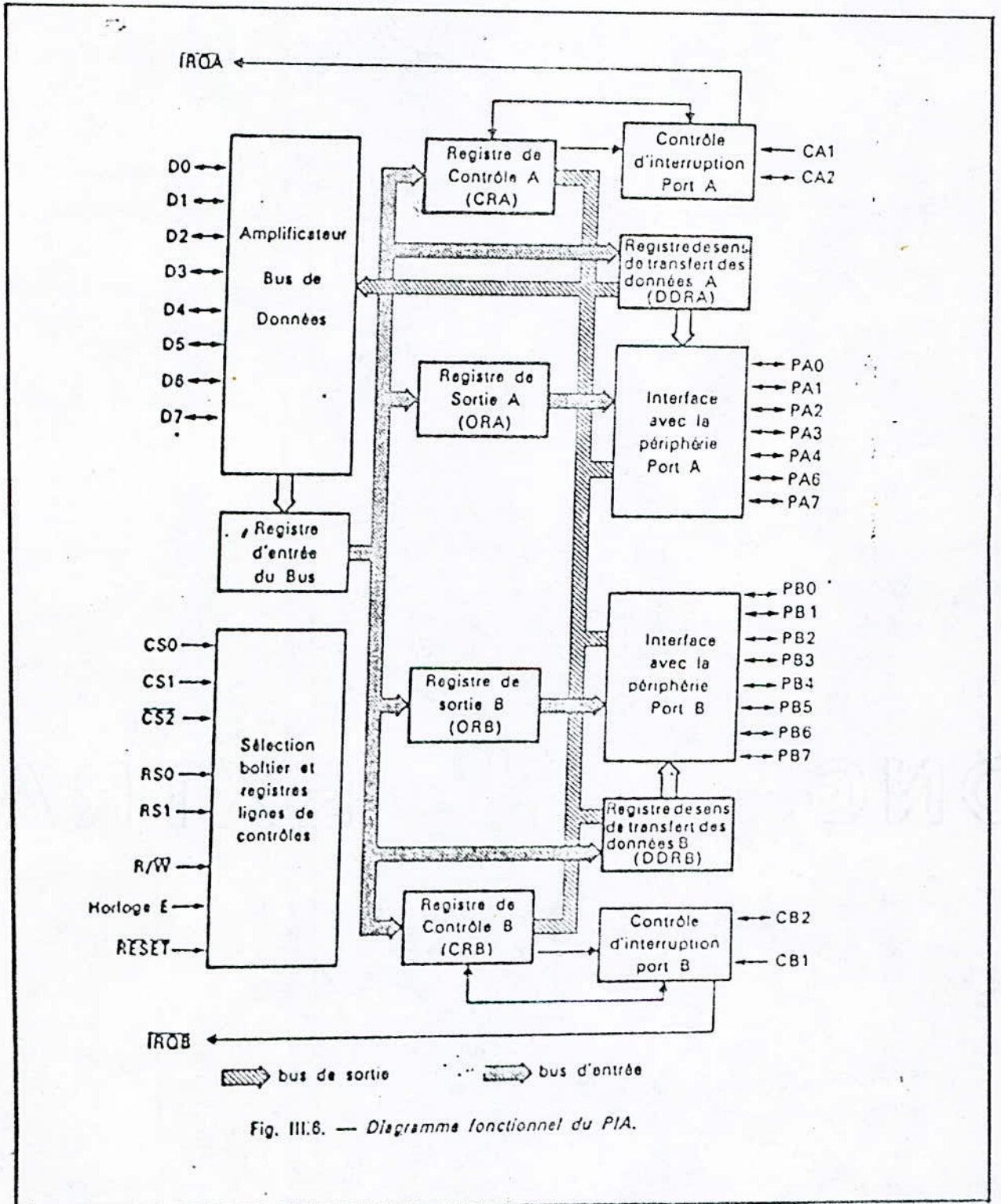
SN 74138

entrees					Sorties							
G <sub>1</sub>	G <sub>2</sub> *	C	B	A	Y <sub>0</sub>	Y <sub>1</sub>	Y <sub>2</sub>	Y <sub>3</sub>	Y <sub>4</sub>	Y <sub>5</sub>	Y <sub>6</sub>	Y <sub>7</sub>
X	1	x	x	x	1	1	1	1	1	1	1	1
0	X	x	x	x	1	1	1	1	1	1	1	1
0	1	0	0	0	0	1	1	1	1	1	1	1
0	1	0	0	1	1	0	1	1	1	1	1	1
0	1	0	1	0	1	1	0	1	1	1	1	1
0	1	0	1	1	1	1	0	1	1	1	1	1
0	1	1	0	0	1	1	1	1	0	1	1	1
0	1	1	0	1	1	1	1	1	1	0	1	1
0	1	1	1	0	1	1	1	1	1	1	0	1
0	1	1	1	1	1	1	1	1	1	1	1	0

Table de Verite

$$G_2^* = G_{2A} + G_{2B}$$

# Architecture Interne du PIA 6821



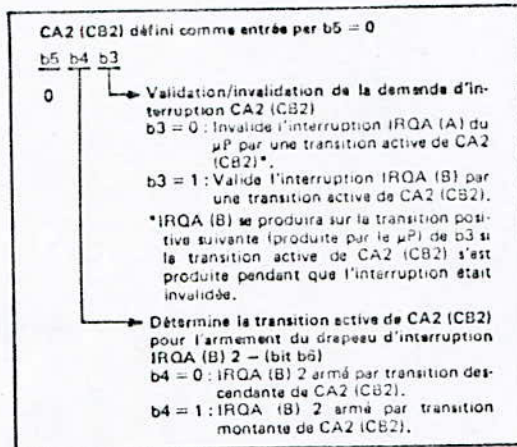


Fig. 4-38. Configuration des bits CA2 et CB2 du registre de commande du PIA. (Avec l'autorisation de Motorola, Inc.)

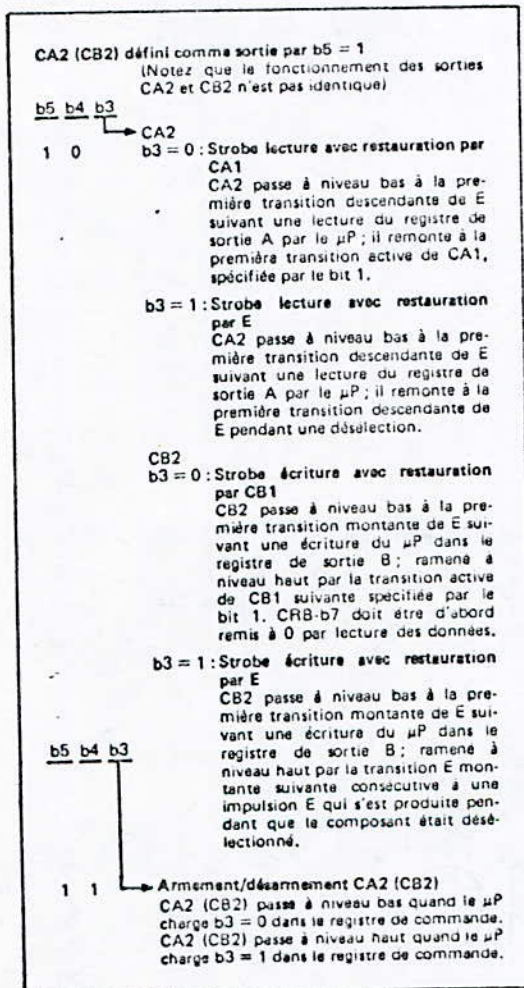


Fig. 4-39. Configuration du CA2 en broche de sortie. (Avec l'autorisation de Motorola, Inc.)

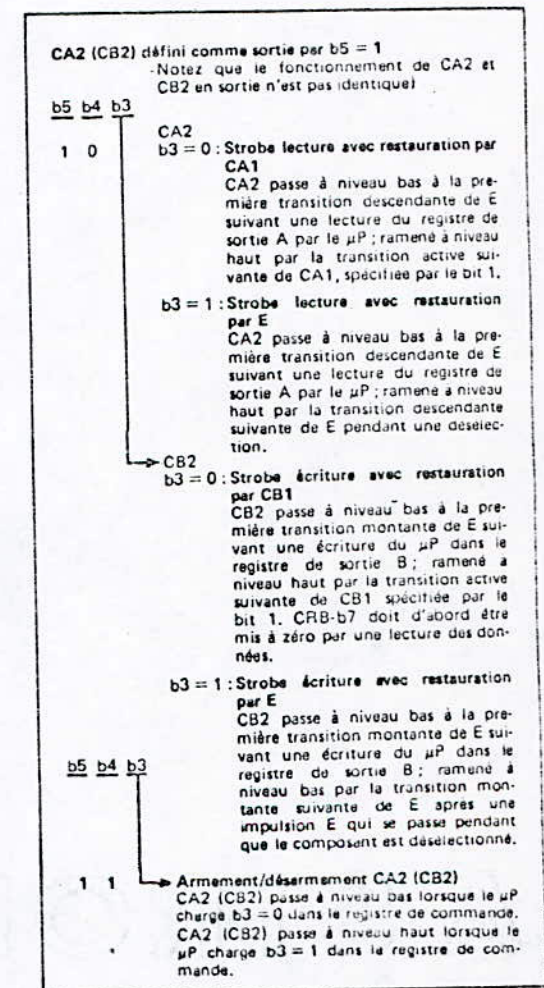


Fig. 4-40. Configuration du CB2 en broche de sortie. (Avec l'autorisation de Motorola, Inc.)



t 1  
des  
t 5  
B.  
ser-  
B1  
les  
nde  
ent  
stat  
eur  
s 0  
B1.  
ons  
: C1  
C2.  
t B  
nde  
ront  
ront  
e de  
istre  
que  
nis à  
iphé-  
gistre  
ande  
idée.  
its 1  
recti-  
taure  
re de  
ue la  
port  
CR1.  
O du  
lables

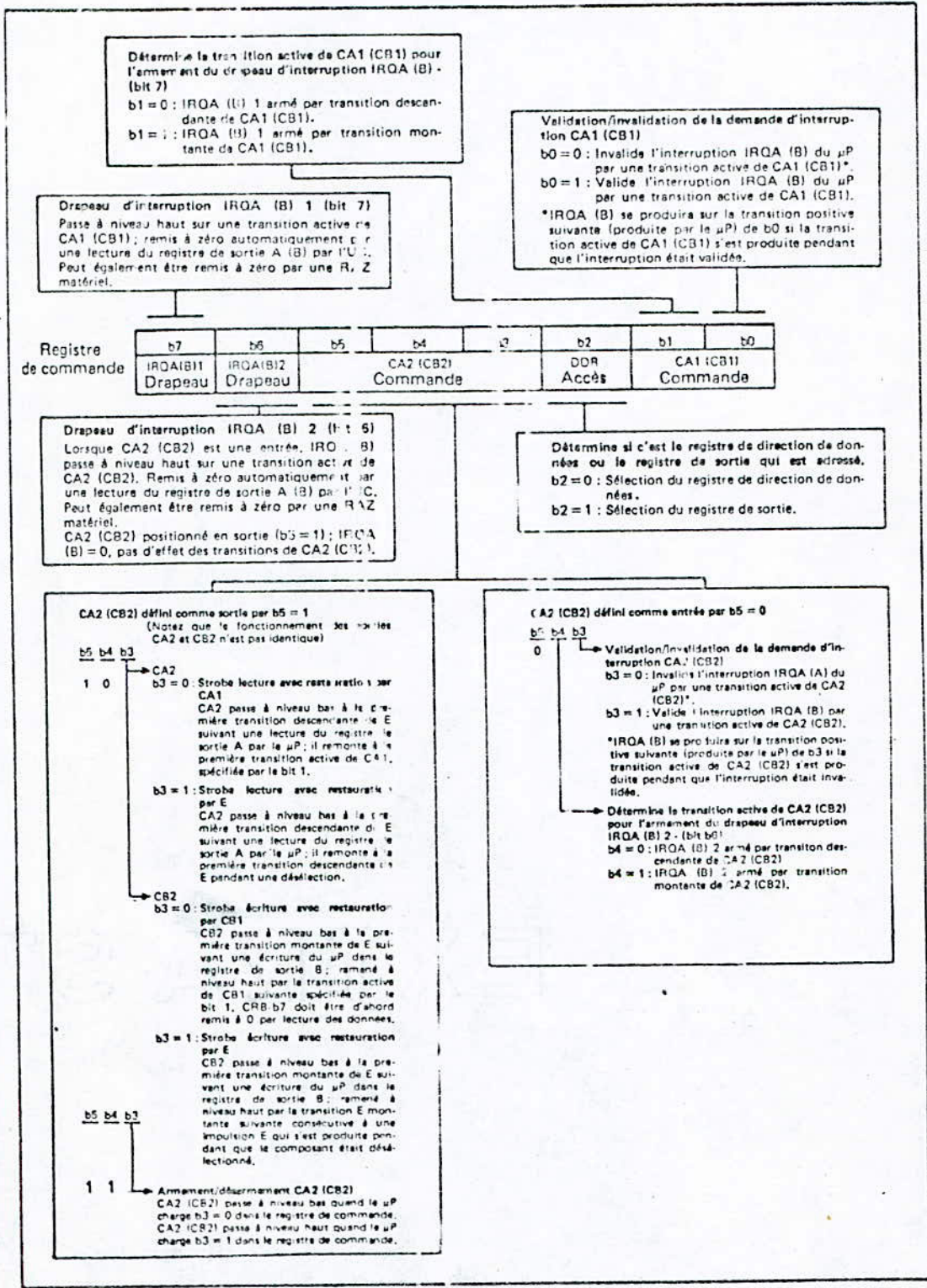


Fig. 4-37. Configuration des bits CA1 et CB1 du registre de commande du PIA. (Avec l'autorisation de Motorola, Inc.)

\*\*\*\*\*  
\*\* ANNEE N \*\*  
\*\*\*\*\*



Télex: 600 778 F

VECSYS 14, avenue de la Gare 91570 Bièvres. Tel. (6) 941 15 04

SARL au capital de 684 200 F. RC Corbeil B316676741

## TRAITEMENT AUTOMATIQUE DE LA PAROLE

Faciliter le dialogue Homme-Machine

tel est l'objectif du

Traitement Automatique de la Parole.

Moyen de communication indépendant des yeux et des mains, la communication orale :

- *simplifie le dialogue avec la machine*
- *facilite la mobilité de l'opérateur*
- *permet d'utiliser les mains et les yeux à une autre tâche*

C'est le canal de dialogue le plus naturel.

VECSYS, en collaboration avec le LIMSI figure parmi les leaders de ce nouveau savoir-faire.

- La Synthèse vocale permet la diffusion de messages :
  - *passifs : ex. annonces dans les lieux publics*
  - *interactifs : en réponse à une action précise.*
- La reconnaissance vocale facilite :
  - *la commande de processus*
  - *la saisie de données*

VECSYS, propose une solution à toutes ces applications.

- synthèse à partir du texte : **ICO85**
- restitution de parole compressée : **LPC, Mic, Delta**
- reconnaissance de mots enchaînés : **RME186**
- reconnaissance de mots isolés : **RMI88, RMI50**
- terminal vocal autonome : **TLV02 TLV50**

Grâce à son expérience de plusieurs années, VECSYS peut également vous conseiller et réaliser à la demande tout système vocal particulier.



## LA PAROLE COMPRESSEE

Différente de la synthèse à partir du texte (IC085), la restitution de la parole compressée est utilisée pour la diffusion de messages fixes de haute qualité.

### Principe :

La parole est numérisée puis codée afin de minimiser l'occupation mémoire (compression). La restitution s'effectue par décodage et conversion numérique/analogique.

Une fois numérisé le signal vocal peut subir différents traitements informatiques :

- transmission numérique dans le réseau PTT
- cryptage pour les communications bancaires ou militaires - etc.

### Les Méthodes :

- codage direct du signal

Le procédé standard utilisé est le Mic (Modulation par impulsions et Codage) qui nécessite un débit de 64 K bits/s.

La modulation delta permet un débit 2 fois plus faible grâce au codage de la différence entre le signal de parole et une prédiction de ce signal (débit 10 à 32 K bits/s).

Ces deux techniques simples à mettre en oeuvre grâce à l'utilisation de circuits spécialisés permettent d'obtenir une parole de très haute qualité au détriment de la taille de la mémoire utilisée.

- codage paramétrique LPC

L'étude de l'appareil phonatoire a permis de le modéliser sous la forme d'un filtre numérique dont les paramètres (une dizaine) sont évalués toutes les 10 à 25 ms.

Les variations lentes du conduit vocal permettent de prédire approximativement un échantillon de parole par une combinaison linéaire d'un certain nombre d'échantillons précédents.

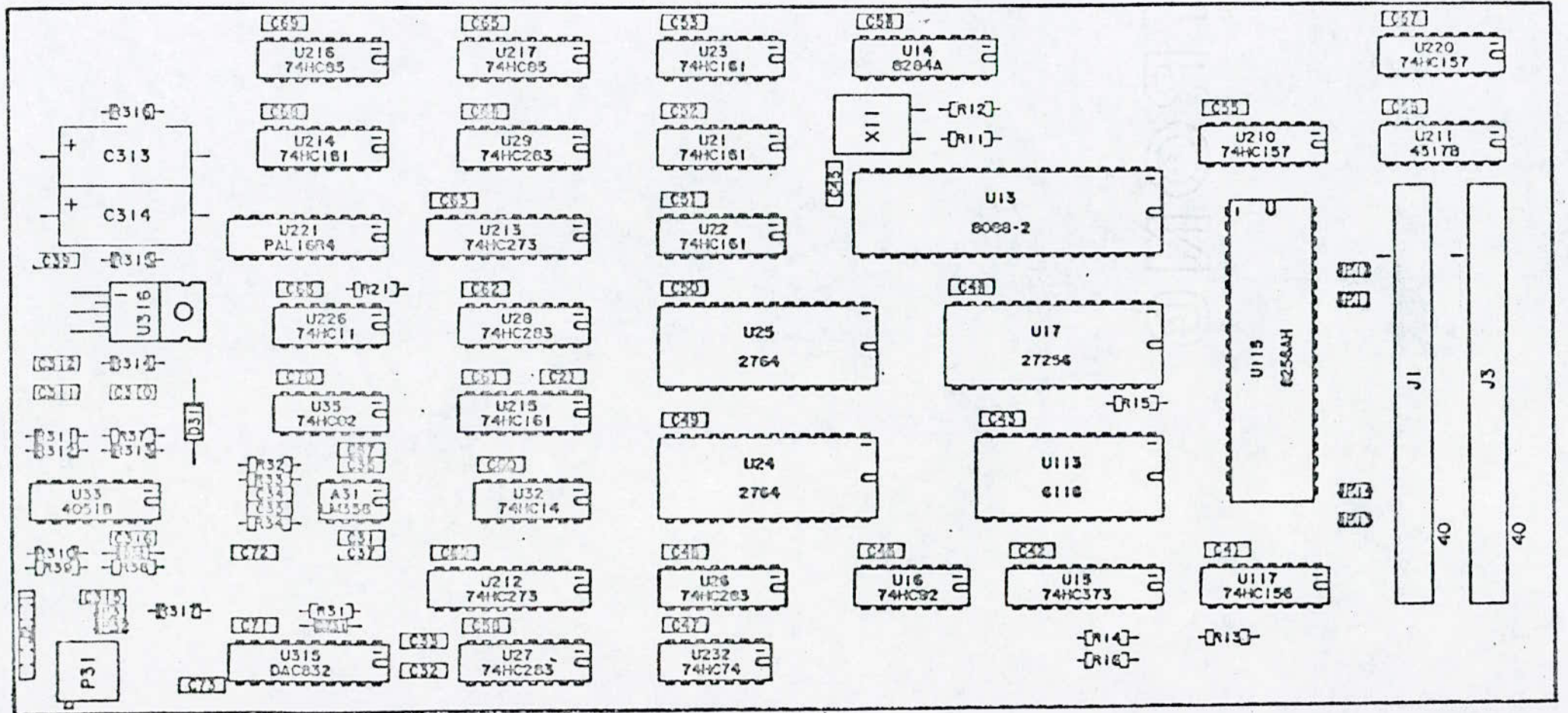
La commercialisation récente de nombreux circuits de synthèse de parole permet de réaliser un système peu onéreux restituant une parole de bonne qualité avec un débit variant de 1200 à 9600 bits/s.

Encombrement réduit et parole de qualité font que la synthèse LPC tend à s'imposer pour la diffusion de messages auprès d'un public exigeant.

VECSYS réalise à la demande le vocabulaire nécessaire à vos applications, notamment pour les chips de synthèse HITACHI

# CARTE SYMPA 3

=====



VECSYS 14, avenue de la Gare 91570 Bièvres, France  
Tél. (1) 69 41 15 04 + - Télex 600 778 F  
S.A. au capital de 686 000 F - R.C. Corbeil B 316676741



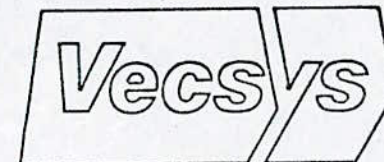
S Y M P A 3  
=+=+=+=+=+

CARTE DE SYNTHÈSE A PARTIR DU TEXTE

- \* SYNTHÈSE DIRECTE D'UN TEXTE ORTHOGRAPHIÉ EN FRANÇAIS
- \* VOCABULAIRE ILLIMITÉ
- \* TRADUCTION DES NOMBRES
- \* SYNTHÈSE PAR DIPHONÈMES
- \* PRISE EN COMPTE DES LIAISONS
- \* REPÉTITION DES MESSAGES
  
- \* INTERFACES : - PARALLÈLE 8 BITS  
                  - SÉRIE TTL (V24 EN OPTION)
  
- \* AMPLIFICATEUR 5W INCORPORÉ
- \* ALIMENTATION +12V/0,2A , +5V/0,6A
- \* FORMAT : 260MM X 120MM

(DOCUMENTATION PRÉLIMINAIRE)

VECSYS - 14, avenue de la Gare 91570 BIEVRES Tél: (6) 941 15 04 - Telex: 600 778 F

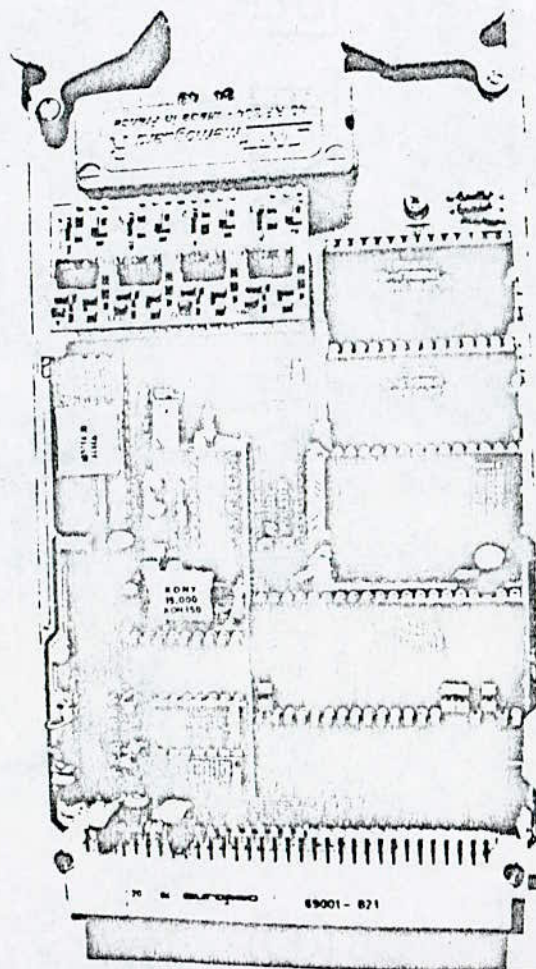


## RMI50A

CARTE DE RECONNAISSANCE DE LA PAROLE  
PAR MOTS ISOLES - MONOLOCUTEUR

### PRINCIPALES CARACTERISTIQUES

VOCABULAIRE .....	version A1 : 47 mots
.....	version A2 : 112 mots
LONGUEUR DES MOTS .....	300 ms à 2s
TEMPS DE REPONSE .....	< 300ms pour 47 mots
TAUX DE RECONNAISSANCE ..	> à 99 %
APPRENTISSAGE .....	une seule passe
NOMBRE DE FILTRES .....	8 (Module hybride VECSYS)
ENTREES AUDIO .....	Micro: 1mV-0.3K $\Omega$
.....	Auxiliaire : 150mV-100K $\Omega$
INTERFACES .....	Série RS232
.....	Parallèle 8 bits
MEMOIRE RAM .....	8 à 16 Ko CMOS
.....	sauvegarde
MEMOIRE ROM .....	16 Ko
MICROPROCESSEUR .....	Intel 8088
ALIMENTATION .....	5V/0.5A $\pm$ 12V/60mA
FORMAT .....	Europe 100x160 mm
SAUVEGARDE VOCABULAIRE ..	100 jours environ





## RME186

### CARTE DE RECONNAISSANCE DE LA PAROLE PAR MOTS ENCHAINES - MONOLOCUTEUR

#### RECONNAISSANCE DE MOTS ENCHAINES :

sans limitation du nombre de mots par séquence  
détection de mots dans un flot de parole continu

#### APPRENTISSAGE :

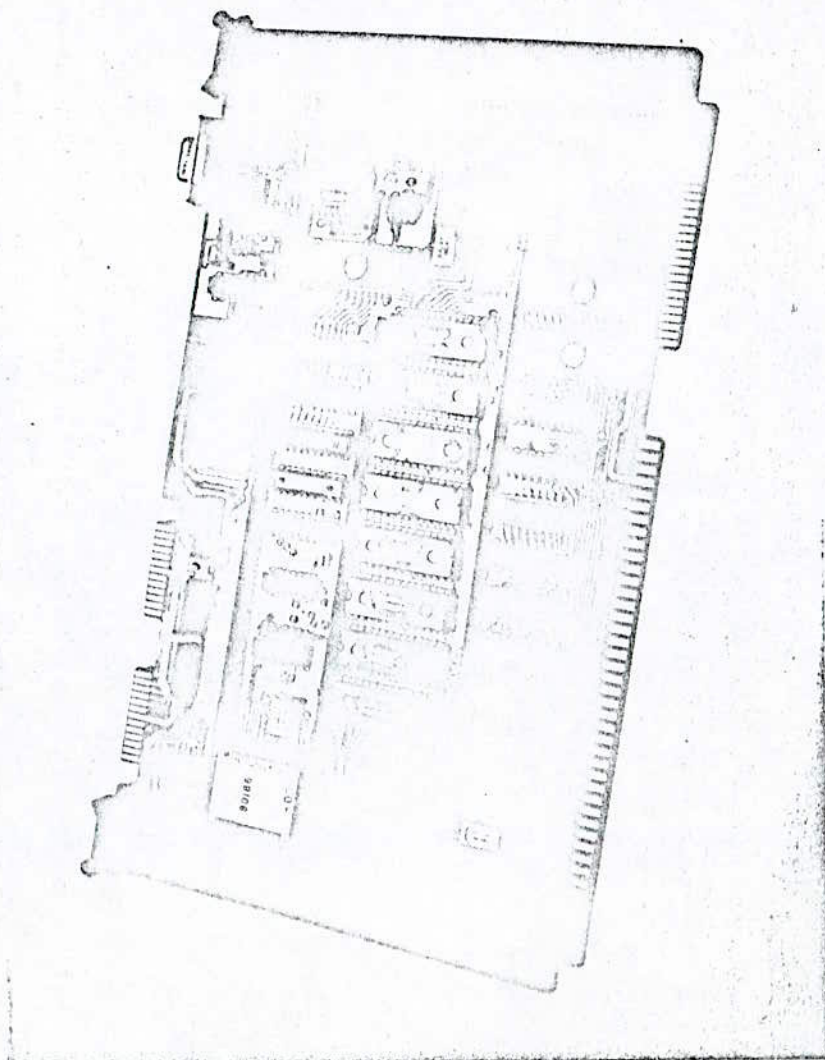
Nombre variable de passes suivant application

Sessions d'apprentissage pour les chiffres et les nombres incluses

#### ANALYSEUR SYNTAXIQUE INTEGRE :

Grammaire hors contexte (structure arborescente)

NOMBRE MOYEN DE REFERENCES .....	200
FACTEUR DE BRANCHEMENT SYNTAXIQUE MOYEN .....	10
NOMBRE DE PARAMETRES AJUSTABLES .....	6
TAUX DE RECONNAISSANCE .....	95% sur les séquences
.....	98% sur les mots
BANC DE FILTRES .....	NEC 7720
MICROPROCESSEUR .....	Intel 80186
MEMOIRE ROM .....	64 Ko
MEMOIRE RAM .....	64 Ko double accès CMOS
INTERFACES .....	serie RS232
.....	IEEE796 (Multibus Intel)
ENTREES AUDIO .....	Micro 1mV-0.3K $\Omega$
.....	Auxiliaire 150mV-100K $\Omega$
FORMAT .....	SBC 304 x 171 mm



spécifications sujettes à modification



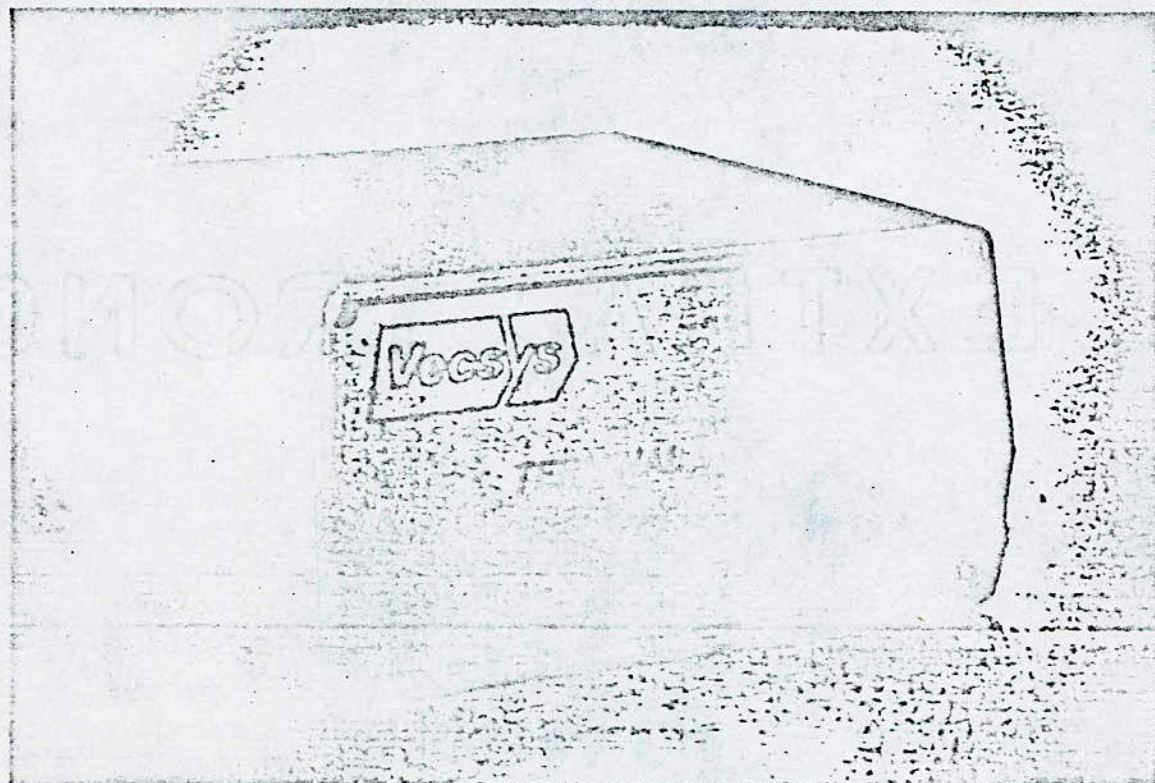
COFFRET DE RECONNAISSANCE DE LA PAROLE  
PAR MOTS ISOLES - MONOLOCUTEUR

PRINCIPALES CARACTERISTIQUES

VOCABULAIRE .....	version A1 : 47 mots
.....	version A2 : 112 mots
LONGUEUR DES MOTS .....	300 ms à 2s
TEMPS DE REPOSE .....	< 300ms pour 47 mots
TAUX DE RECONNAISSANCE ..	> à 99 %
APPRENTISSAGE .....	une seule passe
NOMBRE DE FILTRES .....	8 (Module hybride VECSYS)
ENTREES AUDIO .....	Micro: 1mV-0.3K $\Omega$
.....	Auxiliaire : 150mV-100K $\Omega$
INTERFACE .....	Série RS232
MEMOIRE RAM .....	8 à 16 Ko CMOS
.....	sauvegarde
MEMOIRE ROM .....	16 Ko
MICROPROCESSEUR .....	Intel 8088
ALIMENTATION .....	220v
SAUEGARDE VOCABULAIRE	100 jours environ
DIMENSIONS .....	long:180 prof:200 haut:90

spécifications sujettes à modification.

VECSYS 14, avenue de la Gare 91570 Évryres FRANCE  
Tél. (6) 69 41 15 04 - Télex 600 778 F





T A R I F S A U 1 S E P T E M B R E 1 9 8 6

Prix H.T. en Francs

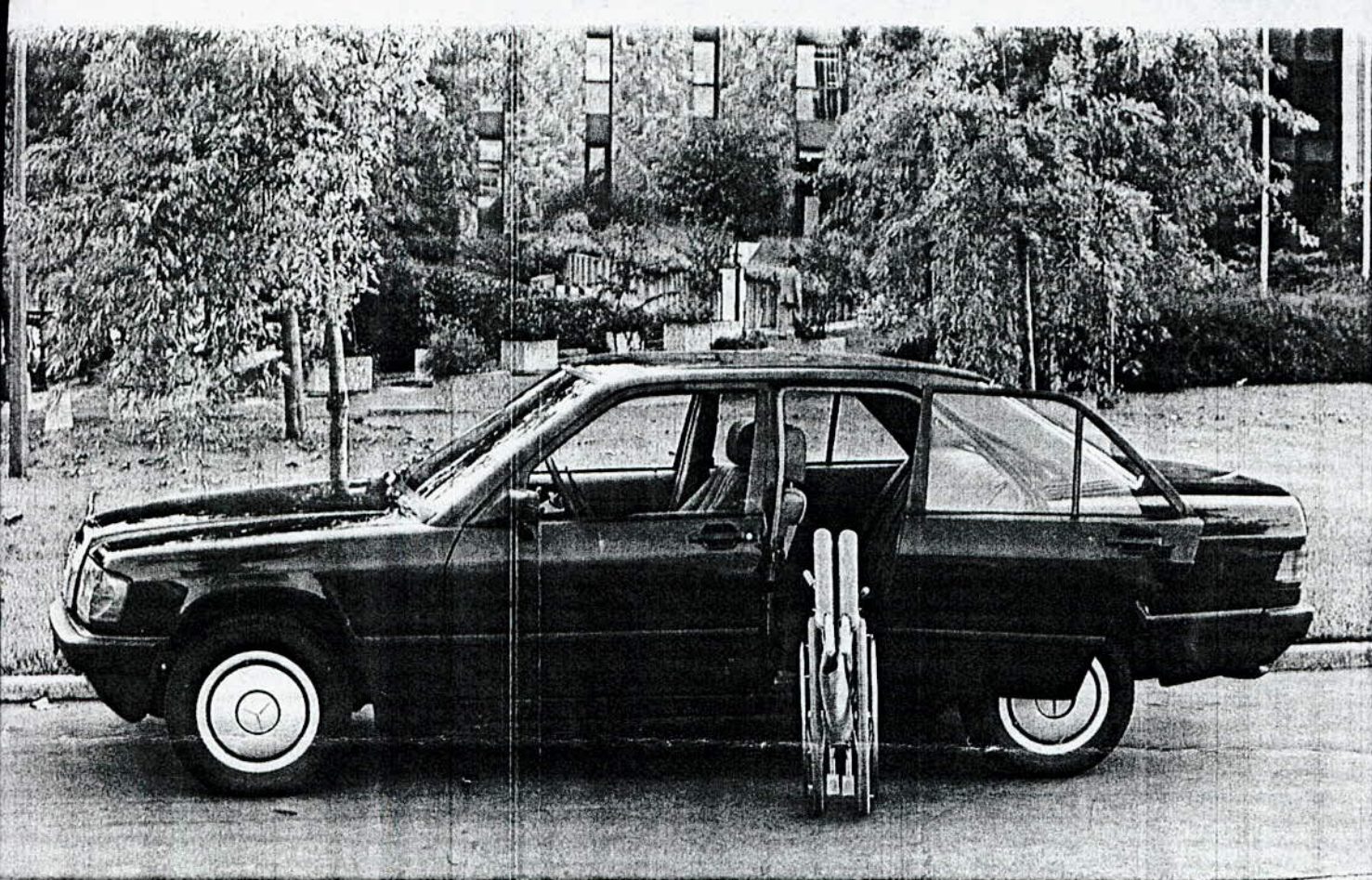
TRAITEMENT AUTOMATIQUE DE LA PAROLE

REFERENCE	I	I	I	2à10	I	11à24	I	25à99	I
I RMI50 A	:	:	:	:	:	:	:	:	I
I Carte 47 mots isolés	:	5.490	:	4.980	:	4.360	:	3.820	I
I Carte 112 mots isolés	:	6.330	:	5.770	:	5.100	:	4.510	I
I RMI88	:	:	:	:	:	:	:	:	I
I Carte 240 mots isolés, Multibus	:	14.900	:	13.400	:	11.920	:	10.430	I
I RME186	:	:	:	:	:	:	:	:	I
I Carte Mots enchaînés, Multibus	:	29.800	:	28.300	:	26.200	:	24.900	I
I IC085	:	:	:	:	:	:	:	:	I
I Carte Synthèse à partir du texte, Multibus	:	9.980	:	8.950	:	7.900	:	6.700	I
I SYMPA 1	:	:	:	:	:	:	:	:	I
I Synthèse à partir du Texte. Coffret Autonome	:	4.980	:	4.530	:	3.984	:	3.486	I
I SYMPA 2	:	:	:	:	:	:	:	:	I
I Idem + Clavier non voyants	:	5.630	:	5.130	:	4.510	:	3.940	I
I SYMPA 3	:	:	:	:	:	:	:	:	I
I Carte synthèse	:	3.970	:	3.650	:	3.340	:	2.990	I
I TLV02	:	:	:	:	:	:	:	:	I
I Coffret Terminal Vocal Format Multibus	:	8.340	:	7.780	:	7.250	:	6.750	I
I TLV50	:	:	:	:	:	:	:	:	I
I Coffret 47 mots isolés	:	6.980	:	6.280	:	5.950	:	5.230	I
I Coffret 112 mots isolés	:	7.820	:	7.040	:	6.670	:	5.865	I
I MICRO SHURE	:	:	:	:	:	:	:	:	I
I SM10	:	1.590	:	:	:	:	:	:	I
I VR230	:	1.040	:	:	:	:	:	:	I

Prix départ Bièvres  
 Garantie 3 mois

Logiciels particuliers sur demande:

- Fonctionnement en mode autonome.
- Sortie des sonogrammes.
- Multilocuteurs.
- "Word spotting".
- Mots enchaînés.
- Logiciels d'application.
- etc.



**ADVANCED TECHNOLOGY  
FOR DISABLED DRIVERS**

**KEMPF**  
P.O. Box 61103  
SUNNYVALE, CA 94086 - 61103  
Tel. (408) 773-0219 U.S.A.

**MODERNE TECHNIK FÜR  
KÖRPERBEHINDERTE AUTOFAHRER**

**KEMPF**  
Lehmkaulenweg 35  
D 5305 ALFTER GIELSDORF  
Tel. (0228) 64 55 82

**EQUIPEMENTS POUR CONDUCTEURS  
HANDICAPES PHYSIQUES**

**KEMPF**  
11, rue Duranton  
F 75015 PARIS  
Tel. (1) 45 54 22 55 - Telex 203 038 F

**VOOR FYSISCH GEHANDICAPTE  
BESTUURDERS**

**KEMPF BENELUX**  
Vuurberg 44 - B 1920 DIEGEM  
Near Brussels Airport  
Tel. (2) 721 42 37



**KEMPF**  
F 67370 DOSSENHEIM-KOCHERSBERG  
Tel. (0033) 88 69 62 87 - Telex 890 970 F  
Boîte Postale 57 - F 67034 STRASBOURG-CEDEX



## **Contrôle d'environnement par commandes vocales :**

**L**e micro-ordinateur KATALAVOX réalisé par Kempf-Equipements, assure un service fiable pour contrôler les différentes commodités d'une chambre à l'usage de personnes atteintes de tétraplégie.

**L**e microphone peut être placé sur la table de chevet par exemple. L'appareil, selon les versions, est soit portatif et sans fils ou bien relié au secteur.

**L**a mise en service est aisée : l'utilisateur prononce la liste des mots de commande, par exemple «lumière, radio, ventilateur, sonnette, télévision» et un code spécial par exemple «Sésame, ouvre, ordinateur». Dès lors KATALAVOX est prêt à fonctionner.

**P**our désactiver KATALAVOX il suffit de prononcer le mot-codé «Sésame». On est ainsi certain qu'une conversation n'active involontairement une des fonctions. De même l'ordinateur redevient utilisable après les mots : «Sésame, ouvre, ordinateur».

Bien entendu il existe un téléphone main-libre actionné par KATALAVOX.

**L**'expérience montre qu'il est nécessaire d'apporter au tétraplégique une aide individuelle ; Kempf-Equipements est en mesure de réaliser le contrôle d'environnement adéquat, y compris la fourniture d'un lit motorisé commandé à la voix.

---

## **KATALAVOX dans l'automobile**

**L**es commandes vocales sont une aide précieuse pour des conducteurs dysméliques ou tétraplégiques incomplets.

En France le premier permis de conduire avec emploi de commandes vocales pour les fonctions électriques a été délivré le vendredi 13 janvier 1984. Bien entendu ce véhicule, équipé par Kempf-Equipements, disposait du KATALAVOX.

**D**epuis le 14 juin 1981, date de la présentation en première mondiale d'un véhicule à commandes vocales, Kempf-Equipements a accumulé une expérience inégalée dans ce domaine spécifique.



# commandes vocal es pour fauteuil roulant

**C**e fauteuil roulant est commandé à la voix ; son usage est moins contraignant que la commande au menton. Même avec une aide respiratoire le tétraplégique peut conduire aisément en parlant à voix basse grâce à l'usage d'un micro-laryngophone. L'arrêt d'urgence et la coupure de désactivation est réalisé par un contacteur placé dans l'appuie-tête.

**A**pprendre à conduire ce fauteuil est aisé : il suffit de prononcer les cinq mots de commande. Le vocabulaire peut être changé instantanément : il suffit de pousser sur un bouton pour effacer la mémoire et prononcer une autre série de cinq mots.

Par exemple :

- en avant
- en arrière
- à droite
- à gauche
- ralentir

L'ordinateur combine ces 5 mots en 66 situations de conduite : il y a quatre vitesses avant, deux vitesses de marche arrière, trois angles de virage, la possibilité de tourner sur place à différentes allures, de quitter ce mouvement de rotation pour entamer une ligne droite sans arrêt préalable, etc...

On peut passer de marche avant en marche arrière sans prononcer le mot stop : l'ordinateur effectue tout seul le freinage avant d'alimenter les moteurs en sens inverse.

**A**ctuellement Kempf-Equipements livre ce fauteuil (matériel de base Ortopédia type 924) après consultation individuelle. Il est possible d'inclure un contrôle d'environnement qui permet depuis le fauteuil d'ouvrir à distance une porte, d'allumer un éclairage, etc... L'assise du fauteuil est soit un modèle «Postura» ou bien un siège d'automobile de la gamme «Recaro». L'ordinateur de commandes vocales est le KATALA-VOX réalisé par Kempf-Equipements : ce matériel a fait ses preuves dans l'automobile.

STRASBOURG

KEMPF  
57  
034 STRASBOURG-CEDEX  
88 69 62 87  
x 890 970 F

PARIS

Ets KEMPF  
11, rue Durantou  
F 75015 PARIS  
Tél. (1) 45 54 22 55  
Télex 203 038 F

BRUSSELS

KEMPF-BENELUX  
Vuurberg 44  
B 1920 DIEGEM  
Tél. (2) 721 42 37

BONN

KEMPF Jean-Daniel  
Lehmkaulenweg 35  
D 5305 ALFTER  
GIELSDORF  
Tél. (0228) 64 55 82





COMMANDES  
VOCALES

**KATALAVOX**

## *Micro-ordinateur de commandes vocales à 15 relais programmables*

Ce système monocuteur utilise des mots isolés préalablement mémorisés par l'utilisateur. Il est muni de 15 relais dont les sorties sans potentiel sont capables de couvrir 2 ampères. Le nombre de mots et le codage des sorties sont déterminés par l'utilisateur, mémorisés en EEPROM, et par conséquent modifiables à volonté. Il est possible de changer un seul mot sans changer le reste du vocabulaire.

A la commande il convient d'indiquer la prise désirée :

7 plots pour 2 sorties par relais (fermeture)  
10 plots pour 3 sorties par relais  
L'alimentation 12 volts passe également par cette prise ainsi qu'un signal de mise à la masse pour couper le micro du type dynamique 200 ohms.

Les 15 relais peuvent :

fermer pendant 0,3 secondes  
fermer et ouvrir alternativement  
être codés par paire (3 paires max.) pour que le mot A ferme le relais A et ouvre B, et que le mot B ferme le relais B et ouvre A, enfin qu'un mot C ouvre A et B (par exemple : «gauche», «droite», «arrêt»)  
être tous ouverts par un seul mot

Le vocabulaire de référence (19 mots maximum) est stocké dans une mémoire EEPROM : dès que l'ordinateur est mis sous tension le contenu est versé en RAM et le système est automatiquement prêt à fonctionner. Un second vocabulaire de 19 mots maximum est disponible en actionnant préalablement une touche.

Bien entendu des codages spécifiques peuvent être étudiés ; par exemple mise en veilleuse par 2 mots de passe, etc...

Depuis 1983 KATALAVOX est utilisé pour des applications d'aide aux handicapés : fauteuil roulant, commande d'environnement, téléphone.

Les premiers permis de conduire catégorie F ont été délivrés début 1984 avec la mention «commandes vocales pour les fonctions électriques».

Des voitures-école pour handicapés utilisant KATALAVOX sont utilisées en France, Benelux et Norvège.

**KATALAVOX** est réalisé par Martine Kempf. (voir photographie ci-dessus).



Ce véhicule est utilisé pour la conduite sans l'aide des deux bras : une direction au pied gauche, des pédales d'accélérateur et de frein au pied droit. Toutes les fonctions secondaires sont effectuées à la voix y compris ouverture-fermeture de porte, pare-soleil, lève-vitres, frein de parcage et bien entendu clignotants, avertisseur, lumières, etc...

STRASBOURG

KEMPF  
7  
34 STRASBOURG-CEDEX  
8 69 62 87  
890 970 F

PARIS

Ets KEMPF  
11, rue Duranton  
F 75015 PARIS  
Tél. (1) 45 54 22 55  
Télex 203 038 F

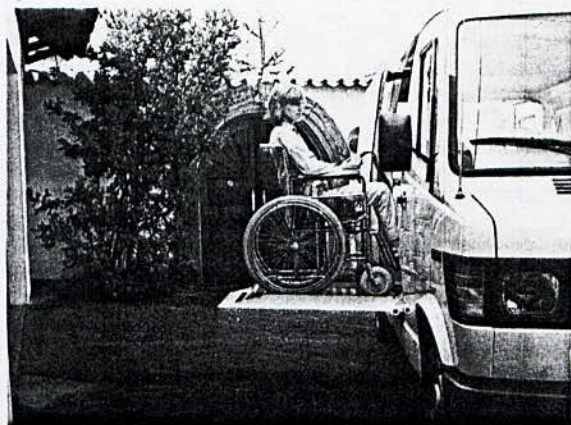
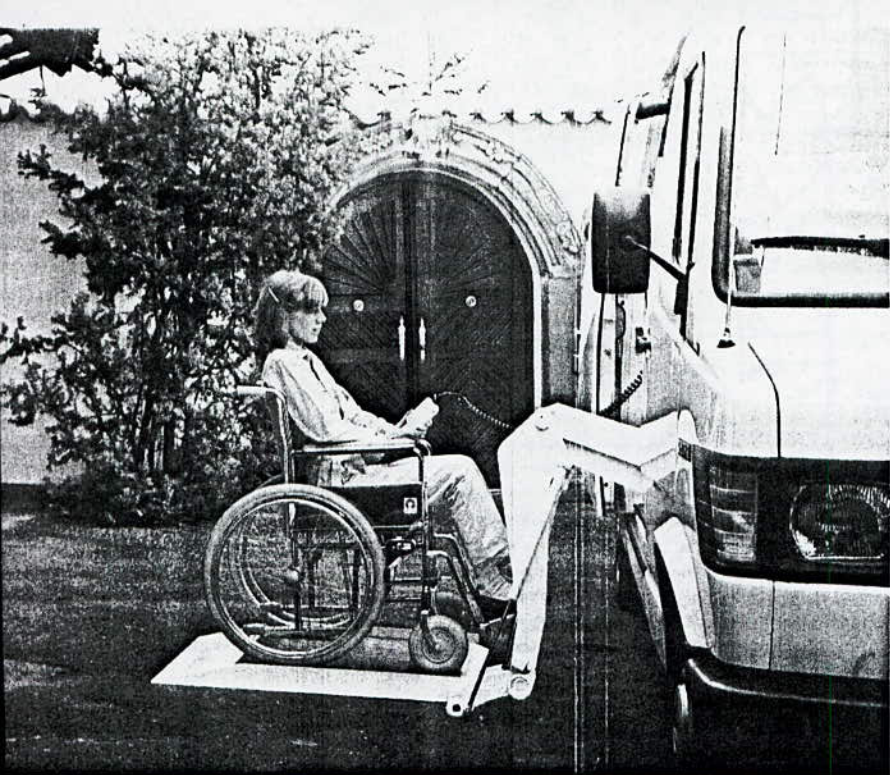
BRUSSELS

KEMPF-BENELUX  
Vuurberg 44  
B 1920 DIEGEM  
Tél. (2) 721 42 37

BONN

KEMPF Jean-Daniel  
Lehmkaulenweg 35  
D 5305 ALFTER  
GIELSDORF  
Tél. (0228) 64 55 82

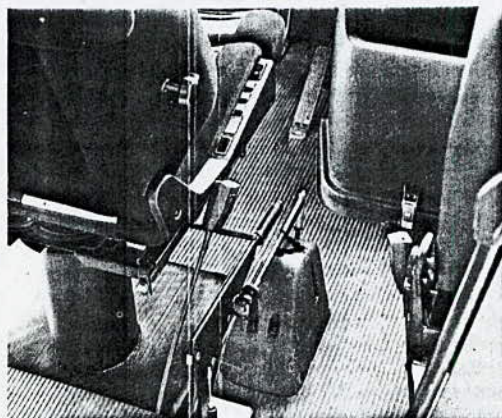
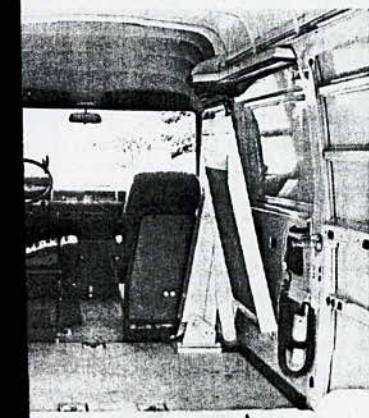
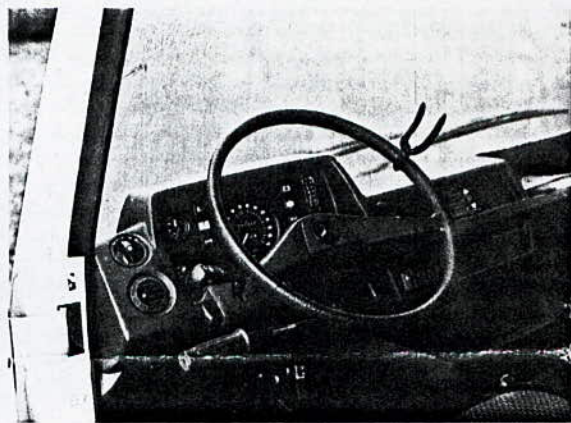
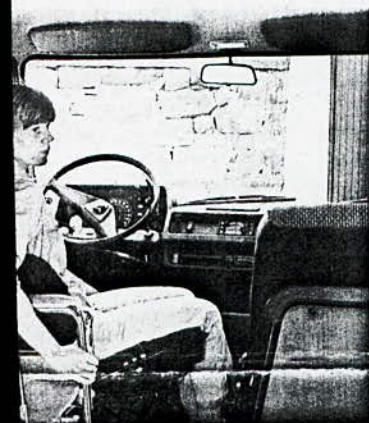




Kempf-Equipments has developed over the last thirty years a complete system which permits a quadriplegic to safely perform all driving tasks without assistance, including entering and leaving the vehicle.

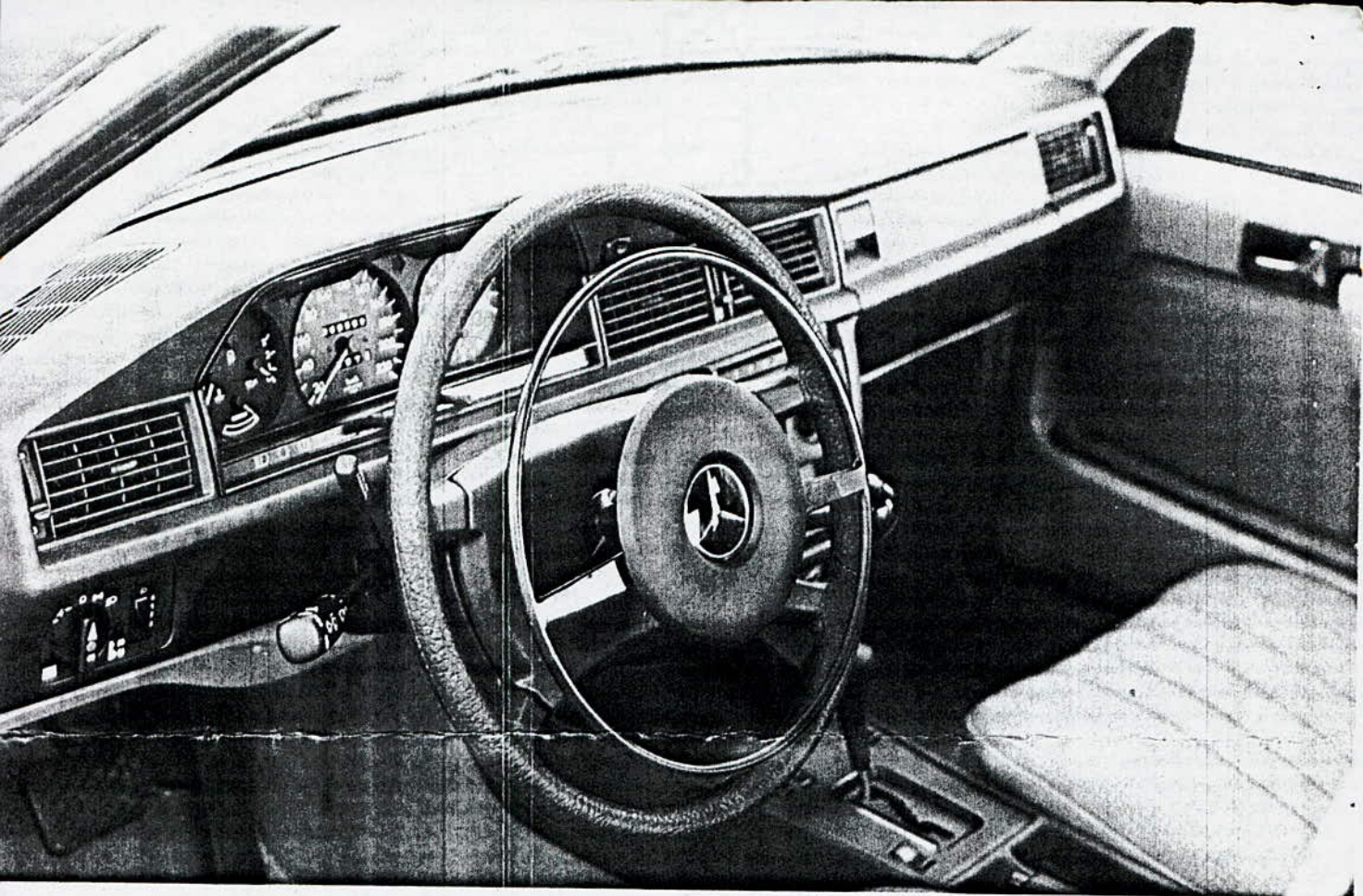
individuelle Anpassungen, Einsteigehilfen und modernste Sitztechnik für Tetraplegiker und Paraplegiker die mit Rollstuhl Unabhängigkeit wünschen.

Cet équipement procure aux conducteurs tétraplégiques indépendance et confort.



...fe independent transportation  
...the severely handicapped





## For paraplegic drivers

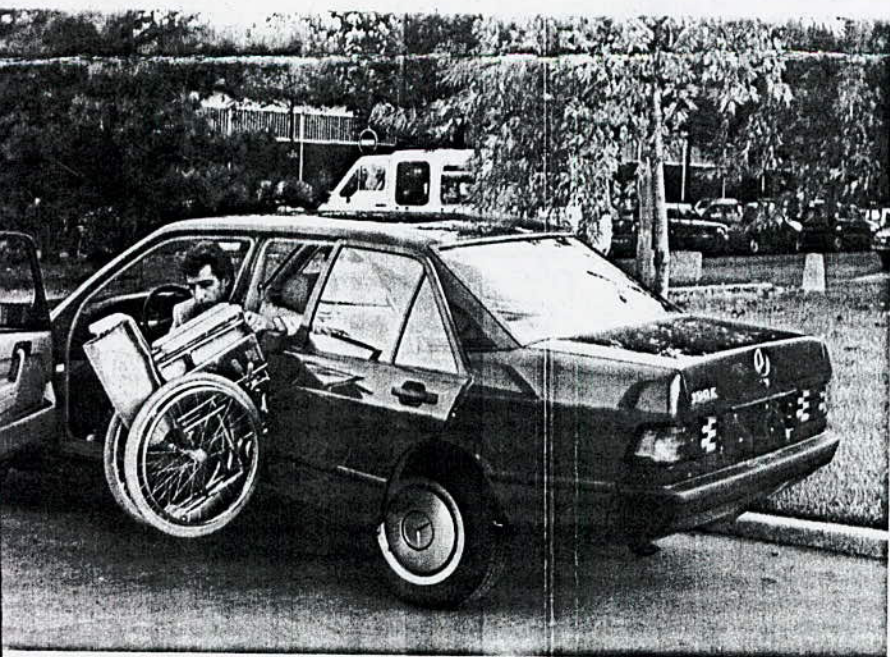
The Accelerator Ring, which revolves freely inside the steering-column, enables the driver to keep both hands on the steering-wheel and thus to drive a long distance without getting tired. The main brake is within easy reach of the right hand and is operated by a downward push.

## Für Paraplegiker

Handbedienung mit Gasring am Lenkrad bringt Sicherheit und ermüdungsfreies Fahren. Die Betriebsbremse befindet sich rechts und gewährleistet absolute Kniefreiheit.

## Pour paraplégiques

La commande manuelle du frein se trouve à la main droite. Le cercle-accelérateur sur le volant est amovible pour un conducteur valide.

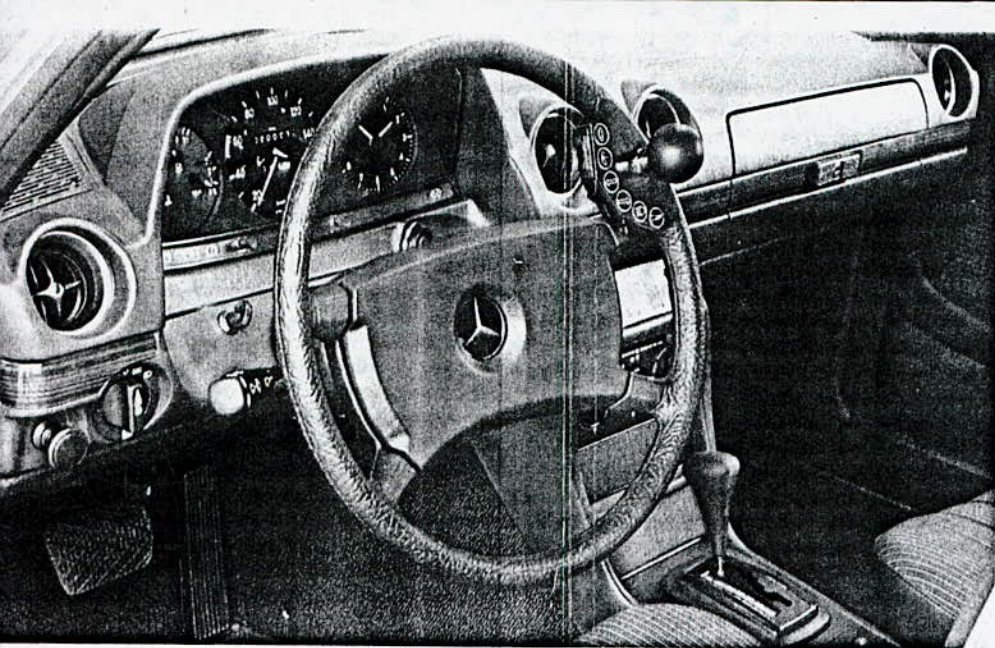


A telescopic back-door enables the paraplegic driver to pull his wheelchair into the back of his car without anybody's help. This door can either slide back and forth or, alternatively, swing on its hinges and thus open and shut in the ordinary way when the car is used by an able-bodied person.

Die Teleskoptür fährt mit Elektromotor ein und aus. In Ruhestellung schwenkt die Tür normal auf und zu.

Porte arrière coulissante permettant l'embarquement aisé du fauteuil roulant.





**For hemiplegic  
drivers**

**Für Armbehinderte**

**Pour amputés  
d'un bras  
ou hémiplégiques**

Thanks to electric switches a one-armed or hemiplegic driver can easily switch on and off the secondary controls such as indicators, lights, horn, windscreen-wipers and washers, without taking his hand off the steering-wheel.

Elektrikbedienung am Lenkrad: ohne Loslassen des Drehknopfs kann der Fahrer Blinker, Licht, Wischer und Wascher betätigen.

Le dispositif COMELDIR permet d'actionner (sans lâcher la poignée tournante) les clignotants, lumières, essuie-glace et lave-glace.

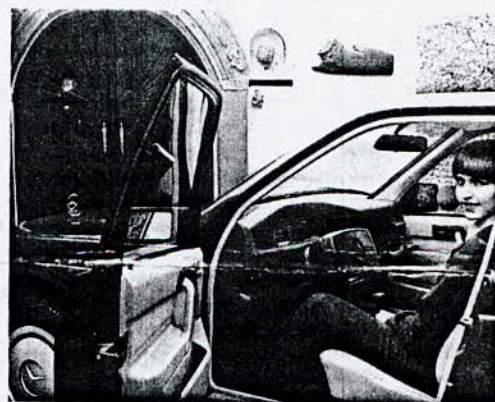
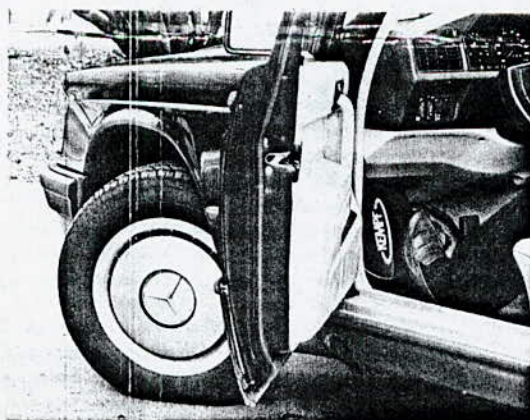
### Electronic voice-input for secondary functions:

Daimler-Benz AG and Martine Kempf have developed a Mercedes 190 E specially adapted for drivers without the use of their arms.

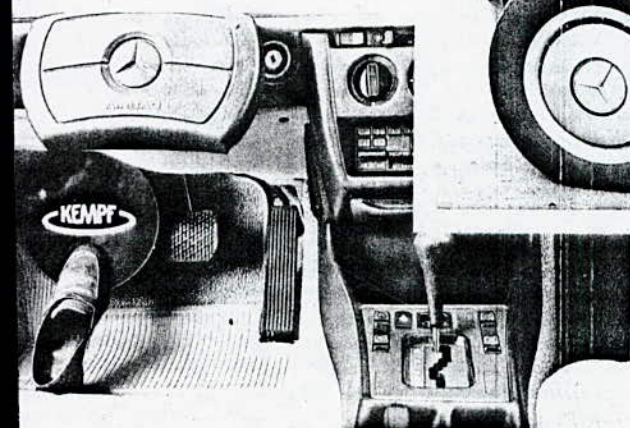
Steering, brake and accelerator are foot controlled. All other functions (up to 10 commands) are voice-controlled.

Some examples are:

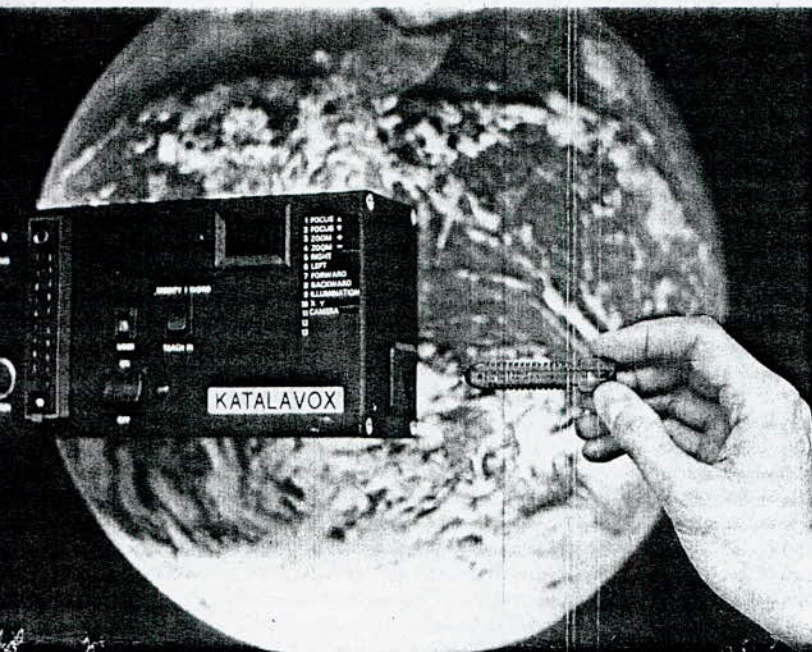
- Open and close the door
- Hand-brake
- Automatic gear shift
- Radio - Wiper
- Windshield-wiper - Seat-positioning
- Turn-indicator - Headlights



Für armloses Fahren: Fußlenkung und Sprachbedienung für elektrische Funktionen.



Pour conducteurs privés de l'usage des deux bras: direction au pied et commandes vocales pour les fonctions secondaires.



THE KEY  
WITH  
A MEMORY ...

... OPENS  
A WHOLE  
NEW WORLD

# KATALAVOX SPEECH-RECOGNITION CONTROL SYSTEM WITH A KEY, FOR SURGICAL MICROSCOPES

KATALA	VOX
▽ Modern greek	Latin ▽
TO UNDERSTAND	THE VOICE

KATALAVOX now with "The Memory Key" allows unlimited individual voice-control of any manufacturer's motorized surgical microscope system.

Each memory key contains one surgeon's vocabulary.

Programm the key once - it never forgets.

Reprogrammable to new vocabulary in less than 60 sec.

Contains an EEPROM from SEEQ Technology (Electrical Erasable Programmable Read Only Memory).

Usable with any KATALAVOX, anywhere - anytime.

- KATALAVOX is a speaker-dependant, isolated word speech-recognition system.
- Its memory or key has to be trained once to each user's voice. To accomplish this the user repeats his commands twice and Katalavox converts the microphone analog signals into digital signals and stores them as samples in memory or key.
- To understand the user's command, Katalavox compares it with the samples in memory or key. If one sample matches, the command is recognized and Katalavox activates the corresponding microscope function.

KEMPF Tomorrow's Technology Today

P.O. Box 61 103  
SUNNYVALE CA 94086-1103  
(408) 773-0219



# KATALAVOX

THE VOICE-CONTROL  
USED IN MICROSURGERY SINCE 1984

ings voice-control to : • FOCUS  
• MAGNIFICATION  
• X-Y MOVEMENTS.

YOUR motorized SURGICAL MICROSCOPE.

ill also at your command, activate :

COAXIAL ILLUMINATION  
5 mm CAMERA RELEASE  
VIDEO SYSTEM.

OT and HANDFREE CONTROL by double function proximity  
ch:

light forward movement of head activates microphone,  
KATALAVOX accepts commands.

light backward movement of head deactivates microphone  
nd stops commanded function.

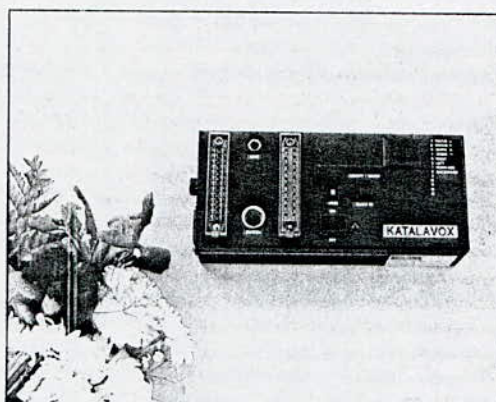
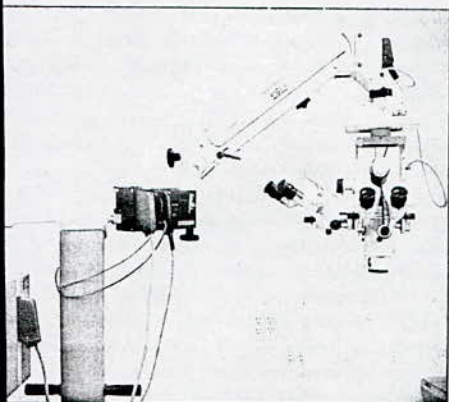
## STANDARD SPECIFICATIONS :

Programmable for up to 12 users.

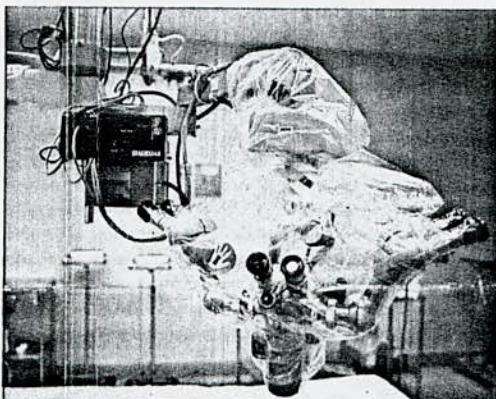
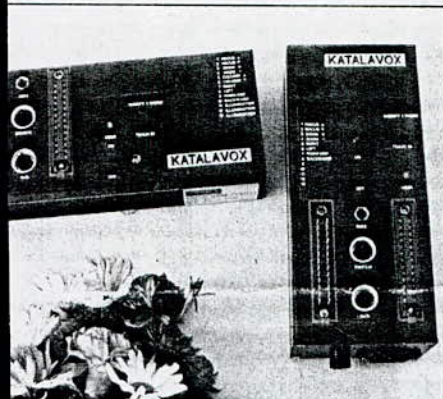
- Adaptable to a new voice in less than 60 sec.
- Switch selectable for each surgeon
- Recognizes only the selected voice
- Retains memory without battery until retrained
- Switchable from voice to hand of foot-control
- Adaptable to all major surgical microscopes
- Automatic compensation for loud background noise.

## OPTION

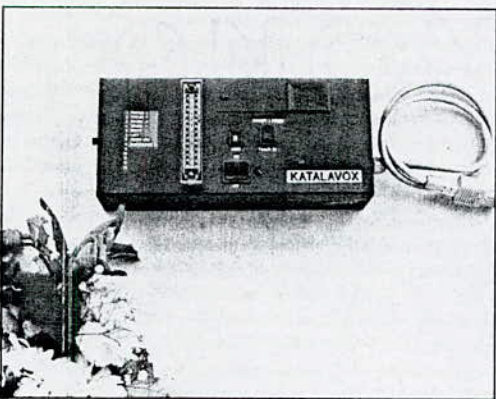
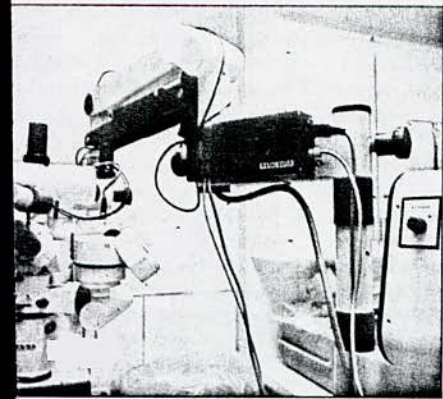
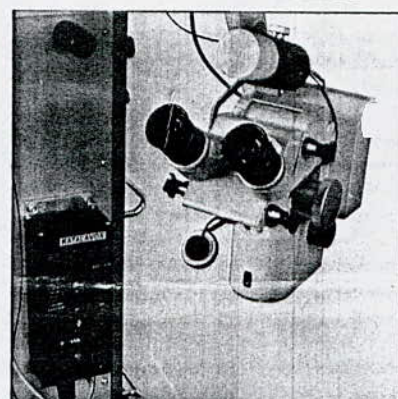
- MEMORY-KEY for each surgeon
- Programmable for unlimited number of users.



◀ MÖLLER  
COOPER VISION



ZEISS ▼



◀ WILD

SPEECH-RECOGNITION CONTROL SYSTEM :

# KATALAVOX



SOCIÉTÉ INDUSTRIELLE DE LIAISONS ÉLECTRIQUES  
DIVISION "SIGNALISATION INDUSTRIELLE"  
69, rue Ampère - 75017 PARIS - Tél.: (1) 42.67.20.60 + - Telex: 280748 F

SY-C

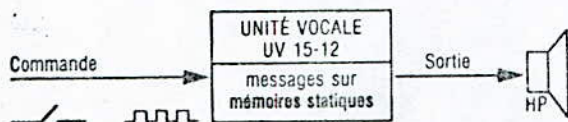
101

## synthèse de la parole

# unité vocale UV 15-12

### GÉNÉRALITÉS

Le système d'unités vocales UV 15-12 permet la diffusion automatique, sur commande extérieure, de messages parlés en toutes langues à partir d'un vocabulaire stocké sous forme numérique sur mémoires statiques. L'utilisation de mémoires statiques, à l'exclusion de tout mécanisme ou support magnétique de lecture assure une qualité et une fidélité inaltérables à la reproduction de la parole.



### AVANTAGES

- Haute fiabilité
- Réponse instantanée
- Pas d'entretien (système entièrement statique)
- Souplesse d'adaptation phonétique
- Transmission de messages suivant ordres de priorité programmés
- Autocontrôle incorporé.
- Excellente qualité vocale, respect du timbre et de la couleur mélodique de la voix.

### PRINCIPE

Utilisation du principe MIC avec compression logarithmique.

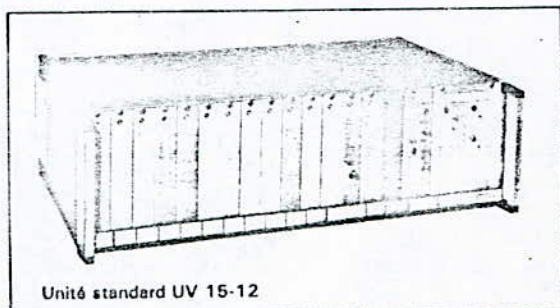
Fréquence d'échantillonnage: 8 KHz.

Stockage des informations numérisées sur REPRON.

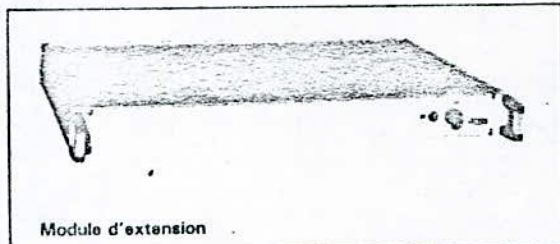
### PRÉSENTATION DES ÉLÉMENTS

L'ensemble du système UV 15 est présenté en racks ou coffrets normalisés:

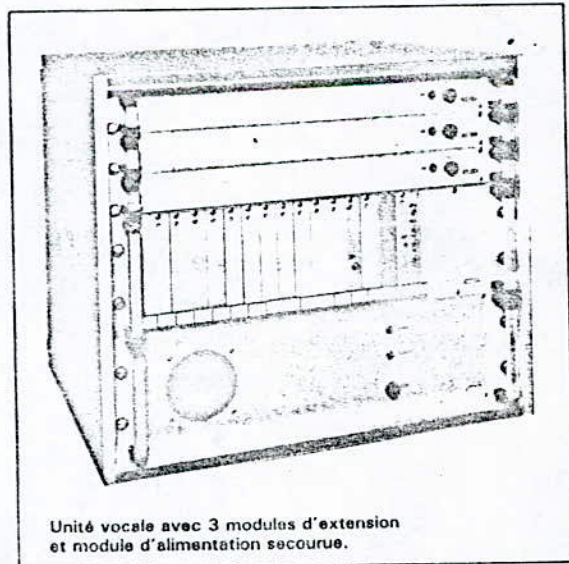
- Unité de base standard UV 15-12
- Modules d'extension
- Module d'alimentation secourue: la photographie n° 3 représente un ensemble comprenant trois modules d'extension, une unité de base standard UV 15-12 et un module d'alimentation secourue.



Unité standard UV 15-12



Module d'extension



Unité vocale avec 3 modules d'extension et module d'alimentation secourue.

#### MODULE D'EXTENSION ④

Le module d'extension est placé à l'extérieur de l'unité de base UV 15-12 et relié à cette dernière par l'intermédiaire de la carte « gestion E/S ». Il assure l'isolement galvanique des entrées.

- Présentation: rack ou coffret 1 U - 19 pouces.
- Entrées:
  - Nombre: 32 par boucle à fermeture de contacts
  - Raccordement: sur bornes à vis.
- Alimentation:
  - 220 V. C.A.: 50 Hz
  - 48 V. C.C.:
    - sur batteries
    - avec module d'alimentation secourue.
- Dimensions:
  - Version rack (profondeur: 240 mm)
    - 1 U 19 pouces 81 T en version 220 V. C.A.
    - 1 U 19 pouces 84 T en version 48 V. C.C.
  - Version coffret de table (profondeur: 240 mm)
    - 220 V. C.A.: l = 81 T h = 1 U
    - 48 V. C.C.: l = 84 T h = 1 U

**N.B.** L'unité vocale UV 15-12 peut être complétée au maximum par 8 modules d'extension, ce qui représente 256 entrées du type boucle.

#### MODULE D'ALIMENTATION SECOURUE

- Présentation: rack ou coffret 3 U.
- Alimentation: 220 V. C.A. - 50 Hz.
- Sortie: 48 V. C.C.
- Batterie d'accumulateurs intérieure: étanche, plomb gélifié.
- Autonomie: environ 30 minutes.
- Dimensions:
  - Version rack (profondeur: 240 mm)
    - 3 U 19 pouces 84 T.
  - Version coffret de table (profondeur: 240 mm)
    - l = 84 T h = 3 U

**N.B.** Un haut-parleur de contrôle peut être incorporé à ce module.

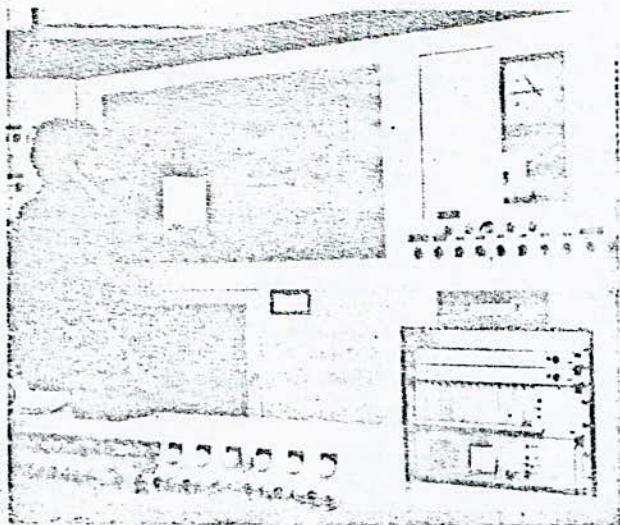
#### QUELQUES APPLICATIONS

- **Sécurité** - Diffusion d'alarmes, message de danger, consignes de sécurité: raffineries, plate-formes de forage, centrales de production d'énergie, immeubles de grande hauteur, magasins, mines, tunnels, espaces souterrains.
- **Exploitation** - Diffusion d'incidents de fabrication, séquences de procédure: aciéries, laminiers, ponts-roulants, ascenseurs.
- **Information** -
  - Diffusion d'annonces: gares, stations de sports d'hiver, magasins, aéroports, autobus, wagons de voyageurs, télépancartage ...
  - Diffusion de messages par téléphone ...
- **Télésurveillance** - Réseaux de distribution d'énergie électrique, d'eau, de gaz...

#### NOS RÉFÉRENCES

- S.N.C.F.: Annonces aux voyageurs  
Paris Austerlitz  
Paris Boulevard Victor.
- E.D.F.: Surveillance de sous-stations  
Centre de distribution de St Mandé.
- HAUTS FOURNEAUX DE LA CHIERS  
Surveillance de sécurités et process.
- SOGITEC  
Simulateurs de vol.
- MÉTROPOLITAIN: Annonces aux voyageurs  
Paris.
- SOLMER  
Messages d'évacuation en halls de manutention.
- SOFRESID  
Centralisation d'alarmes en IGH et diffusion de messages de sécurité.
- OGER-DEF  
Diffusion automatique de messages d'évacuation et de sécurité (Arabie Saoudite).
- SILEC  
Répétiteurs piétons parlants (Signalisation urbaine)  
Système «SYMPHONIE».

#### UTILISATION D'UNE UNITÉ VOCALE SILEC SUR CENTRALE D'ALARME EN I.G.H.



Ci-contre, unité vocale SILEC intégrée dans le poste de surveillance et sécurité d'un Immeuble de Grande Hauteur.

- Diffusion automatique des alarmes:
  - une unité vocale de base UV 15-12
  - deux modules d'extension
  - un module d'alimentation 48 V C.C. secourue.
- Système commandé par boucles.
- Alarmes diffusées en local au poste de surveillance et sur centrale d'émission « Bip-Bip » pour information directe aux personnes responsables.
- Mémorisation des appels et diffusion d'alarmes par ordre d'importance.

## CONFIGURATION DU SYSTÈME

Le système est constitué d'une unité vocale de base standard UV 15-12 à microprocesseur ① pouvant fonctionner seule ou bien associée à un ou plusieurs modules d'adaptation assurant une liaison rationnelle avec la nature de l'équipement de l'utilisateur.

Deux types de modules d'adaptation :

- Passif : assure simplement l'isolement galvanique des Entrées/Sorties ②.
- Actif : assure la gestion des Entrées/Sorties (files d'attente, réédition cyclique, hiérarchisation, etc...) ③.

Dans ce dernier cas, l'accès au système peut se faire par un bus codé 10 bits (1024 entrées) ou par boucle (256 entrées) via le module d'extension ④ à 32 entrées.

L'interface assure également l'adaptation d'une ou plusieurs sorties analogiques, ainsi que la surveillance du lecteur (Autocontrôle).

Interfaces :

- Carte d'isolement galvanique ②
- Carte de gestion E/S ③
- Module d'extension ④.

Le système est normalement alimenté sur secteur en courant alternatif ou continu. Un module « alimentation secourue » est proposé en option.

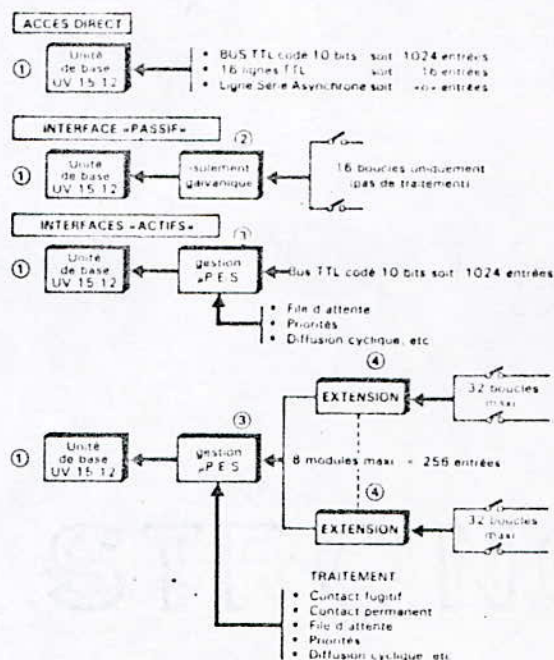
## DESCRIPTION - CARACTÉRISTIQUES

### UNITÉ VOCALE DE BASE ① TYPE UV 15-12

- Présentation: rack 3 U 19 pouces ou coffret de table.
- Capacité mémoire maximum: 96 ou 192 secondes (en fonction de la taille du dictionnaire). Mémoires sur 12 cartes débrochables au format européen.
- Capacité parole: fonction de la finesse de l'analyse des termes du dictionnaire.
- Mode d'accès: niveau logique TTL - 1024 entrées (accessibles sur câble plat).
- Sorties analogiques:
  - 8 sorties OdB sur 600  $\Omega$
  - plus une voie de contrôle amplifiée 5 W sur 4  $\Omega$ .
- Bande passante: 300 à 3000 Hz.
- Correction: par égaliseur 4 octaves.
- Alimentation:
  - 220 V. C.A. - 50 Hz
  - 48 V. C.C.:
    - sur batteries (20 à 60 V. C.C.)
    - sur module d'alimentation secourue.
- Dimensions:
  - Version rack (profondeur: 240 mm)
    - 3 U 19 pouces 81 T en version 220 V. C.A.
    - 3 U 19 pouces 84 T en version 48 V. C.C.
  - Version coffret de table (profondeur: 240 mm)
    - 220 V. C.A. l = 81 T h = 3 U
    - 48 V. C.C. l = 84 T h = 3 U

### CARTE D'ISOLEMENT GALVANIQUE ②

Ce circuit, incorporé à l'unité vocale de base UV 15-12 a pour fonction d'adapter les informations d'entrée sous forme de contacts, en niveau logique directement assimilable par l'unité vocale de base.



### CARTE DE GESTION ENTRÉE/SORTIES ③

Ce circuit, incorporé à l'unité de base UV 15-12 assure la communication avec l'utilisateur :

- soit directement si cette dernière est assurée par un bus TTL Codé (10 bits maximum)
- soit par l'intermédiaire de modules d'extension ④ si l'accès est du type « boucle », contacts libres de potentiel.

#### Fonctions logiques

- **File d'attente** - Mémorisation des appels intervenant en cours de diffusion, ces derniers étant restitués à l'unité vocale de base dans l'ordre chronologique de leur apparition.
- **Priorité** - Appels enregistrés affectés d'un numéro de priorité (8 maximum). Si cette fonction est validée, l'interface assure la hiérarchisation des appels et restitue à l'unité vocale de base les appels prioritaires, puis les non prioritaires dans l'ordre chronologique de leur apparition.
- **Priorité relative ou absolue** - Un appel prioritaire est validé après la dernière diffusion en cours (priorité relative) ou interrompt la diffusion en cours pour être immédiatement validé (priorité absolue).
- **Rediffusion cyclique** - Si un appel est maintenu il est diffusé cycliquement après un temps « t » défini par programme.

#### Fonctions analogiques

- **Sorties OdB** - Le circuit de gestion est équipé de quatre sorties OdB sur transformateur 600  $\Omega$ . Il est possible d'affecter un numéro d'annonce à une sortie analogique. Dans ce cas, le circuit de gestion assure automatiquement le multiplexage de la diffusion (une voie parmi quatre). Si cette fonction n'est pas programmée, la sortie n° 1 est prise en défaut.
- **Autocontrôle** - En phase de veille, l'unité vocale de base est soumise à un balayage cyclique des annonces. Si une rupture de modulation est détectée, (durée supérieure à 1 seconde) le système fournit à l'utilisateur une alarme sous forme d'un contact libre de potentiel.
- **Message en cours « MEC »** - Un contact fermé, libre de potentiel est fourni à l'utilisateur pendant toute la durée d'une diffusion (ex.: commutation de sonorisation).

## DEFINITION DU SYSTEME

Un système parlant est constitué d'un lecteur à micro-processeur (UV 15/12 ou UV 15/5) associé éventuellement à un ou plusieurs modules d'adaptation assurant une liaison rationnelle à l'équipement utilisateur.

Les modules d'adaptation peuvent être du type "passif" s'ils n'ont qu'à assurer l'isolement galvanique des Entrées/Sorties du système ou "Actif" si le système doit assurer une gestion des E/S (File d'attente, réédition cyclique, hiérarchisation, etc ...). Dans ce dernier cas, l'accès au système peut se faire par un bus codé 10 bits (1024 entrées) ou par boucle (256 entrées) via module d'extension 32E. L'interface assure également l'adaptation d'une ou plusieurs sorties analogiques, ainsi que la surveillance du lecteur (Autocontrôle).

Les possibilités respectives de ces différentes versions sont décrites dans la section "INTERFACES".

# UNITES VOCALES UV.15/12 UV.15/5

## DESCRIPTION GENERALE

### Présentation

- . UV 15/12 Version standard 220 V  
Coffret de table où bac à cartes en 3U x 81T \*
  - . UV 15/12 Version 48 V =  
Coffret de table où bac à cartes en 3U x 84T
  - . UV 15/5 Version 220 V uniquement  
Coffret de table en 3U x 40T
- \* 1T = 5,08 mm.

### Modularité

Chaque version est prévue pour recevoir des cartes enfi-  
chables au format "Europe" 100 x 160.

#### UV 15/12

- . 12 Modules mémoire (64K ou 128K)
- . 1 Module Microprocesseur I
- . 1 Module Microprocesseur II
- . 1 Module Conversion "A"
- . 1 Module Amplification

#### UV 15/5

- . 5 Modules mémoire (64K ou 128K)
- . 1 Module Microprocesseur I
- . 1 Module Microprocesseur II
- . 1 Module Conversion "A"



## FONCTIONS ANALOGIQUES

- Sorties 0 db. Le circuit de gestion est équipé de quatre sorties 0 db sur transformateur 600 $\Omega$  /600 $\Omega$  : Il est possible d'affecter un numéro d'annonce à une sortie analogique, dans ce cas le circuit de gestion assure automatiquement le multiplexage de la diffusion (une voie parmi quatre). Si cette fonction n'est pas programmée la sortie 1 est prise par défaut.
- Autocontrôle.  
En phase de veille, le lecteur est soumis à un balayage cycliqué des annonces. Si une rupture de modulation est détectée (durée supérieure à une seconde) le système fournit à l'utilisateur un contact libre de potentiel "DEF" (lecteur en défaut).
- Message en cours. "MEC"  
Fonction standard. Un contact fermé libre de potentiel est fourni à l'utilisateur pendant toute la durée d'une diffusion.  
  
Cas particulier : l'information "MEC" peut être délivrée en pré-déclenchement ou temporisée, dans ce cas, c'est le programme lecteur qui gère cette fonction.
- Haut-parleur supplémentaire. (HPS)  
(Voir module amplification)

## CARACTERISTIQUES PRINCIPALES

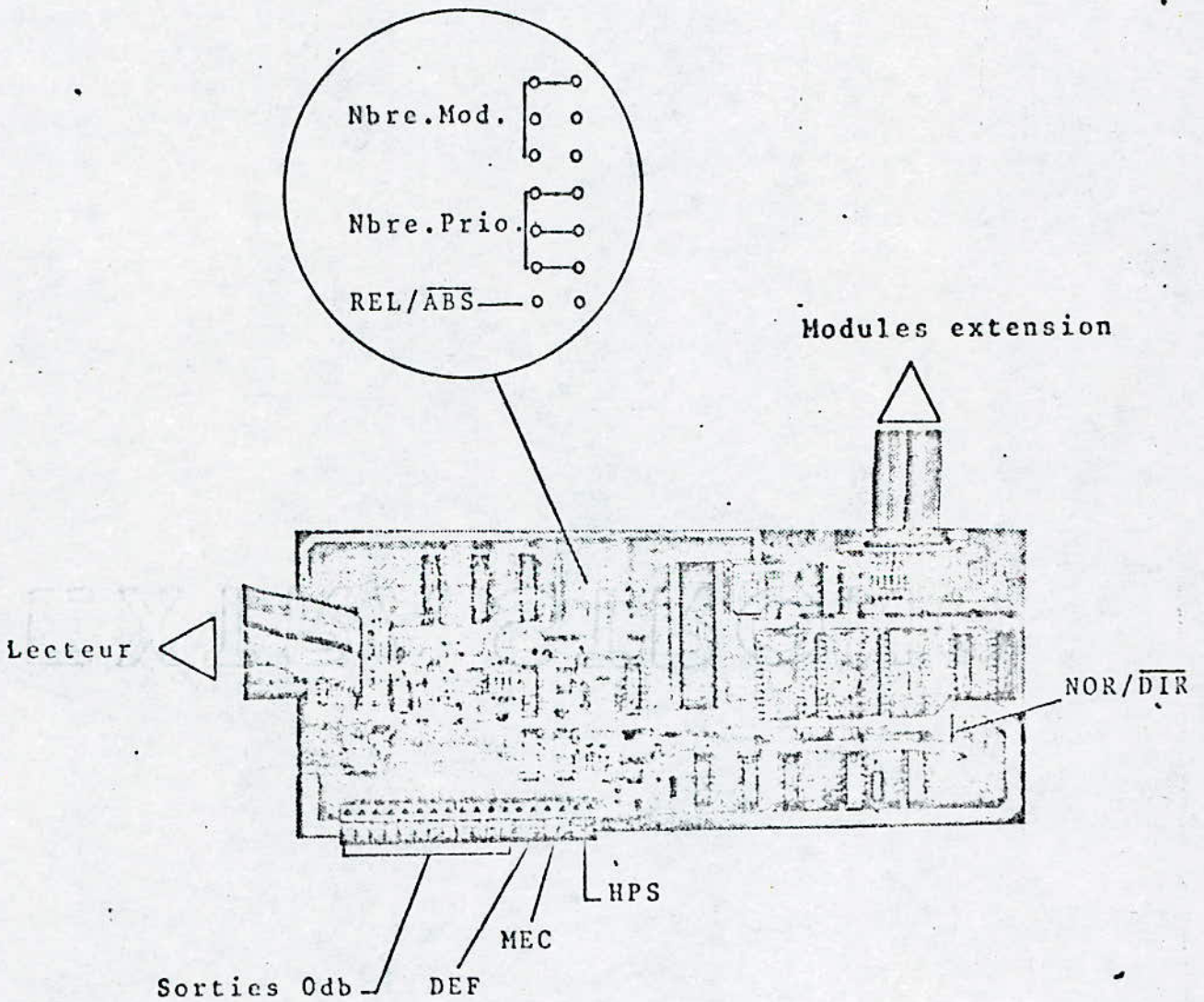
### FONCTIONS LOGIQUES

- File d'attente. L'interface assure la mémorisation des appels intervenant en cours de diffusion, ces derniers étant restitués au lecteur dans l'ordre chronologique de leur apparition.
- Notion de priorité. Les appels enregistrés peuvent être affectés d'un numéro de priorité (8 maximum) si cette fonction est validée, l'interface assure la hiérarchisation des appels et restitue au lecteur dans l'ordre qui leur a été attribué : les appels prioritaires puis les appels non prioritaires dans l'ordre chronologique de leur apparition.
- Priorité relative ou absolue. Suivant le mode de programmation adopté, un appel de type prioritaire sera validé après la dernière diffusion en cours (priorité relative) ou interrompt la diffusion en cours pour être immédiatement validé (priorité absolue).
- Rediffusion cyclique. Chaque appel est mémorisé en même temps que son origine (Heure-Minute). Si cet appel est maintenu il sera diffusé cycliquement après un temps "T" défini par programme.
- Type d'accès. Le mode d'exploration des entrées est défini par le sélecteur de fonction NOR/DIR. La fonction "Normal" correspondant à un accès par modules d'extension, dans ce cas il convient également de définir le nombre de modules effectivement raccordés (sélecteur Nbr. Mod.) ce mode d'exploration permet l'accès de 256 entrées sous forme de boucle. La fonction "Direct" permet le raccordement direct d'un bus TTL codé, de "n" bits (10 maximum) soit l'accès de 1024 entrées.

# GESTION INTERFACE

Ce circuit, solidaire du lecteur, assure la communication avec l'utilisation soit directement si cette dernière est assurée par un bus TTL codé (10 bits maximum) ou par l'intermédiaire de "modules d'extension" si l'accès est du type "boucle" (contacts libres de potentiel).

Le programme de gestion étant polyvalent, l'accès aux différentes fonctions est validée par programmation Hardware (voir fig. XII).



# MEMOIRE

Deux versions - MP01 - 64K (fig. I)  
- MP01 - 128K.

Equipement en fonction du volume d'annonces à traiter.

L'unité vocale peut recevoir indifféremment l'une ou l'autre version.

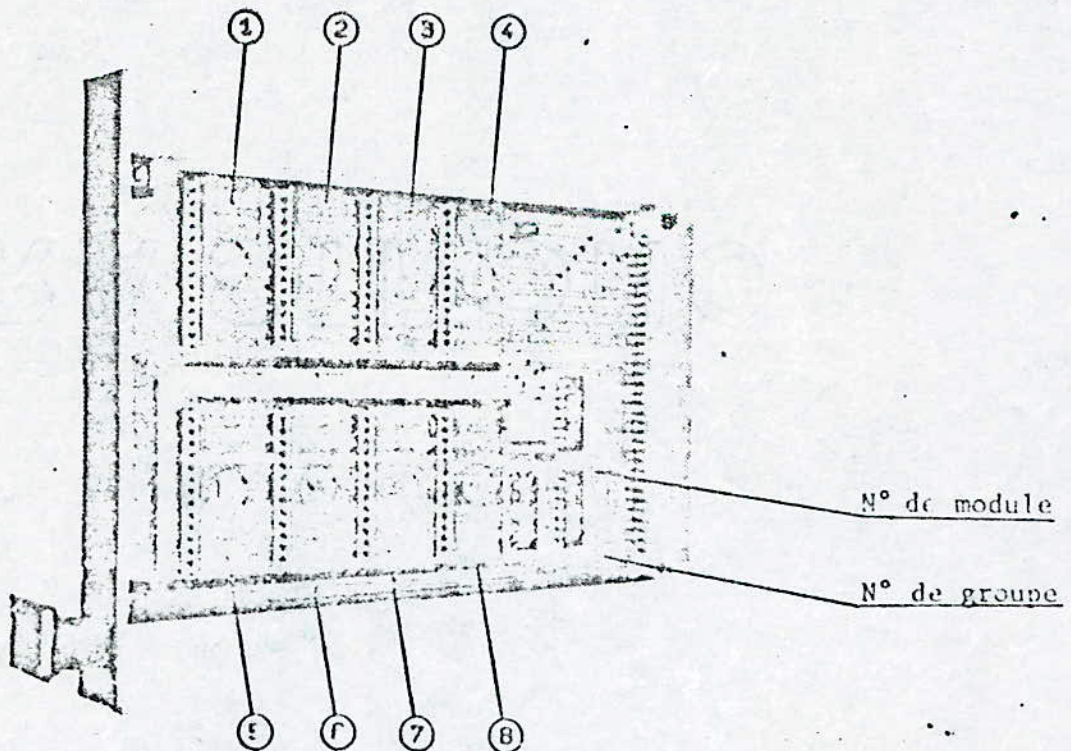
capacité par module	MP01 64K	8 boitiers	2764	=	64 Koctets	=	8 secondes
	MP01 128K	8 boitiers	27128	=	128 Koctets	=	16 secondes

Le champ mémoire est constitué de deux groupes de modules:

Groupe 1 = modules de 1 à 8

Groupe 2 = modules de 9 à 12

Chaque module est banalisé et reçoit (par positionnement de cavaliers) son n° de groupe ainsi que son n° de module lors de l'implantation mémoire (voir fig. I)



# MICROPROCESSEUR.1

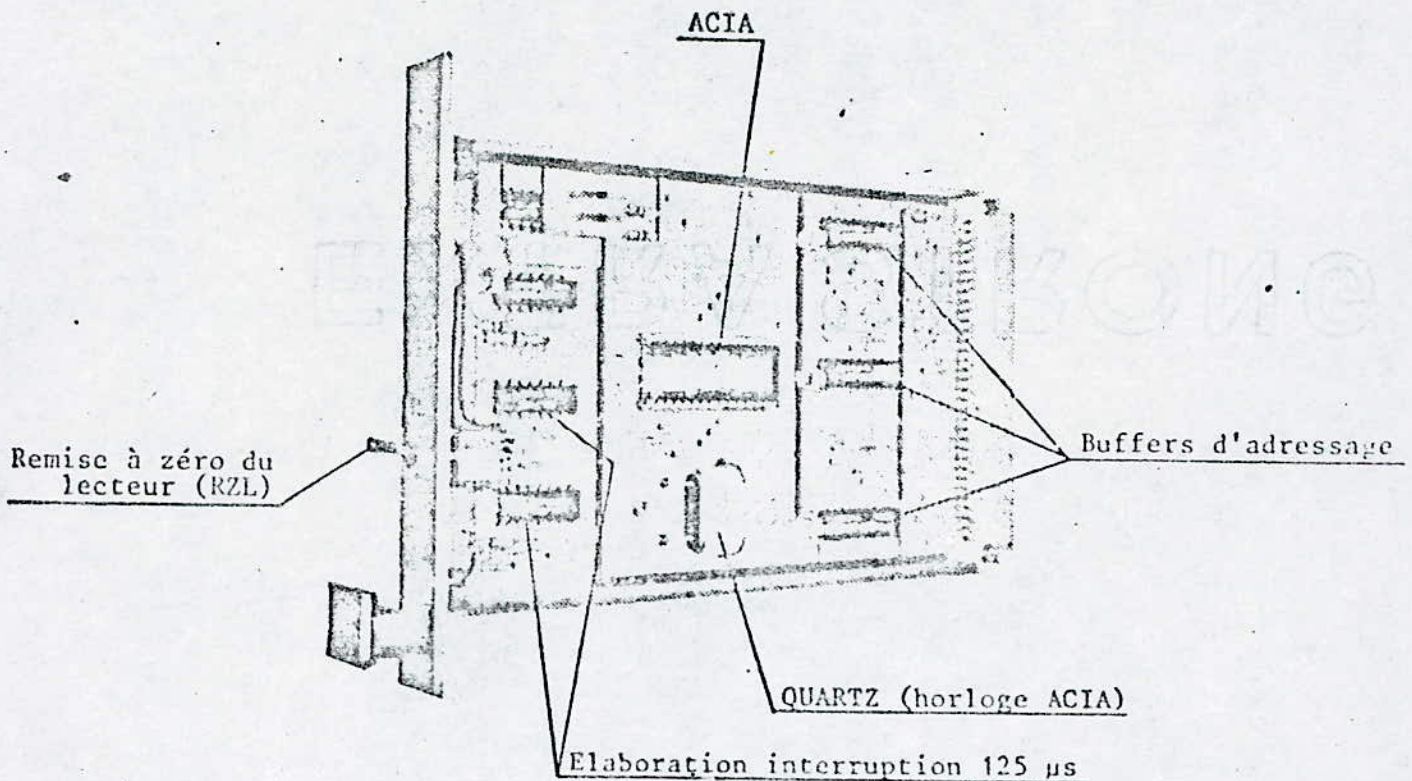
# MICROPROCESSEUR.2

La gestion du Lecteur assurée par le Microprocesseur et ses boîtiers périphériques ne pouvant être regroupée sur un seul module, a été scindée en deux modules MICRO I et MICRO II assurant respectivement les fonctions suivantes :

## MICROPROCESSEUR I

- Générateur d'interruption 125  $\mu$ s (fréquence de restitution)
- Buffers d'adressage du dictionnaire
- Remise à zéro du Lecteur (commutation RZL)
- ACIA (Asynchronous Communication Interface Adapter)

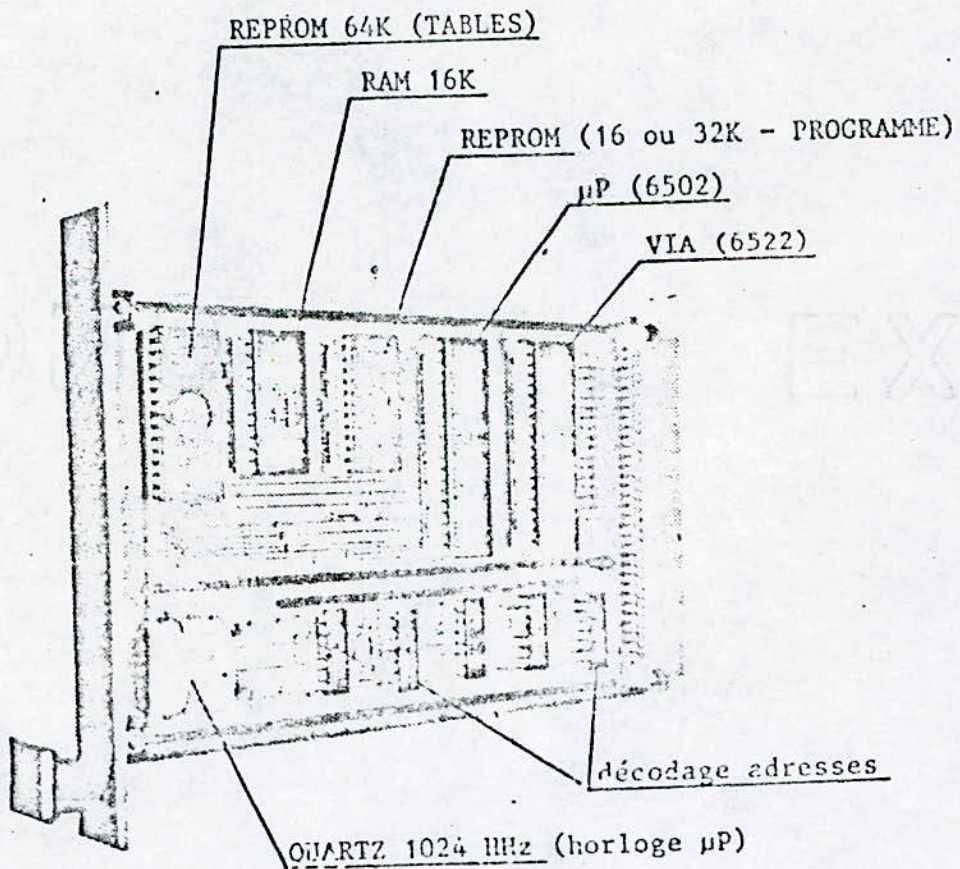
Le boîtier ACIA délivre toutes les informations nécessaires à la communication en mode série asynchrone, l'accès à ces lignes est situé sur le fond de panier (voir distribution FIG. III)



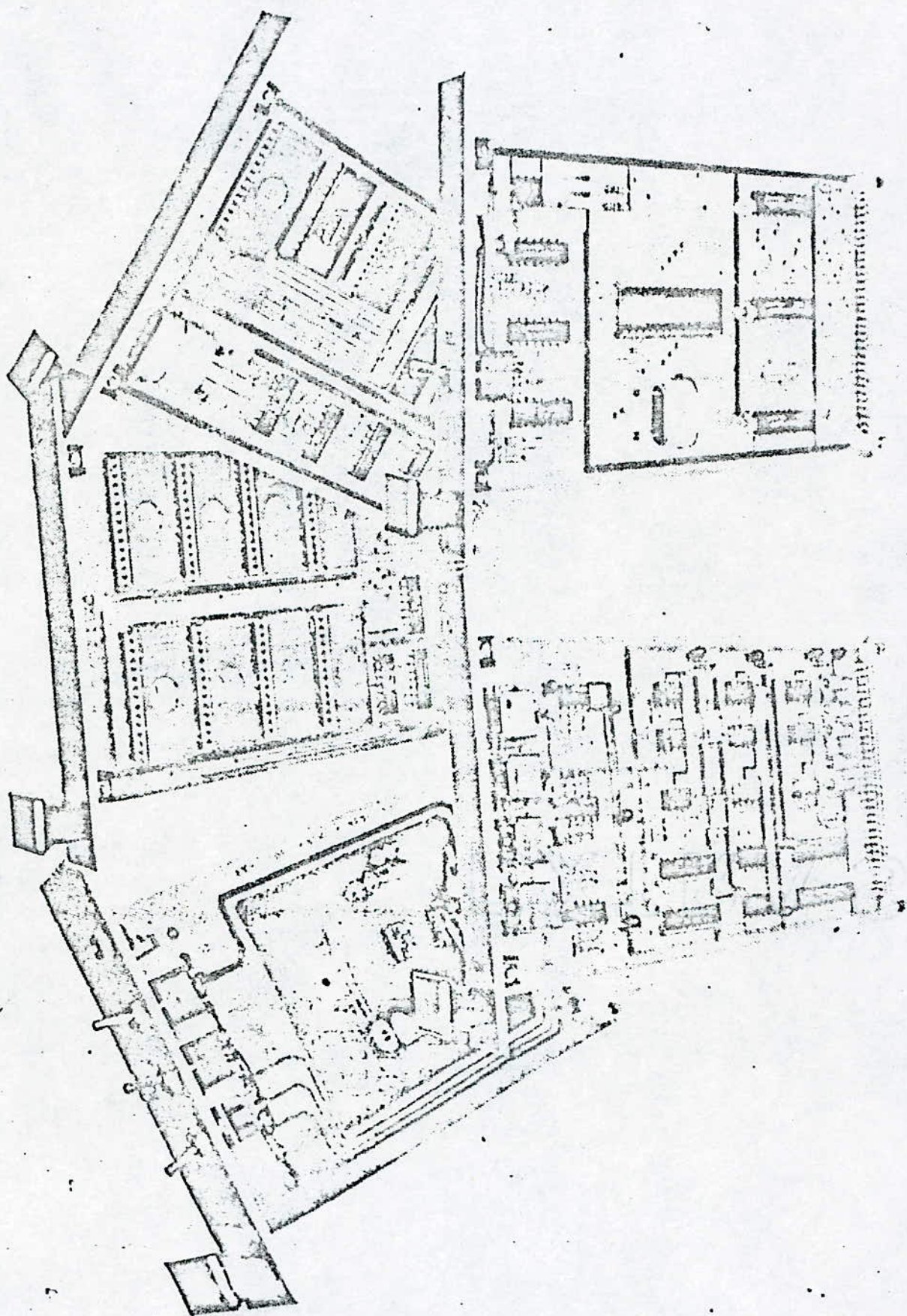
## MICROPROCESSEUR II

Ce module reçoit :

- le microprocesseur (Rockwell 6502)
- une mémoire REPR0M (16K ou 32k) contenant le programme de gestion du lecteur
- une mémoire REPR0M 64K contenant les tables d'assemblage (lois d'association des termes du dictionnaire)
- une mémoire RAM pour le stockage des informations en cours de traitement
- le VIA (Versatile Interface Adapter) assurant la liaison en mode parallèle des informations - Entrées - Sorties niveau logique
- les décodeurs d'adresses des boîtiers périphériques du microprocesseur.



# MODULES



# CONFIGURATIONS DE SYSTEMES

## ACCES DIRECT

UV 15/12  
ou  
UV 15/5

- . BUS TTL codé 10 bits soit : 1024 entrées
- . 16 lignes TTL " : 16 entrées
- . Ligne Série Asynchrone " : "n" entrées

## INTERFACE "PASSIF"

UV 15/12  
ou  
UV 15/5

isolement galvanique



16 boucles uniquement  
(pas de traitement)

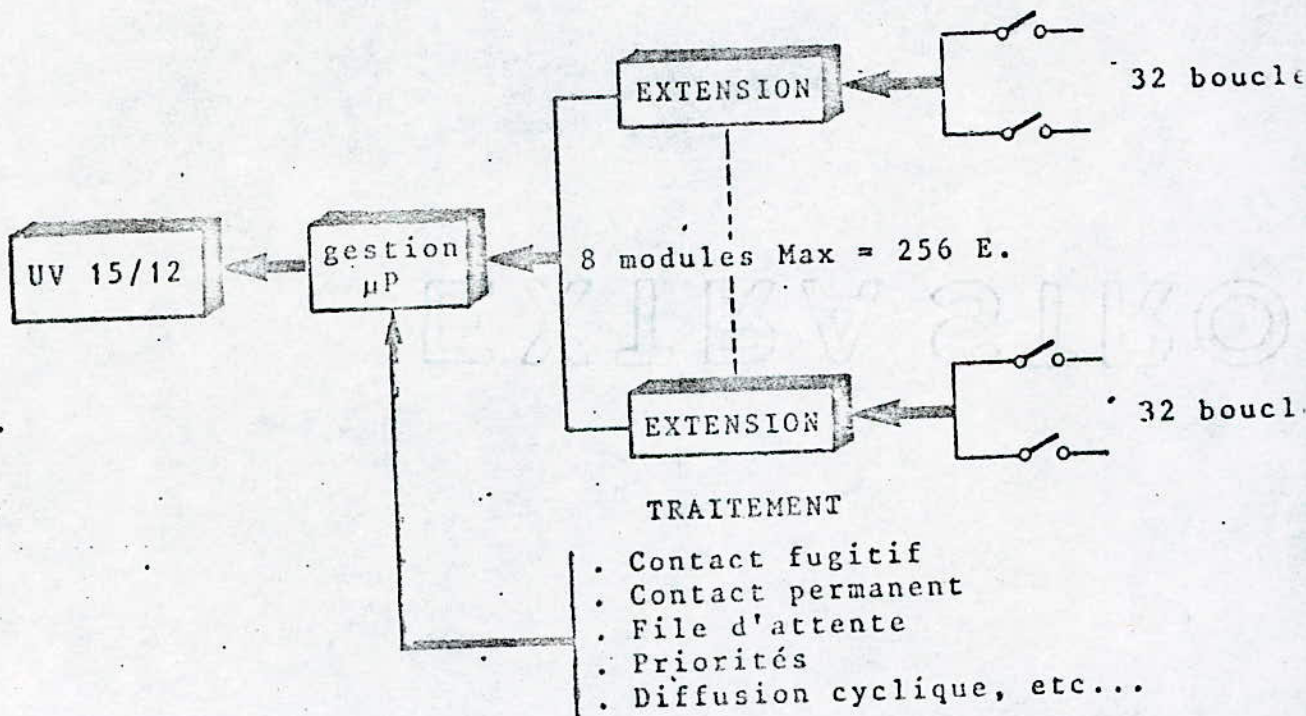
## INTERFACES "ACTIFS"

UV 15/12

gestion  $\mu$ P

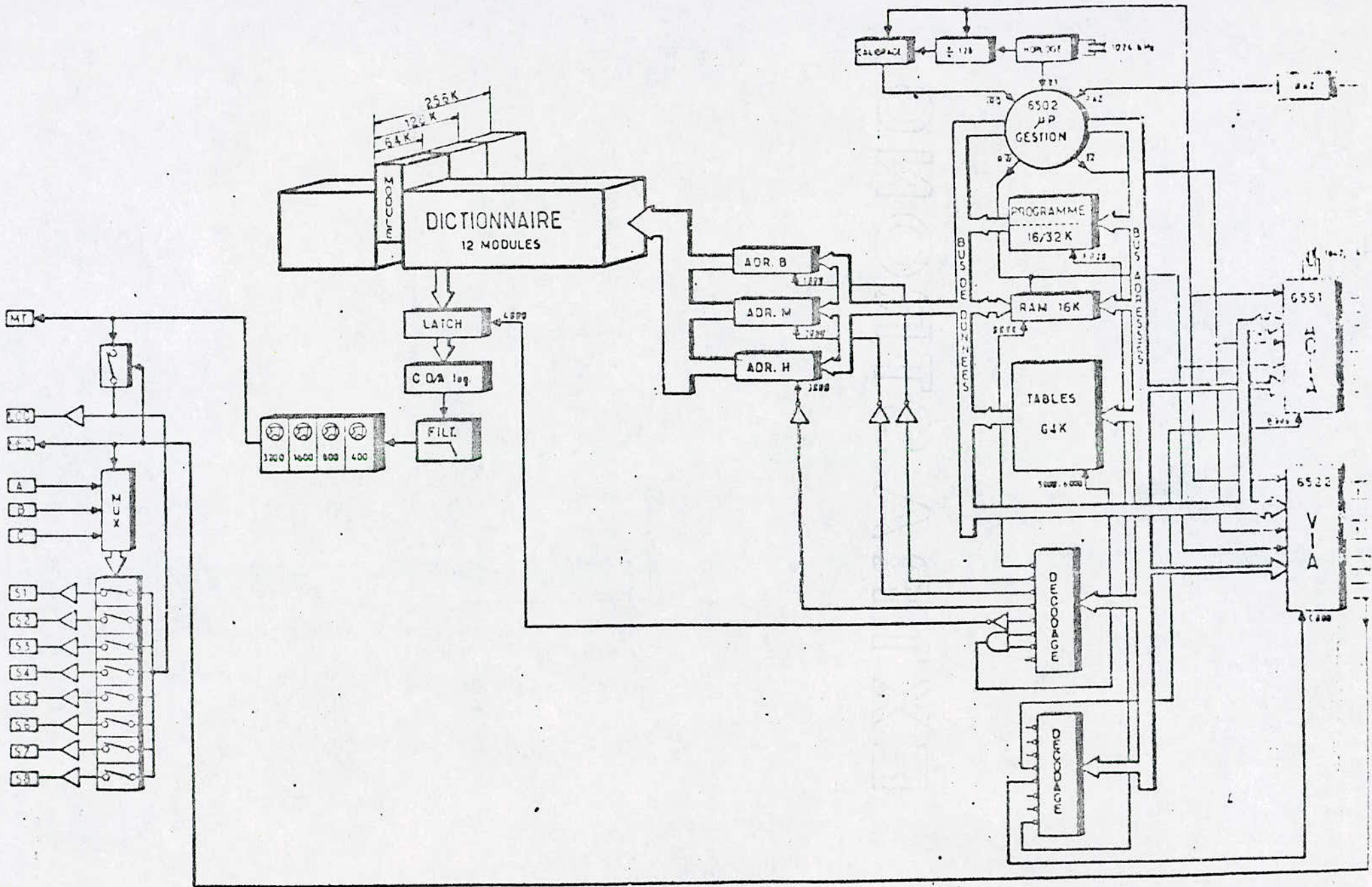
BUS TTL codé 10 bits soit : 1024 E.

- . File d'attente
- . Priorités
- . Diffusion cyclique, etc...





SYNOPTIQUE UV 15/12



\* BIBLIOGRAPHIE \*

---

OUVRAGES :

---

- 1 - J. MAX : TOME I ET II  
METHODES ET TECHNIQUES DE TRAITEMENT DU SIGNAL ET APPLICATION  
AUX MESURES . / edi MASSON 1981  
COTE : TOME I -----> 62.52  
TOME II -----> 621.391
  
- 2 - J. LIFERMANN  
LES METHODES RAPIDES DE TRANSFORMATION DU SIGNAL . / edi MASSON 1980  
COTE : 517.3
  
- 3 - J. GUIBERT  
LA PAROLE : COMPREHENSION ET SYNTHESE PAR LES ORDINATEURS . / 1979  
COTE : 681.3
  
- 4 - M. BELLANGER  
TRAITEMENT NUMERIQUE DU SIGNAL . / edi MASSON 1984  
COTE : 621.391
  
- 5 - J. FLORINE  
AUTOMATISMES A SEQUENCES ET COMMANDES NUMERIQUES . / edi DUNOD 1969  
COTE : 62.52
  
- 6 - S. LEIBSON  
MANUEL DES INTERFACES . / edi Mc GRAW HILL 1984  
COTE : 681.3
  
- 7 - N. AUMJEAUX  
PRATIQUE DE L'ELECTRONIQUE : COMPARATEURS - FILTRES ACTIFS  
COTE 621.375
  
- 8 - J. AUVRAY : TOME I ET II  
ELECTRONIQUE DES SIGNAUX ECHANTILLONNES ET NUMERIQUES . / edi DUNOD  
ELECTRONIQUE DES SIGNAUX ANALOGIQUES . / edi DUNOD
  
- 9 - J. MORLAND  
SCHEMAS ELECTRONIQUES

10 - M.FERRITTI ET F.CINARE

SYNTHESE.RECONNAISSANCE DE LA PAROLE . / edi tests 1983

11 - L.R.RABINER ET R.W.SHAFER

DIGITAL PROCESSING OF SPEECH SIGNALS . / PRENTICE HALL 1960

REVUES :

1 - TOUTE L'ELECTRONIQUE N 489 DEC 83

2 - TOUTE L'ELECTRONIQUE N 490 JAN 84

3 - TOUTE L'ELECTRONIQUE N 496 AOUT - SEP 84

4 - SCIENCES ET TECHNIQUES N 2 HORS SERIE 1894

"L'ORDINATEUR PREND LA PAROLE "

NOTES TECHNIQUES :

1 - L.R.RABINER ET M.J.CHENG  
A.E.ROSENBERG ET C.A. GONEGAL

A COMPARATIVE PERFORMANCE STUDY OF SEVERAL PITCH DETECTION ALGORITHMS

IEEE TRANS ACOUST SPEECH-SIGNAL-PROCESSING  
VOL ASSP 24 PP 399-418 /OCT 1976

2 - M.J.ROSS ET H.L.SCHAFFER ET A.COHEN  
R.FREUDBERG ET H.J.MAULEY

AVERAGE MAGNITUDE DIFFERENCE FUNCTION - PITCH EXTRACTOR

IEEE TRANS ACOUST SPEECH-SIGNAL-PROCESSING  
VOL ASSP 22 PP 353-362 /AUG 1974

3 - J.J.MATRAS

ACOUSTIQUE -ELECTRO ACOUSTIQUE -PERCEPTION DES SONS

TECHNIQUES DE L'INGENIEUR . / E 2602 1979

4 - J.Y.JEGOU ET P.COMBECURE  
C.GAGNOULET ET R.DESCOUT

TRAITEMENT DE LA PAROLE

NT/LAA/TSS/168 JUI 83

5 - M.COVRAT

SERAPHINE : RECONNAISSANCE DE LA PAROLE CONTINUE PAR DES METHODES  
GLOBALES

NT/LAA//TSS/79 AOUT 81

6 - I.EL-MALLAWANY

SYNTHESE ET EVALUATIONS COMRARATIVES DE MELOGRAPHES NUMERIQUES

NT/DAS/SST/10 AOUT 78

7 - C.CESSOU

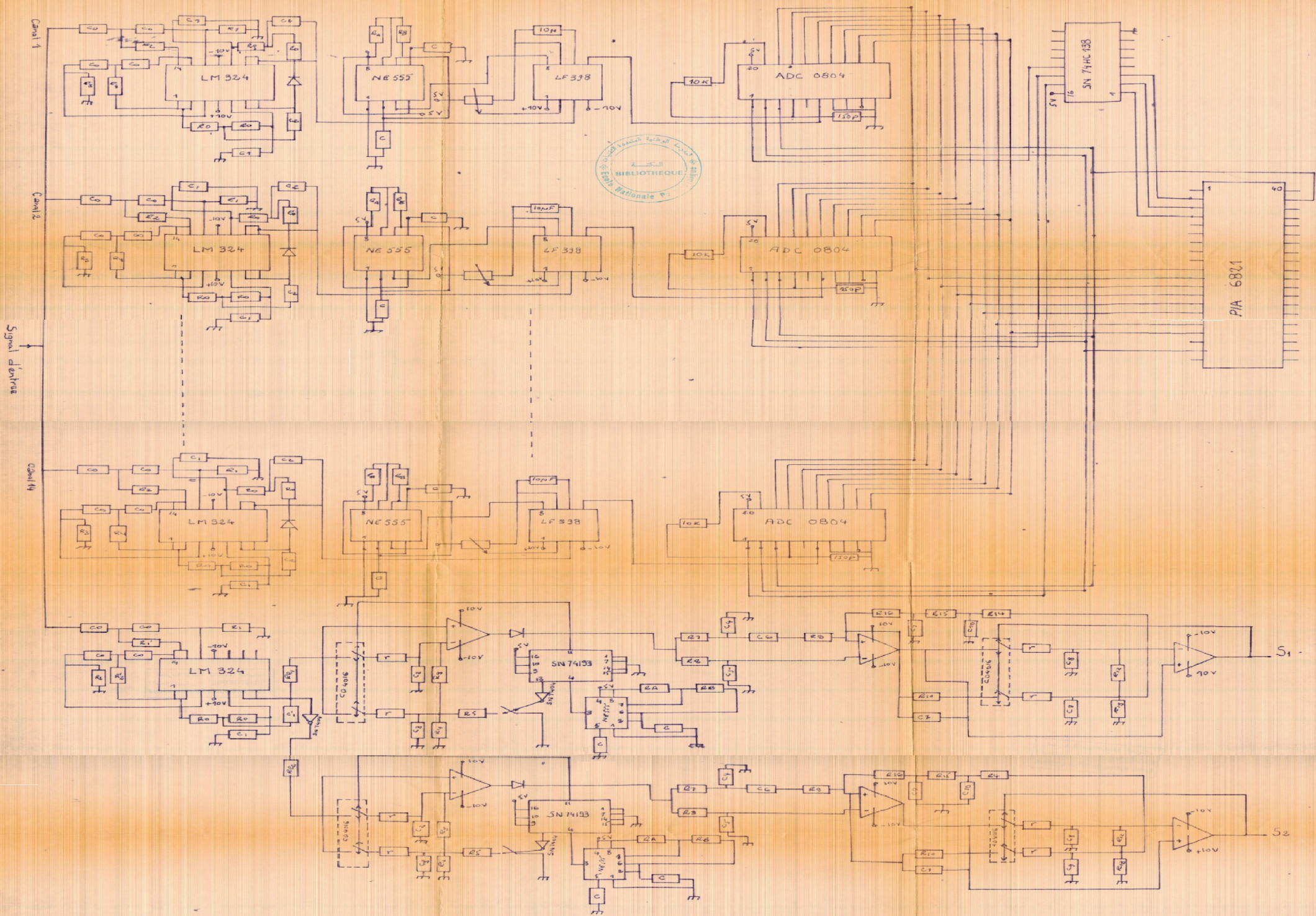
SIMULATION D'UN VOCODEUR A CANAUX. SYNTHSE DE LA PAROLE

RP/DAS/SST/15 OCT 78

8 - I.EL-MALAWANY ET A.COMBET

DETECTEUR NUMERIQUE DE VOISEMENT

NT/CEI/CSE/70 NOV 76



ECOLE

NATIONALE

POLYTECHNIQUE

SCHEMA  
DETECTEUR  
ET

ELECTRIQUE  
DE  
DE

DU  
MELODIE  
L'ANALYSEUR

EXECUTE PAR : M<sup>rs</sup>

K. SEMMAR ET S. MEKAOU