

# Introduction à l'audio numérique

Christine Sénac – christine.senac@irit.fr

Bibliographie:

- Haton, Cerisara and al: « Reconnaissance automatique de la parole », Dunod, 2006
- Rabiner, Juang: « Fundamentals of speech recognition » Prentice Hall Signal, 1993
- Calliope: « La parole et son traitement automatique », Mason, 1989
- cours de Régine André-Obrecht....



## Plan du cours

- I. A la découverte du signal de parole : un signal aléatoire
- II. Analyse du signal
- III. Reconnaissance de la parole
  - C'est quoi? message, locuteur, ...
  - L'importance de la dimension temporelle
  - Reconnaissance (probabilistique, discriminative)

# I. A la découverte du signal de parole : un signal aléatoire

- I.1 Communication Homme-Machine
- I.2 Rappels sur le signal audio
- I.3 Production de la parole
- I.4 Perception des sons
- I.5 Description acoustique des sons de parole
- I.6 Classification des sons français
- I.7 Exercice: Lecture de sonagrammes

UPS - M1 Informatique - Christine Sénac

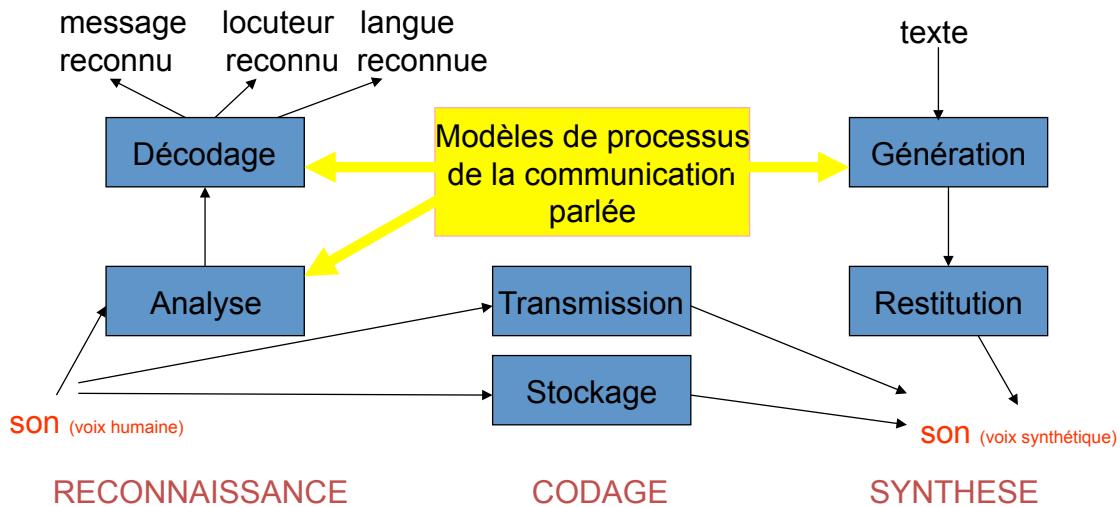
## Communication Humaine

Modes de communication	Critères
✓ Geste	- rapidité
✓ Dessin	- liberté de la vue, du mouvement, des mains
✓ Écriture	- permanence de l'information
✓ Parole	- facilité d'accès, secret

=> Efficacité de la parole

# I.1 Communication Homme-Machine (1/2)

## Les différents domaines du Traitement Automatique de la Parole



# I.1 Communication Homme-Machine (2/2)

## Les domaines scientifiques

### Production de la parole

- ✓ Mécanique (mouvement des articulateurs)
- ✓ Propagation des ondes (génération de l'effet sonore)
- ✓ Physiologie (organes de la production)

### Perception de la parole

- ✓ Physiologie (organes du système auditif)
- ✓ Psycho-acoustique (perception des sons)
- ✓ Processus cognitifs (acquisition des informations sur l'environnement)

### Linguistique

- ✓ phonétique (articulatoire), phonologie (phonèmes), lexique, sémantique, pragmatique (étude de l'usage de la langue en situation de communication et des conditions de celle-ci)

### Acoustique

- ✓ méthodes de traitement du signal vocal: analyse spectrale, statistiques et probabilités, théorie de l'information, ...

**Interdisciplinarité nécessaire...pour modéliser les processus de la communication parlée**

## I.2 Quelques rappels sur le signal Audio (1/4)

- ✓ Le son est le résultat de la vibration de l'air à partir de sa source d'émission

- ✓ Il est défini par deux caractéristiques:

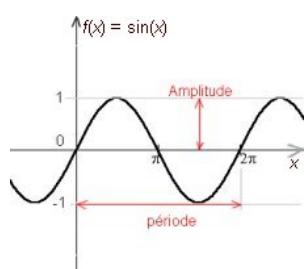
- ✓ Sa **fréquence**

- ✓ Son **intensité**

UPS - M1 Informatique - Christine SéDAC

## I.2 Quelques rappels sur le signal Audio (2/4)

### Son pur

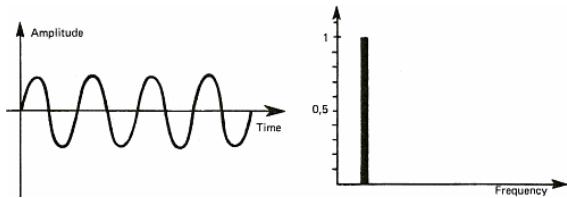


- n'existe pas dans la nature
- correspond à une **seule fréquence**
- représenté par une sinusoïde : périodique, **période** =  $T$  (durée entre 2 crêtes et exprimée en secondes)
- défini par la **fréquence F** de cette sinusoïde ( $F =$  nombre de périodes en une seconde, exprimée en Hertz (**Hz**))
- $F < 20\text{Hz} \Rightarrow$  infrasons;
- $F$  faible ( $20\text{ Hz} - 200\text{Hz}$ )  $\Rightarrow$  son grave;
- $F$  moyenne ( $200\text{ Hz} - 2000\text{ Hz}$ );
- $F$  élevée ( $2000\text{ Hz} - 20000\text{ Hz}$ )  $\Rightarrow$  son aigu ;
- $F > 20000\text{ Hz} \Rightarrow$  ultrasons.

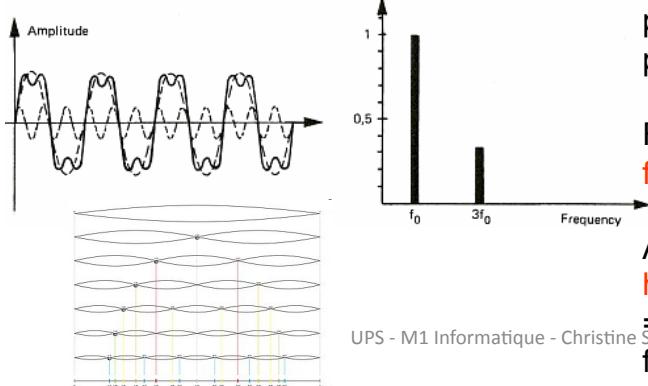
UPS - M1 Informatique - Christine SéDAC

## I.2 Quelques rappels sur le signal Audio (3/4)

### ✓ Son pur



### ✓ Son complexe



Son complexe = son composé de plusieurs sinusoïdes (si périodique)

Fréquence la plus basse = **fréquence fondamentale**

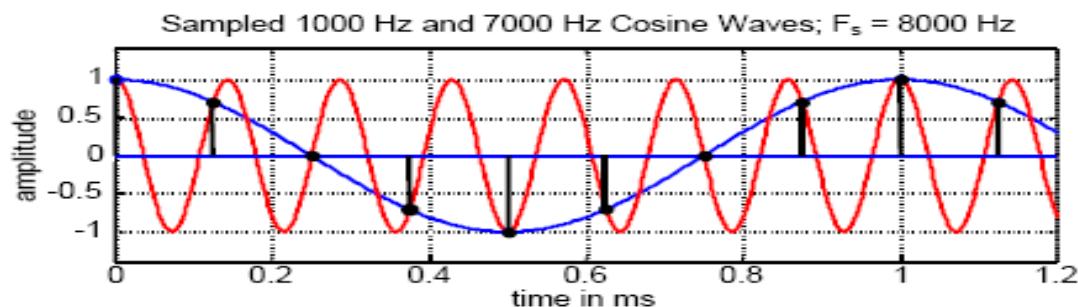
Autres fréquences = **harmoniques**  
= multiples de la fondamentale

UPS - M1 Informatique - Christine Sénaç

## I.2 Quelques rappels sur le signal Audio (4/4) Conversion Analogique/Digitale

### *Le théorème de Nyquist-Shannon pour l'échantillonnage:*

La fréquence d'échantillonnage doit être 2 fois supérieure à la fréquence la plus haute du signal (sinon perte d'information)

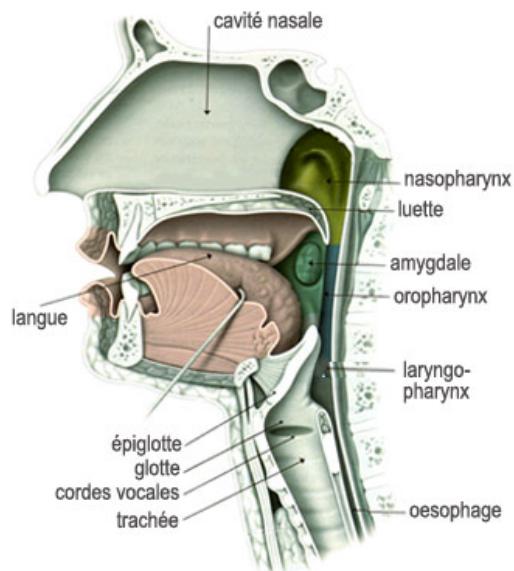


- A bandlimited signal can be reconstructed exactly from samples taken with sampling frequency

$$\frac{1}{T} = F_s \geq 2f_{\max}$$

UPS - M1 Informatique - Christine Sénaç

## I.3 Production de la parole (1/7)



Appareil vocal humain

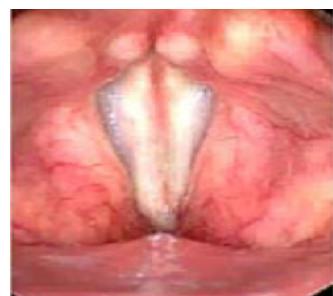
UPS - M1 Informatique - Christine SéDAC

## I.3 Production de la parole (2/7) les cordes vocales

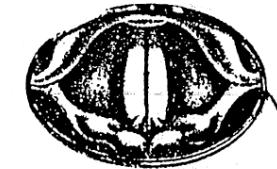
Respiration



Inspiration profonde



Phonation



Voix chuchotée



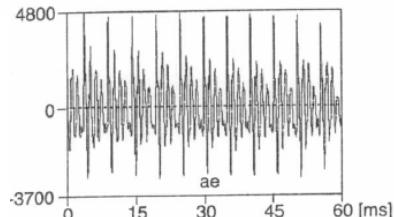
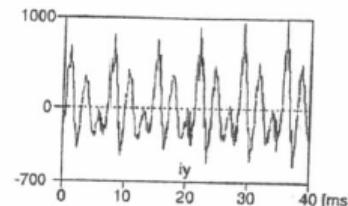
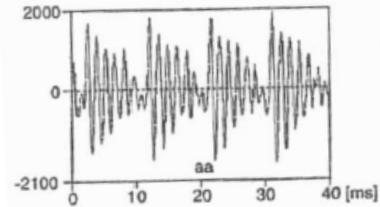
UPS - M1 Informatique - Christine SéDAC

## I.3 Production de la parole (3/7)

### la fréquence fondamentale $F_0$

- La période fondamentale  $T_0 = 1$  cycle d'ouverture/fermeture des cordes vocales
- La fréquence fondamentale  $F_0 = 1/T_0$

	$F_0$ ave (Hz)	$F_0$ min (Hz)	$F_0$ max (Hz)
Men	125	80	200
Women	225	150	350
Children	300	200	500



Démo:

[http://www.oddcast.com/home/demos/tts/tts\\_example.php](http://www.oddcast.com/home/demos/tts/tts_example.php)(faire varier le pitch)

UPS - M1 Informatique - Christine Sénac

## I.3 Production de la parole (4/7)

### Très grande variabilité du signal de parole

#### ✓ Facteurs intra-locuteur

- Facteurs articulatoires:
  - Fréquence de vibration des cordes vocales (=hauteur des sons: + vibration rapide, + son aigu)
  - Force d'expiration (intensité des sons -> accentuation / accent tonique)  
hauteur des sons + intensité => intonation prosodique (par ex: ?, !)
  - Vitesse d'articulation (durée des sons -> rythme / allongement)
  - Variabilité articulatoire (tous les gestes articulatoires ne sont pas réalisés à l'identique)
  - Coarticulation (interaction des gestes les uns sur les autres -> imprécision motrice, inertie mécanique, loi du moindre effort)
- Facteurs paralinguistiques:
  - Type de phonation = mode de vibration des cordes vocales (Voix chuchotée, normale, criée ...)
  - Type d'articulation (articulation normale/soignée/relâchée...)

#### ✓ Facteurs externes

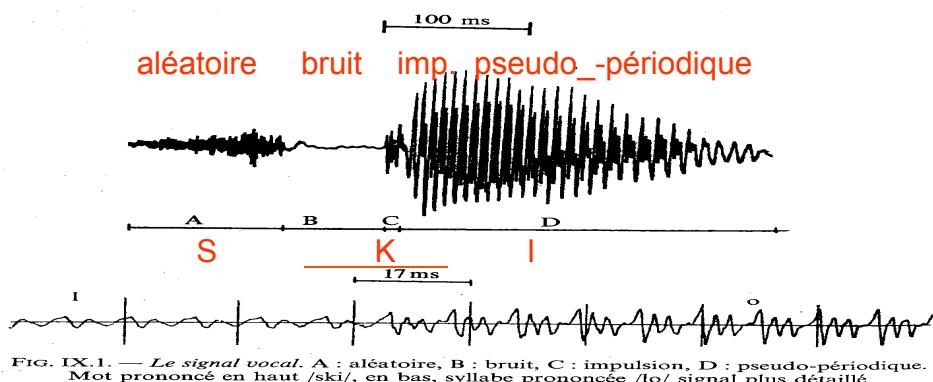
- Facteurs inter-locuteurs:
  - différences physiologiques, socio-culturelles, psychologiques...
- L'environnement:
  - Le bruit ambiant, le microphone, le canal de transmission ...
  - La musique superposée, les voix superposées ...
  - La tâche (stress,...)

UPS - M1 Informatique - Christine Sénac

## I.3 Production de la parole (5/7)

Très grande variabilité du signal de parole

✓ **Signal continu, d'énergie finie, non stationnaire :** structure complexe et variable dans le temps, tantôt pseudopériodique, tantôt aléatoire, tantôt impulsionnelle. Cette structure reflète l'organisation temporelle des gestes de production et sur l'onde vocale apparaissent quelques caractéristiques de la source et du conduit vocal.



✓ **Bandé passante de l'oreille :** [20Hz – 20000Hz]

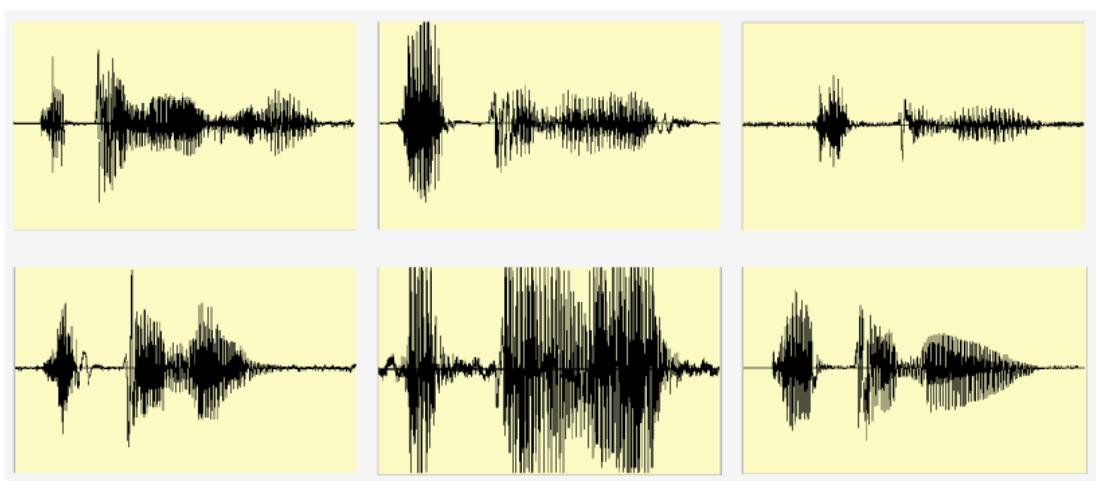
✓ **Information linguistique :** [50Hz – 4000Hz]

✓ **Canal téléphonique :** [300 – 3300Hz] => perte d'information

UPS - M1 Informatique - Christine Sénac

## I.3 Production de la parole (6/7)

Très grande variabilité du signal de parole



- 6 locuteurs différents prononçant le même mot : « appeler »

UPS - M1 Informatique - Christine Sénac

## I.3 Production de la parole (7/7)

Très grande variabilité du signal de parole

Mais ....

Bertil Malmberg(1913-1994), phonéticien(1950), linguiste(1969):

“Les phonéticiens disent: Ce serait pourtant une grave erreur que de se représenter les voyelles et les consonnes comme des unités fixes et invariables, alignées... comme les perles d'un chapelet”

“Les phonologues disent: En réalité, si nous avions à notre disposition un nombre illimité d'unités infiniment variables, aucune communication ne serait possible”

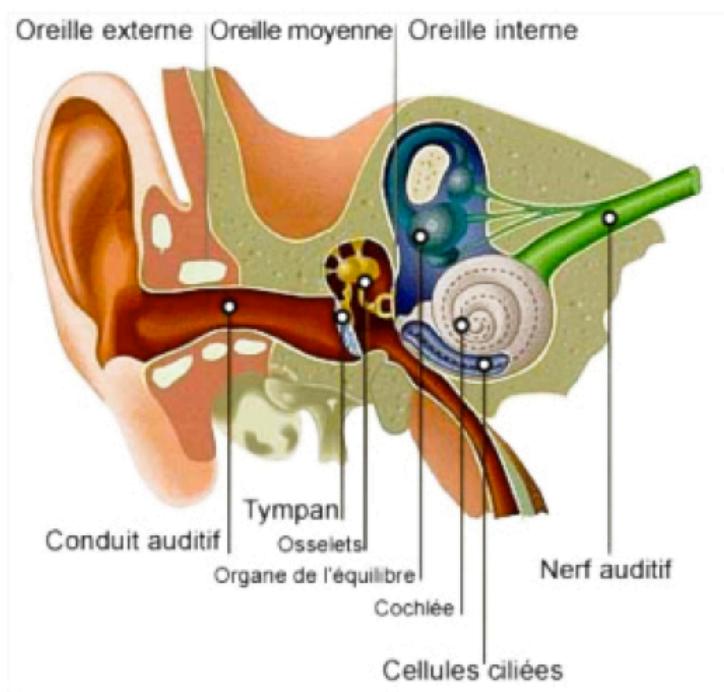
D'où

- **Recherche d'invariants (catégorisation des sons)**
- **Domaine pluridisciplinaire.**

UPS - M1 Informatique - Christine Sénac

## I.4 Perception des sons (1/3)

Variation de pression → Mouvement mécanique → Décharge électrique



UPS - M1 Informatique - Christine Sénac

## I.4 Perception des sons (2/3)

### Mesure de l'intensité acoustique :

- L'intensité acoustique c'est **l'amplitude de la vibration** qui se mesure en décibels
- une intensité acoustique  $I$  est convertie en  $X$  **décibels** par:

$$X \text{ dB} = 20 \log_{10}(I/I_0)$$

où  $I_0$  est le seuil d'audibilité de l'oreille ( $I_0 = 10^{-16}$  Watt/cm<sup>2</sup> pour un son de fréquence 1000Hz)

$I_0$  le seuil d'audibilité correspond à 0 dB

**La perception de la hauteur des sons est** proportionnelle au logarithme de la fréquence fondamentale => notion d'octave

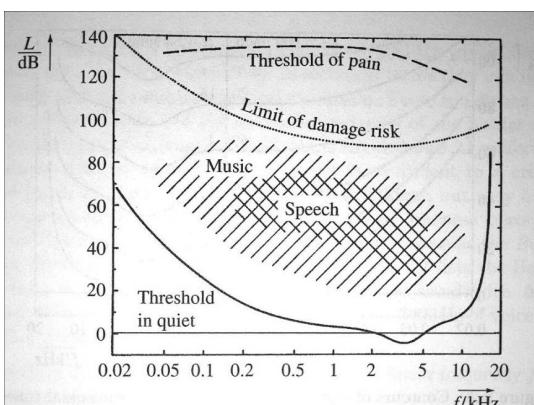
**Octave:** désigne un doublement de fréquence

$$\Delta H_{\text{oct}} = \log_2(1 + \Delta F/F)$$

Bandé passante de l'oreille: [20 Hz - 20 kHz]

UPS - M1 Informatique - Christine SéDAC

## I.4 Perception des sons (3/3)



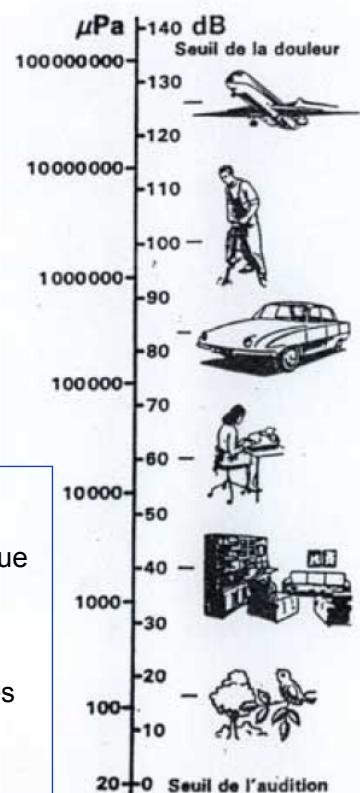
**Dynamique de l'oreille:** 130dB douleur; dès 85 dB dommages

**Imperfections de l'oreille:**

- **Sensibilité non uniforme en fréquence:** par ex à 20dB, 1kHz '+ fort' que 10kHz et 100 Hz non perçu
- **Sensibilité 'fausse' pour les sons purs:** bande + large sur le système auditif qu'en réalité
- **effet de masque:** sur 2 sons de fréquences proches et d'intensité très différentes, seul le + fort sera détecté.

**Imperfections utilisées pour le codage MP3:**

[20Hz, 20kHz] + effet de masque + basses fréquences (mal localisées par l'oreille) en mono. Et pour les sonneries téléphoniques des jeux!



## I.5 Description Acoustique des sons de parole

### I.5.1 La transformée temps-fréquence : Fourier

### I.5.2 Les sons

### I.5.3 Voyelles

### I.5.4 Consonnes

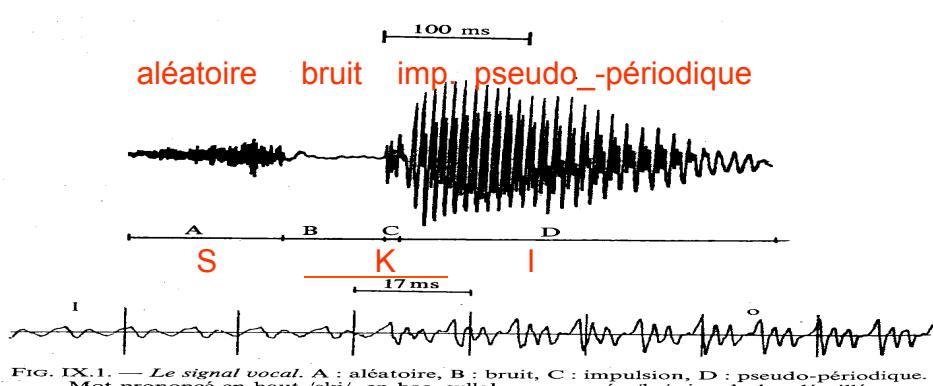
### I.5.5 Paramètres acoustiques/articulatoires

### I.5.6 Lecture de spectrogramme

UPS - M1 Informatique - Christine SéDAC

## I.5 Description acoustique des sons de parole

✓ **Signal continu, d'énergie finie, non stationnaire :** structure complexe et variable dans le temps, tantôt pseudopériodique, tantôt aléatoire, tantôt impulsionale. Cette structure reflète l'organisation temporelle des gestes de production et sur l'onde vocale apparaissent quelques caractéristiques de la source et du conduit vocal.



✓ **Bande passante de l'oreille :** [20Hz – 20000Hz]

✓ **Information linguistique :** [50Hz – 4000Hz]

✓ **Canal téléphonique :** [300 – 3300Hz] => perte d'information

UPS - M1 Informatique - Christine SéDAC

### I.5.1 Le théorème de Fourier

- Le théorème de Fourier exprime mathématiquement le fait qu'un signal quelconque peut être considéré comme la superposition d'un nombre fini ou infini de signaux sinusoïdaux.
- Les conséquences sont:
  - Un signal quelconque est décomposable en une série de signaux sinusoïdaux
  - Si le **signal est périodique**, il peut s'exprimer sous forme d'une série de Fourier: les termes de la série sont des signaux sinusoïdaux dont les fréquences varient comme multiples d'une fréquence de base  $f_0$
  - Si le **signal n'est pas périodique**, il peut s'exprimer sous forme d'une intégrale de Fourier (extension continue de la série de Fourier); les signaux sinusoïdaux constituant ont des fréquences continument réparties

UPS - M1 Informatique - Christine Sénac

### I 5.1 Transformée de Fourier discrète

**L**ORSQU'ON désire calculer la transformée de Fourier d'une fonction  $x(t)$  à l'aide d'un ordinateur, ce dernier n'ayant qu'un nombre fini de mots de taille finie, on est amené à:

- discréteriser la fonction temporelle,
- tronquer la fonction temporelle,
- discréteriser la fonction fréquentielle.

$$X(f) = \int_{-\infty}^{+\infty} x(t) e^{-j2\pi ft} dt$$

En approchant l'intégrale par une somme d'aires de rectangles de durée  $T_e$  et en limitant la durée d'intégration à l'intervalle  $[0, (N-1)T_e]$ , on obtient:

$$X(f) \approx T_e \sum_{n=0}^{(N-1)} x(nT_e) e^{-j2\pi f n T_e}$$

Ce qui donne pour les valeurs de fréquences  $f_k = kf_e/N$ :

$$X(f_k) \approx T_e \sum_{n=0}^{(N-1)} x(nT_e) e^{-j2\pi \frac{nk}{N} f_e T_e} \approx T_e \sum_{n=0}^{(N-1)} x(nT_e) e^{-j2\pi \frac{nk}{N}}$$

UPS - M1 Informatique - Christine Sénac

## I 5.1 Transformée de Fourier et son inverse

### Définition de la TFD

On appelle transformée de Fourier discrète d'une suite de  $N$  termes  $x(0), x(1), \dots, x(N-1)$ , la suite de  $N$  termes  $X(0), X(1), \dots, X(N-1)$ , définis par

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

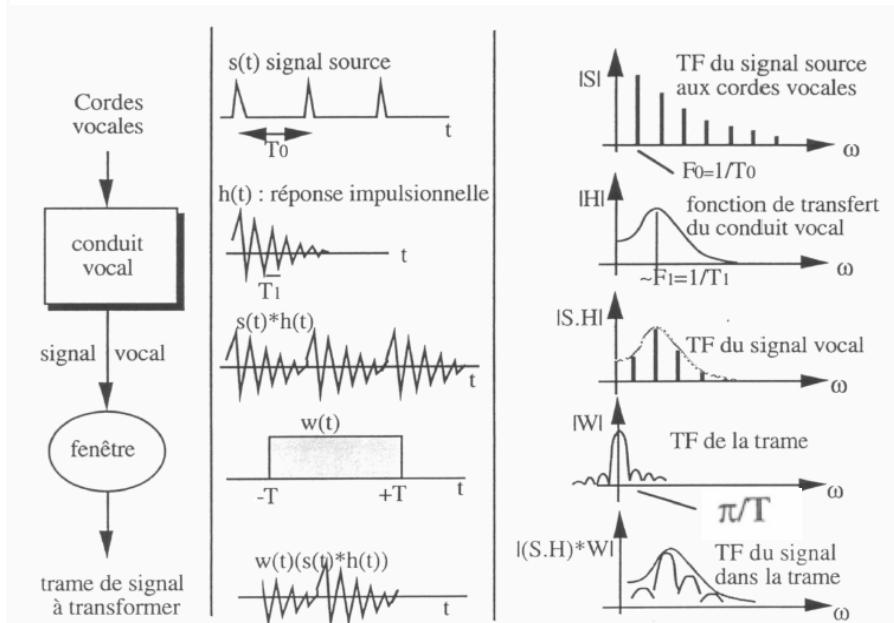
En pratique, les  $N$  termes  $x(n)$  peuvent être  $N$  échantillons d'un signal analogique échantillonné:  $x_n = x(nT_e)$ , et les  $N$  termes  $X(k)$  correspondre à une approximation (à un facteur multiplicatif  $T_e$  près) de la transformée de Fourier de ce signal aux  $N$  points de fréquence  $f_k = kf_e/N$ , avec  $k$  entre 0 et  $N-1$ , c'est à dire  $f$  entre 0 et  $f_e$ .

### Inversion de la TFD

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

UPS - M1 Informatique - Christine SéDAC

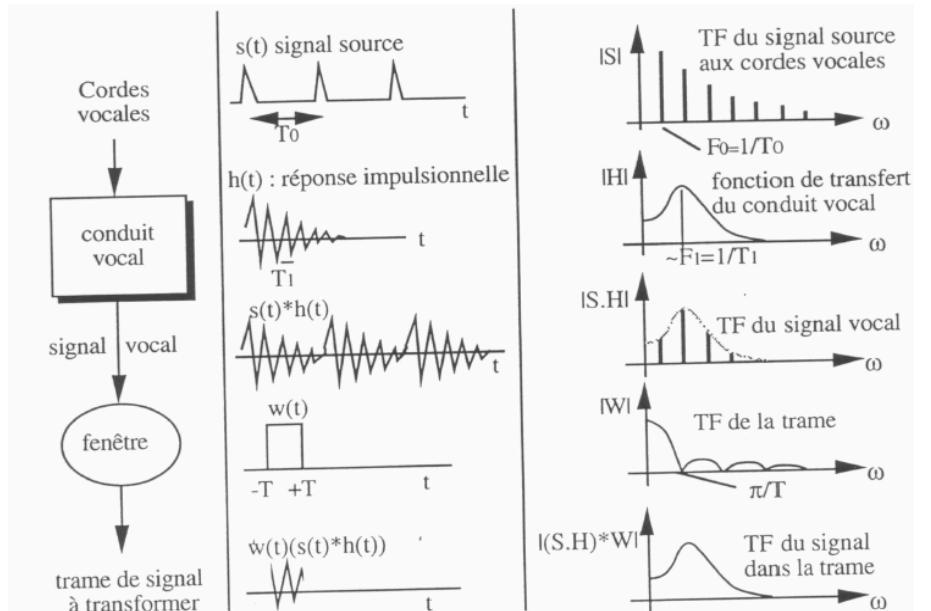
### I.5.1 Transformée de Fourier d'une fenêtre plus longue que la période du fondamental



Faible lissage fréquentiel et visualisation des harmoniques  
=> Spectre à bande étroite

UPS - M1 Informatique - Christine SéDAC

## I 5.1 Transformée de Fourier d'une fenêtre plus courte que la période du fondamental



Fort lissage fréquentiel et pas d'harmoniques  
=> Spectre à large bande

UPS - M1 Informatique - Christine Sénac

## I 5.1 Passage des transformées de Fourier à la parole sous forme de Spectrogramme

Calcul d'un spectrogramme:

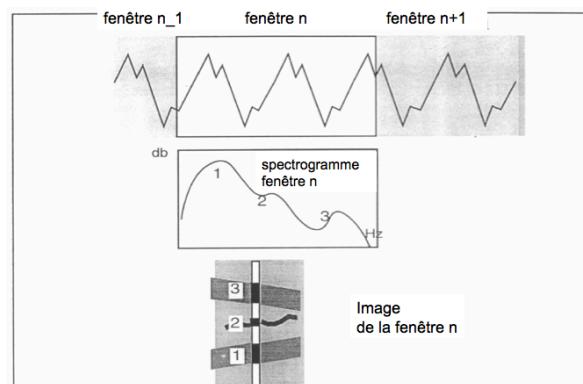
DFT d'une fenêtre de 4 à 32ms qui se déplace de la moitié de sa durée.

Avec une fréquence d'échantillonnage de 16Khz, cette fenêtre a donc entre 64 et 512 points.

⇒DFT entre 64 et 512 points

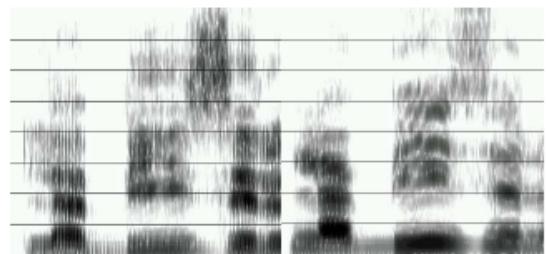
⇒Pour lisser le spectre, on utilise une DFT avec plus de points (au moins 256) ce qui permet d'interpoler le spectre plus finement.

⇒« zero padding »: le signal de départ est complété par des zéros. So on utilise une DFT de 512 points et que la fenêtre a 64 points, on complète par (512-64) points.

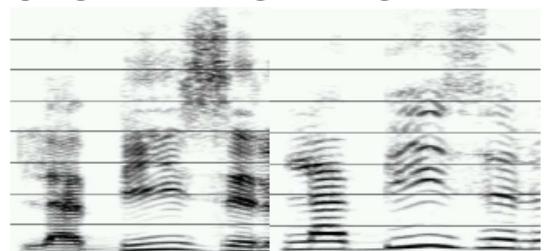


UPS - M1 Informatique - Christine Sénac

### I.5.1 Bande large vs. Bande étroite



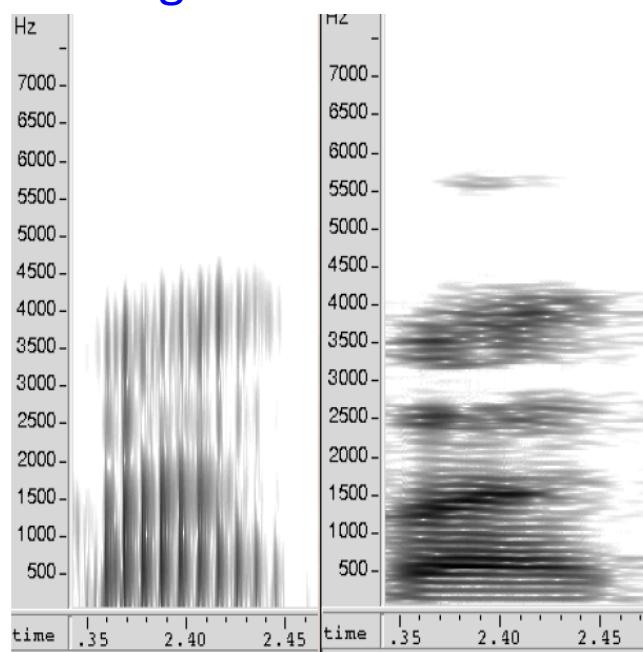
Spectrogrammes à bande large (locuteur à gauche, locutrice à droite)



Spectrogrammes à bande étroite (locuteur à gauche, locutrice à droite)

UPS - M1 Informatique - Christine Sénac

### I.5.1 Bande large vs. Bande étroite



bande large : 500 Hz, fenêtre 32 pts - bande étroite : 32 Hz, fenêtre 512 pts

UPS - M1 Informatique - Christine Sénac

## I.5.1 Illustration du spectre

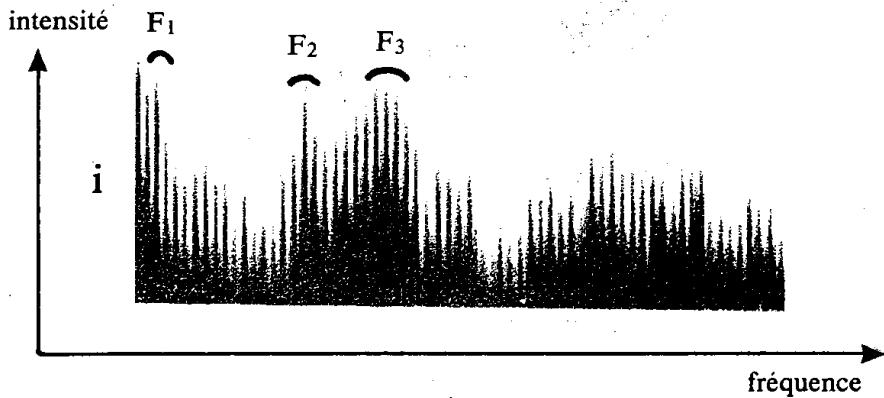


FIG. III.2. — « Section » spectrographique d'une voyelle [i]. Les 3 zones d'harmoniques renforcées (formants) sont identifiées.

UPS - M1 Informatique - Christine Sénaç

## I.5.1 Spectrogramme

- ✓ Représentation de l'intensité des fréquences en fonction du temps
- ✓ Intensité des composants spectraux liée au degré de noirceur du tracé
- ✓ Les **formants** = fréquences de résonance du conduit vocal ( $F_1$ ,  $F_2$ ,  $F_3$ )  
= maxima spectraux correspondant à des harmoniques (multiples de  $F_0$ )

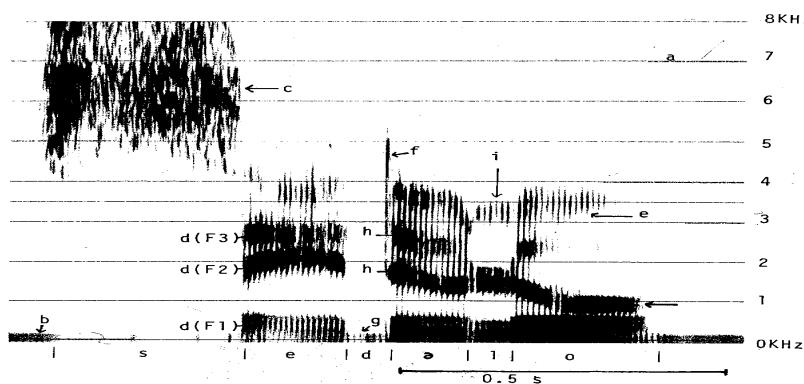


FIG. III.1. — Spectrogramme de la phrase « C'est de l'eau » [sed ə lo] Voiceprint A700  
0-8 kHz.

- (a) ligne de calibration ;
- (b) « ligne de base » produite par l'appareil en l'absence de signal.
- (c) bruit de friction [s]
- (d) formants : la mesure se fait au centre de la bande noire.
- (e) modulation de l'énergie à la fréquence fondamentale (cf. texte)
- (f) barre d'explosion de l'occlusive
- (g) barre de voisement
- (h) transitions formantiques
- (i) formant faible de la sonante [t]

UPS - M1 Informatique - Christine Sénaç

## I.5.2 Les sons du français (niveau phonologique)

**Phonème** = élément sonore le plus bref qui permet de distinguer différents mots

### Consonnes

[p] paie	[t] taise	[k] quai
[b] baie	[d] dais	[g] gai
[m] mais	[n] nez	[ɲ] gagner
[f] fait	[s] sait	[ʃ] chez
[v] vais	[z] zéro	[ʒ] geai
[w] ouais	[ɥ] huer	[j] yéyé
	[l] lait	[R] raie

### Voyelles

[i] lit	[y] lu	[u] loup
[e] les	[ø] leu	[o] lot
[ɛ] lait	[œ] leur	[ɔ] lotte
[a] là	[ɔ] le	[œ̃] brun
[ə] lin	[ɑ̃] lent	[õ] long

*Note : Les distinctions vocaliques [e]-[ɛ], [ø]-[œ] et [o]-[ɔ] ne sont pas faites dans tous les contextes et par tous les locuteurs du français. Par contre, certains locuteurs font aussi des distinctions entre patte et pâte, ([a]-[ɑ]) ainsi qu'entre brin et brun ([ɛ]-[œ]).*

UPS - M1 Informatique - Christine Sénaç

## I.5.3 Les voyelles (1/4)

### • Les voyelles orales [a e i u o y] et nasales: [ã ë õ]

- ✓ Se différencient les unes des autres par leur lieu d'articulation, leur aperture et leur nasalisation.
- ✓ Structure acoustique caractérisée principalement par la **présence de formants**.
- ✓ Pour des fréquences < 5KHz, les voyelles d'un sujet masculin possèdent 5 formants: le 1er (F1) correspond au pic spectral ayant la fréquence la plus basse.
- ✓ MAIS... il existe parfois des 'faux formants'. Ex le formant glottal de FigIII.3
- ✓ Pour les **nasales**: abaissement du voile du palais => mise en parallèle des cavités buccales et nasales => modification des formants.

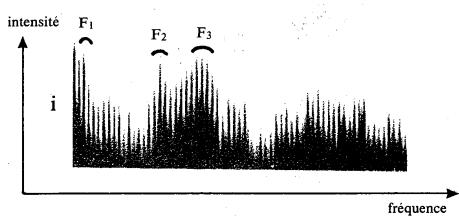


FIG. III.2. — « Section » spectrographique d'une voyelle [i]. Les 3 zones d'harmoniques renforcées (formants) sont identifiées.

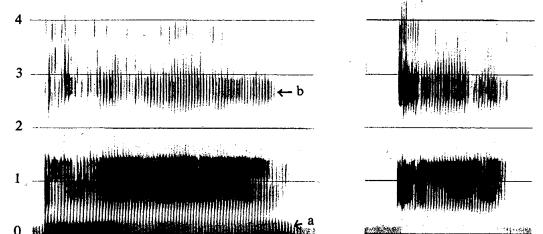


FIG. III.3. — Spectrogramme d'une voyelle [a] : voix faible. Notez le formant glottal et la faiblesse de F3 (b). FIG. III.4. — Spectrogramme d'une voyelle [a] par le même sujet que celui de la figure III.3 : voix forte ; pas de formant glottal et F3 intense.

UPS - M1 Informatique - Christine Sénaç

### I.5.3 Le triangle articulatoire vocalique des voyelles orales (2/4)

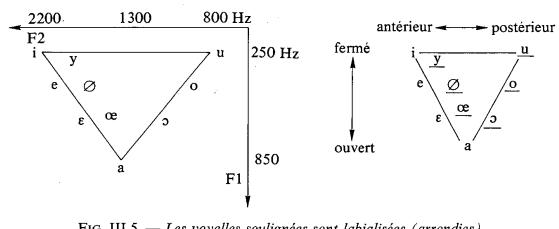


FIG. III.5. — Les voyelles soulignées sont labialisées (arrondies).

- Ouverture articulatoire <-> augmentation de F1
- Antériorisation de l'articulation <-> augmentation de F2
- Augmentation de la labialisation (ouverture des lèvres) <-> baisse de F2 et F3

- 200 prononciations par nuage
- Variabilité intra et inter-locuteur dûe à la physiologie, à la coarticulation...

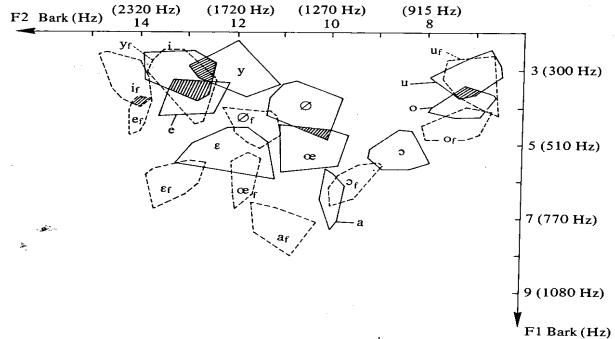


FIG. III.6. — Zones de dispersion des voyelles non nasales du français sur le plan  $F_1/F_2$  (échelle de Bark).

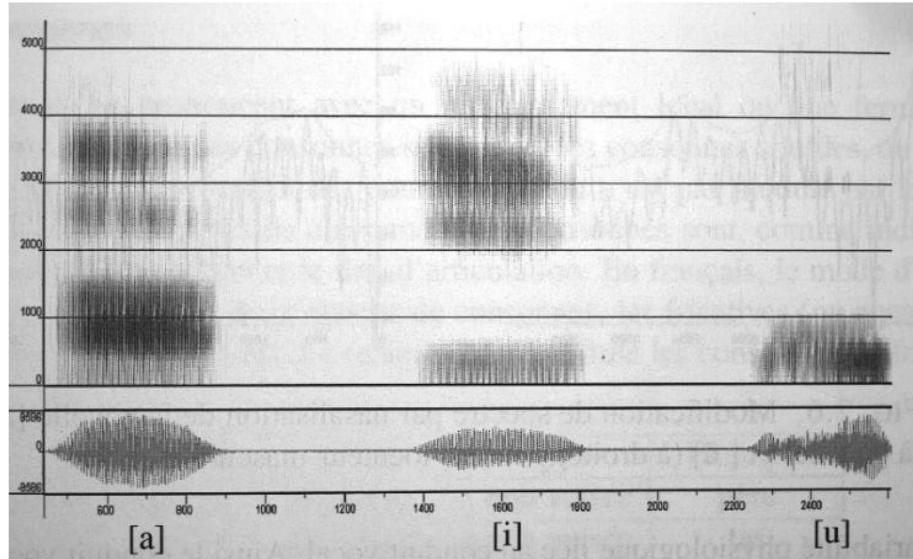
Sujets masculins : —————

Sujets féminins : - - - - -

Les hachures délimitent les zones de recouvrement pour un même sexe.

UPS - M1 Informatique - Christine Sénaç

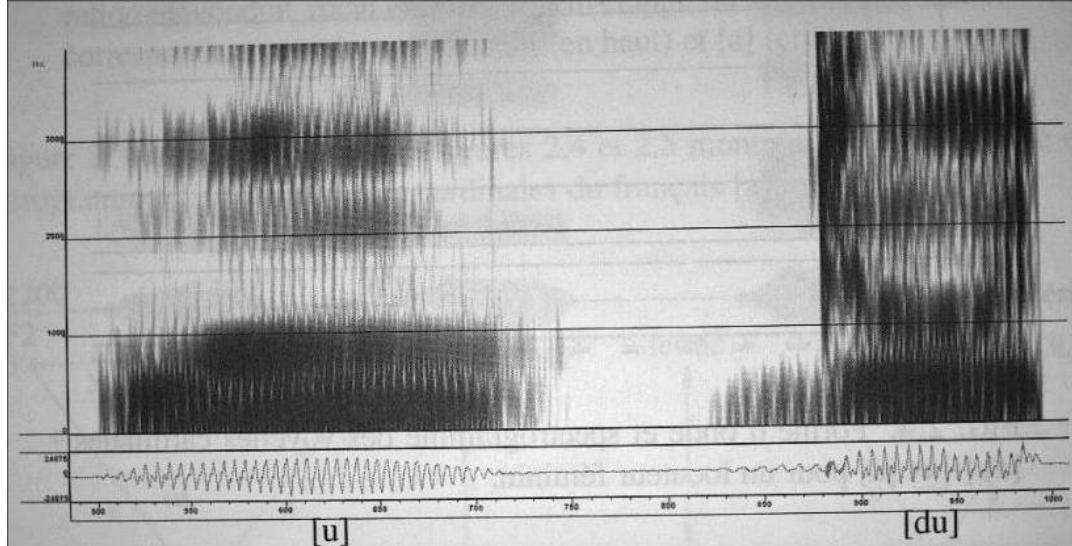
### I.5.3 Description acoustique des voyelles (3/4)



Voyelles extrêmes /a/, /i/, /u/

UPS - M1 Informatique - Christine Sénaç

### I.5.3 Description acoustique des voyelles (3/4)



Une illustration de la loi du moindre effort : /dut/ (mot “doute”)

UPS - M1 Informatique - Christine Sénac

#### I.5.4 Les consonnes (1/12)

- Se comportent comme des événements apériodiques, bruits de turbulence.
  - En français la distinction du mode d'articulation conduit à deux classes majeures: les fricatives et les occlusives.

**Les fricatives** : créées par un rétrécissement du conduit vocal au niveau du lieu d'articulation qui peut être:

le palais [ʃ (chez), ʒ (gilet)], les dents [s (cela), z (zut)] ou les lèvres [f (filet), v (vilain) ].

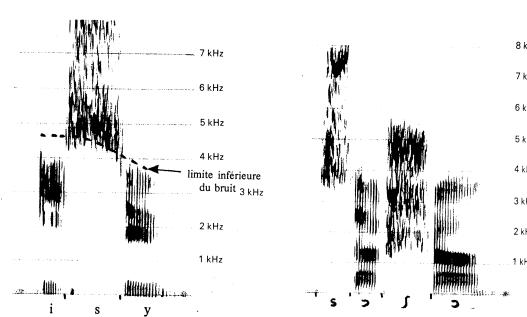
- Les fricatives non voisées [s ſ tʃ] sont caractérisées par un écoulement d'air turbulent à travers la glotte.
  - Les fricatives voisées [z ʒ v] combinent des composantes d'excitation périodique et turbulente: les cordes vocales s'ouvrent et se ferment périodiquement , mais il y a obstruction en aval.

### Les consonnes fricatives ou constrictives

Consonances et dissonances dans le jeu de l'interprétation

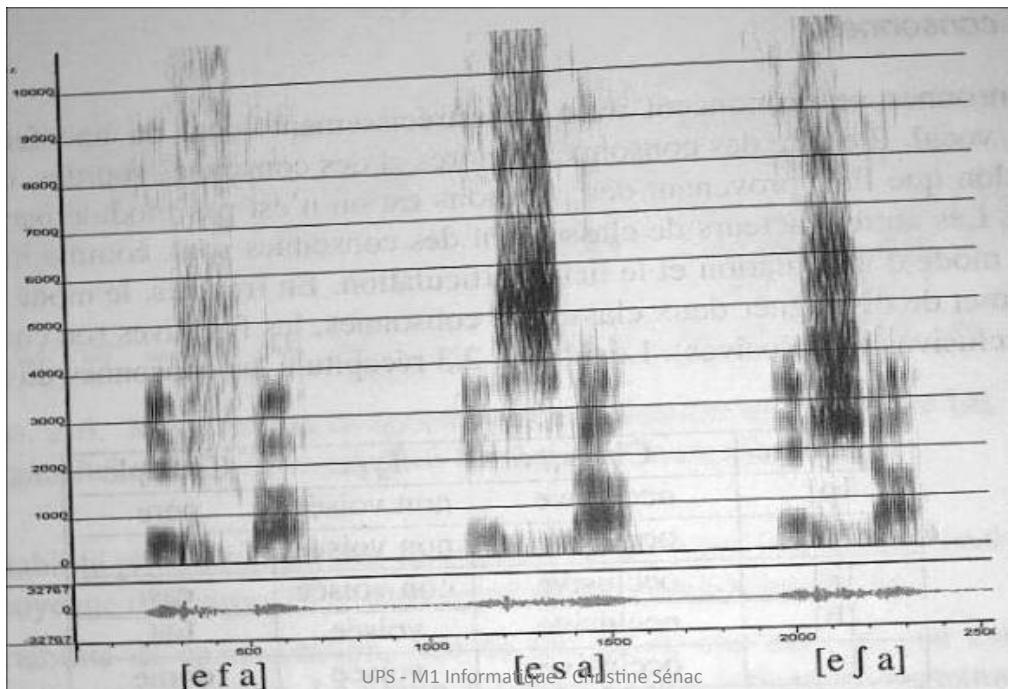
lабио-дентал (*/f/*, */v/*)

## Examen du conduit oral



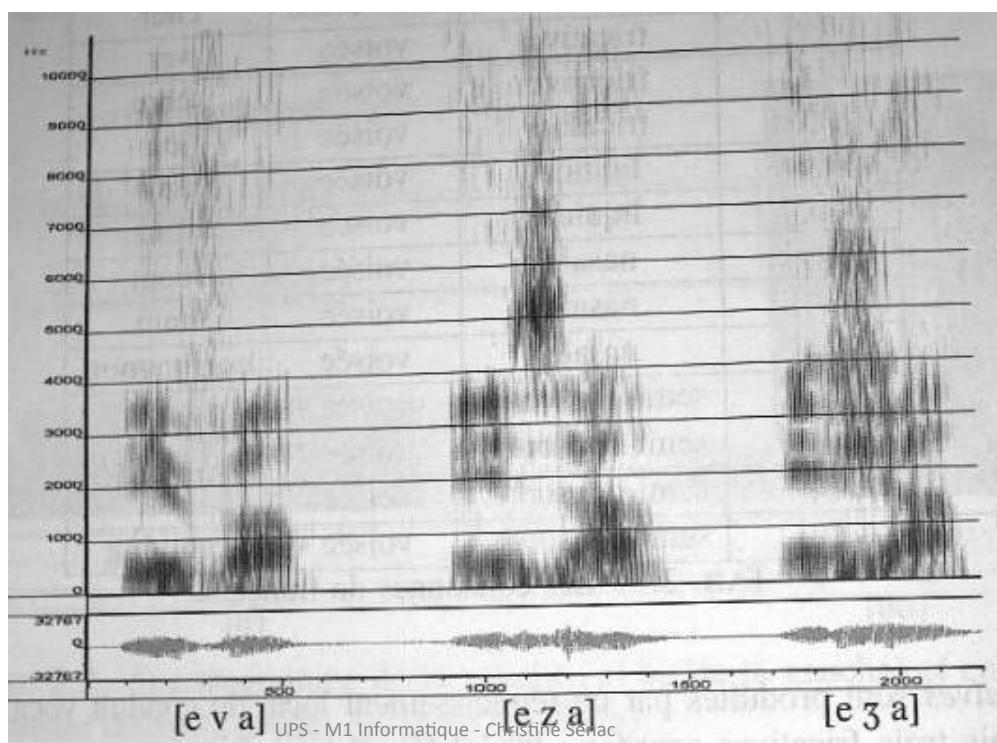
## I.5.4 Les consonnes (2/12)

Consonnes fricatives non voisées



## I.5.4 Les consonnes (3/12)

Consonnes fricatives voisées



## I.5.4 Les consonnes (4/12)

**Les occlusives [b d g p t k]:** elles correspondent à des sons essentiellement dynamiques. Une forte pression est créée en amont d'une **occlusion (fermeture)** maintenue en un certain point du conduit vocal (qui peut être ici le palais [k,g], les dents[t,d], ou les lèvres[p,b]), puis **relâchée brusquement**.

- Pour les **occlusives voisées [b d g]**, un **son basse fréquence** est émis durant l'occlusion alors que pour les **non voisées [p t k]**, l'occlusion correspond à un **silence**.

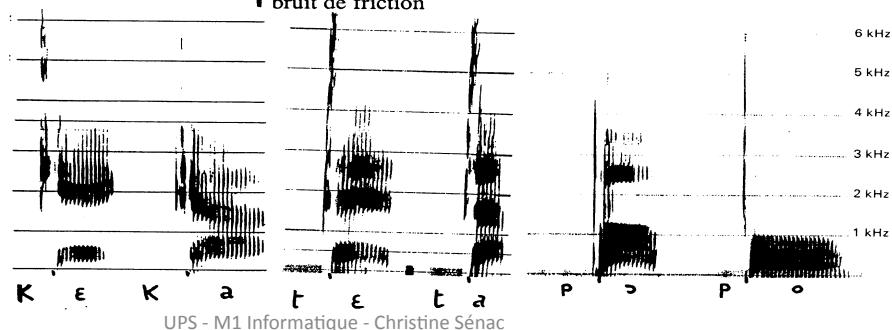
### Les consonnes occlusives

voisées (/b/, /d/, /g/), non voisées (/p/, /t/, /k/)       
 occlusion complète du conduit vocal     
 

labiale	<i>p</i> , <i>t</i>
dentale	<i>t</i> , <i>d</i>
palato vélaire ou vélaire	<i>k</i> , <i>g</i>

suite d'événements     
 

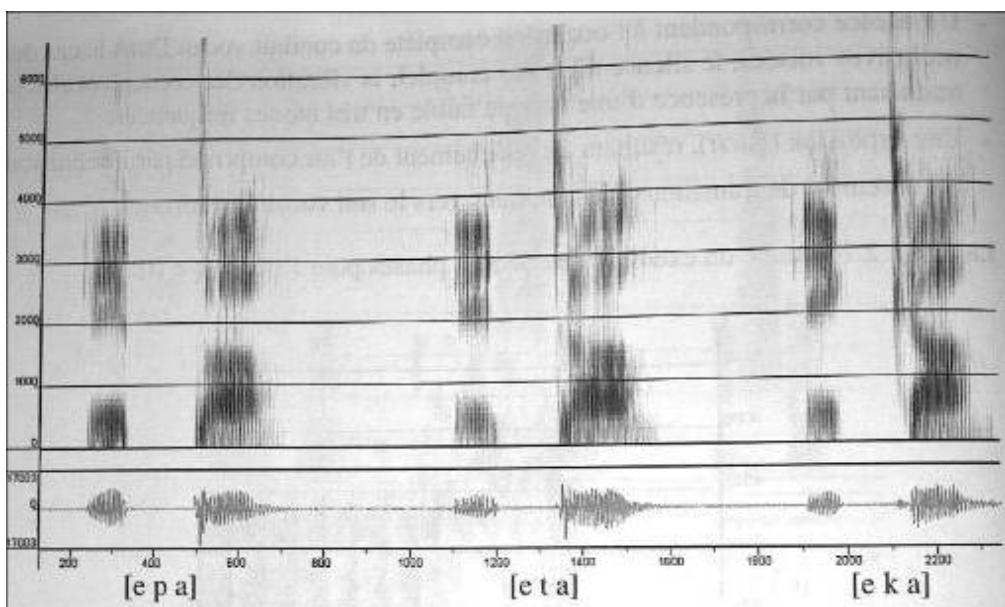
silence ou barre de voisement
barre d'explosion
bruit de friction



UPS - M1 Informatique - Christine Sénac

## I.5.4 Les consonnes (5/12)

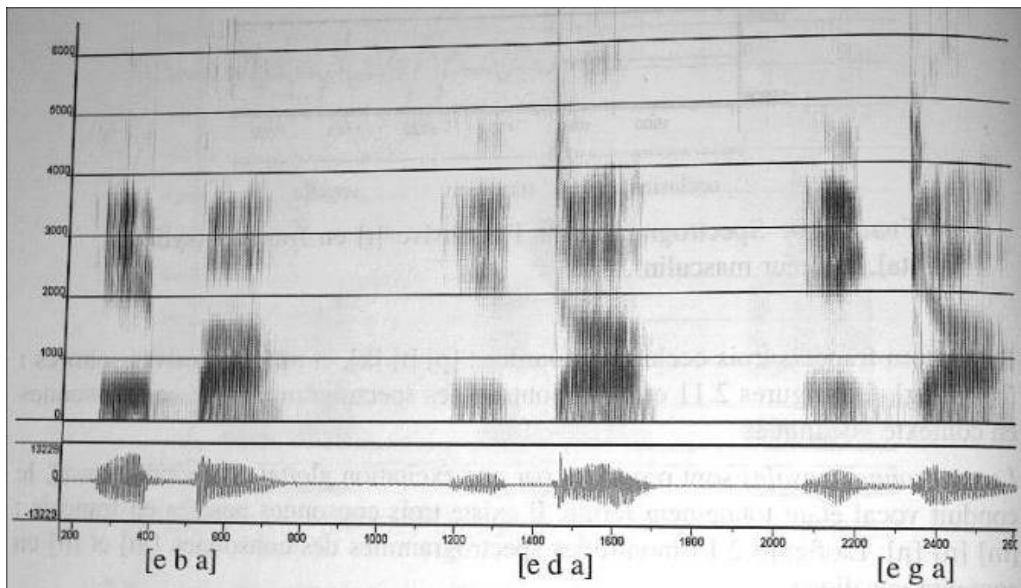
### Occlusives non voisées



UPS - M1 Informatique - Christine Sénac

## I.5.4 Les consonnes (6/12)

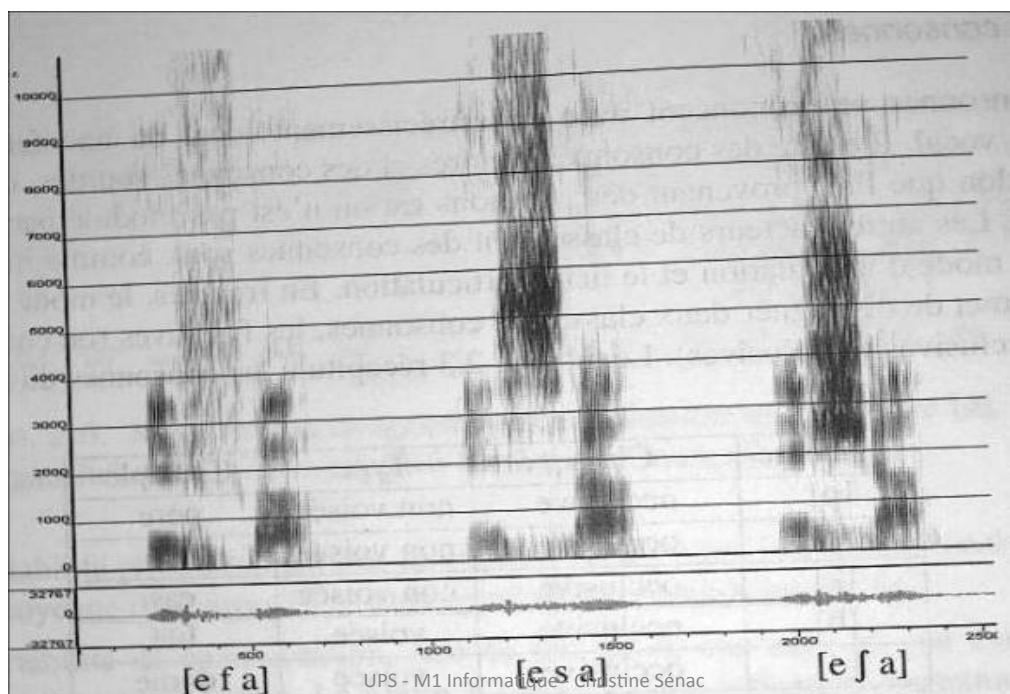
### Occlusives voisées



UPS - M1 Informatique - Christine SéDAC

## I.5.4 Les consonnes (7/12)

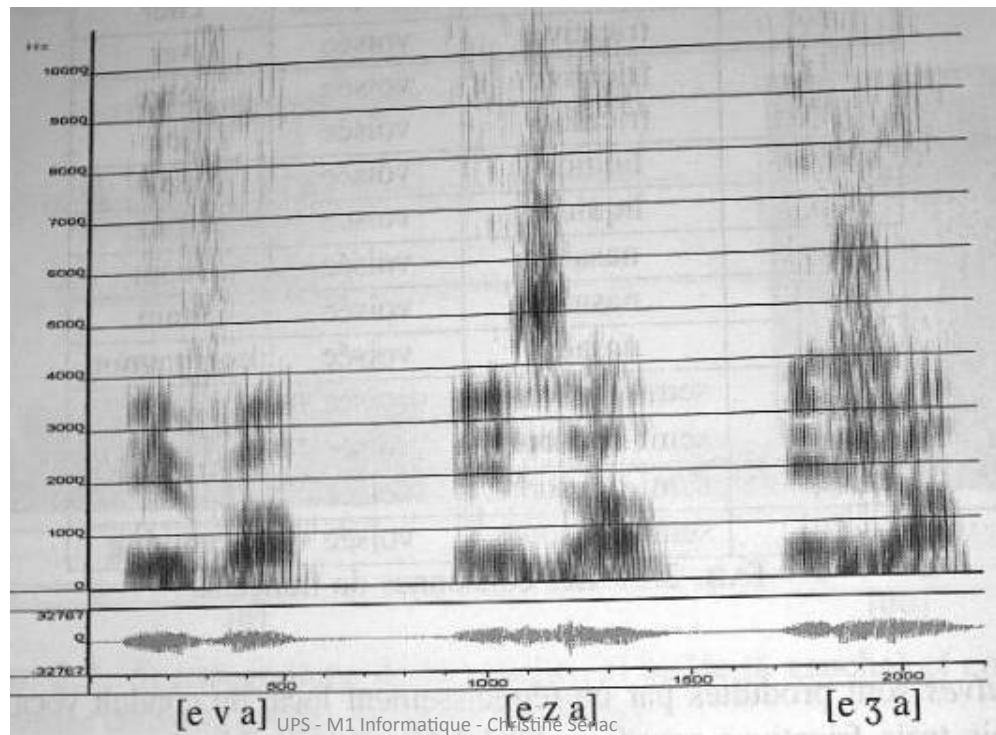
### Fricatives non voisées



UPS - M1 Informatique - Christine SéDAC

## I.5.4 Les consonnes (8/12)

### Fricatives voisées



## I.5.4 Les consonnes (9/12)

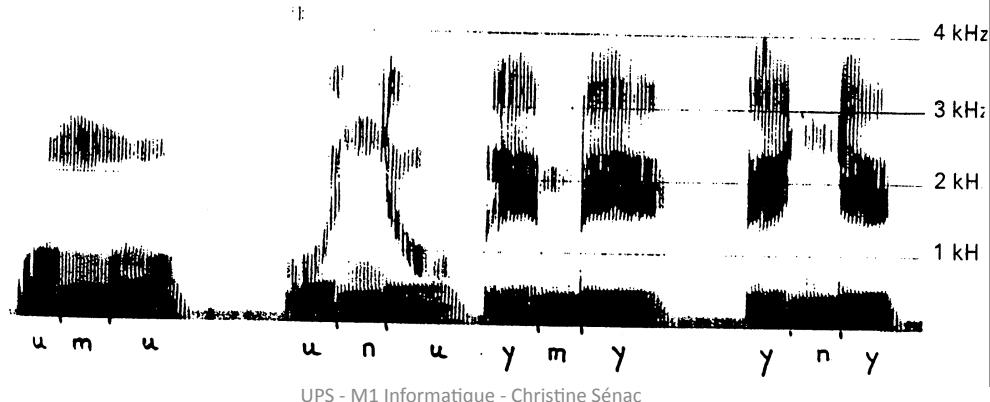
### Consonnes nasales [m (mon), n(nez) et h (gagnant)]

➤ (rarement dévoisées) font intervenir les cavités nasales par abaissement du voile du palais

les consonnes nasales      m, n, ɲ.

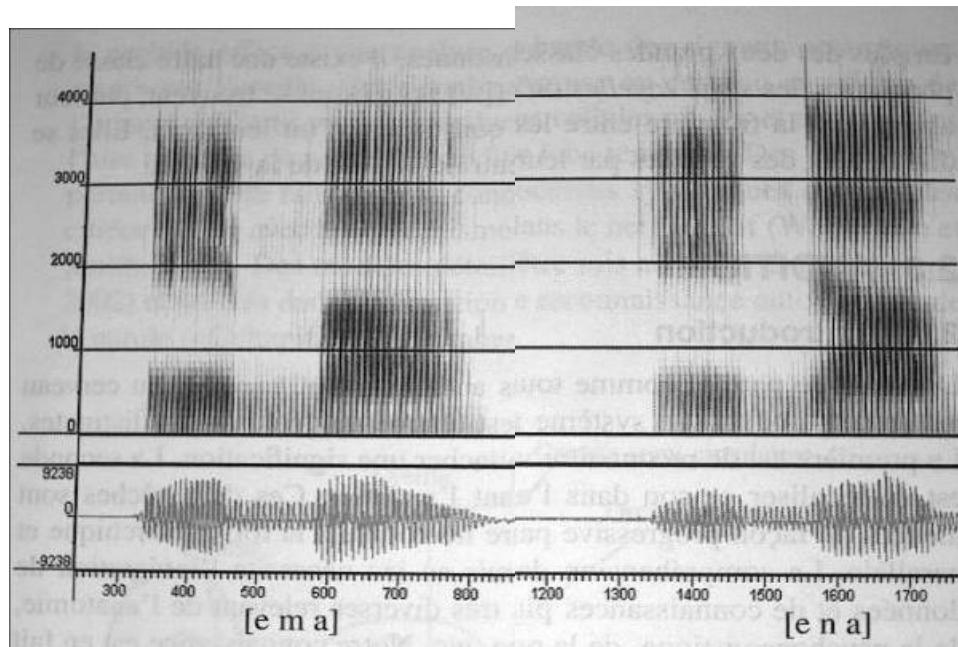
- couplage      conduit oral/ conduit nasal
- occlusion      labiale /m/, dentale /n/

↳ les voyelles s'enchaînent  
en ignorant !m!



## I.5.4 Les consonnes (10/12)

### Consonnes nasales

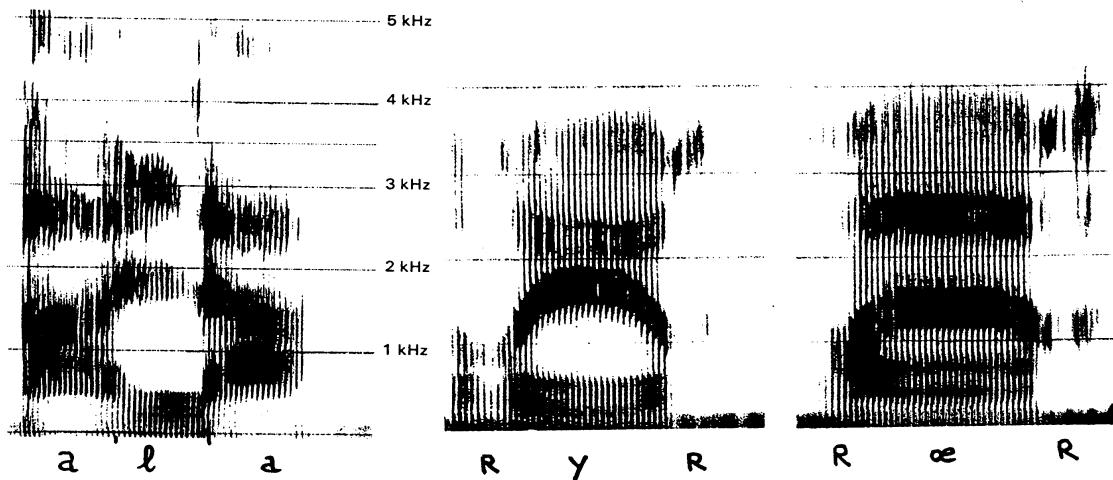


UPS - M1 Informatique - Christine Sénac

## I.5.4 Les consonnes (11/12)

### Les liquides [ l (le), R (récit) ]

- les liquides → /l/ articulation alvéolaire, passage latéral de l'air (en général 1 seul côté)  
 $F_1 = 300 \text{ Hz}$  (stable)  
 $F_2$  très variable dépend du contexte
- /r/ articulation dorso-vélaire (i.e. au fond du palais) } voisé  
dépend du contexte de manière dramatique } et mon voisé



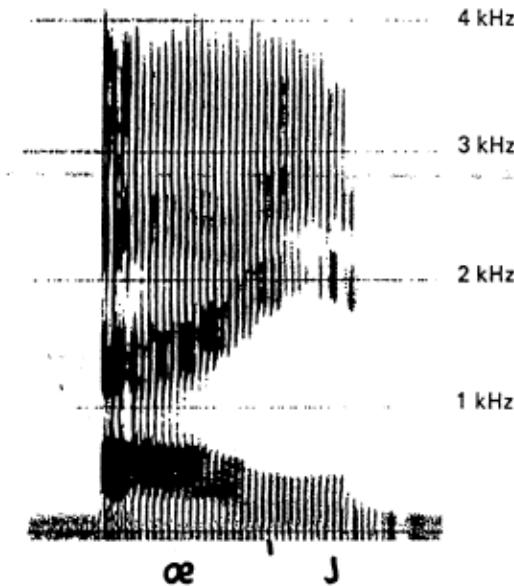
UPS - M1 Informatique - Christine Sénac

## I.5.4 Les consonnes (12/12)

Semi consonnes ou semi voyelles:

- Voisée
- Transitoire
- Voyelle non réalisée

- /j/ → /i/
- /w/ → /u/
- /ɥ/ → /y/



UPS - M1 Informatique - Christine Sénac

## I.6 Classification des sons de parole français (1/2)

Paramètres: acoustiques et/ou similarité articulatoire  
(Jacobson, Fant, Halle 1952)

Voyelles :

- orale/nasale
- Degré d'ouverture du conduit vocal
- Articulation
- Labialité

Consonnes:

- voisée/non voisée
- Mode d' articulation (constrictive, fricative, nasale, liquide)
- Articulation (labiale, dentale, palatale, vélaire)

UPS - M1 Informatique - Christine Sénac

## 1.6 Classification des sons de parole français (2/2)

TABLEAU II. — <i>Classification des phonèmes du français en traits distinctifs</i>			
CONSONNES Mode d'articulation ↓	Labiales	Dentales	Vélo-palatales ← Lieu d'articulation
Occlusives			
non voisées	[p]	[t]	[k]
voisées	[b]	[d]	[g]
Nasales	[m]	[n]	[ɲ]
Fricatives			
non voisées	[f]	[s]	[χ]
voisées	[v]	[z]	[ʒ]
Glissantes	[w]	[ɥ]	[j]
Liquides		[l]	[R]
VOYELLES			
Orales			
	Antérieures		Postérieures
	Non arrondies		Arrondies
Fermées	[i]	[y]	[u]
	[e]	[ø]	[o]
	[ɛ]	[œ]	[ɔ]
Ouvertes	[a]		
Nasales		Antérieures	Postérieures
Fermées	[ã]		[õ]
Ouvertes		[ã]	

UPS - M1 Informatique - Christine SéDAC

## 1.7 Lecture de spectrogramme

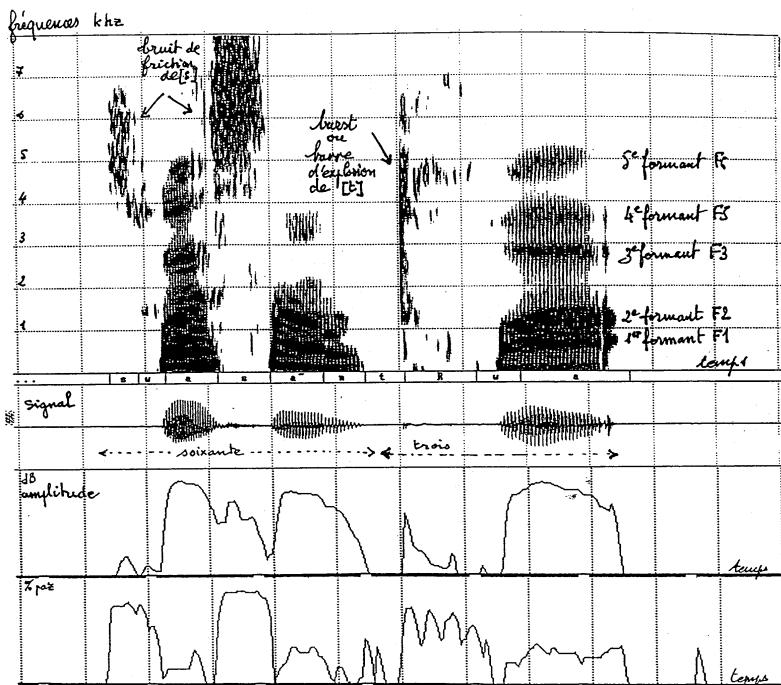
Placer dans l'ordre:

➤ Voyelles – Triangle vocalique

➤ Consonnes :

- Occlusives non voisées
- Occlusives voisées
- Fricatives non voisées
- Fricatives voisées
- autre...

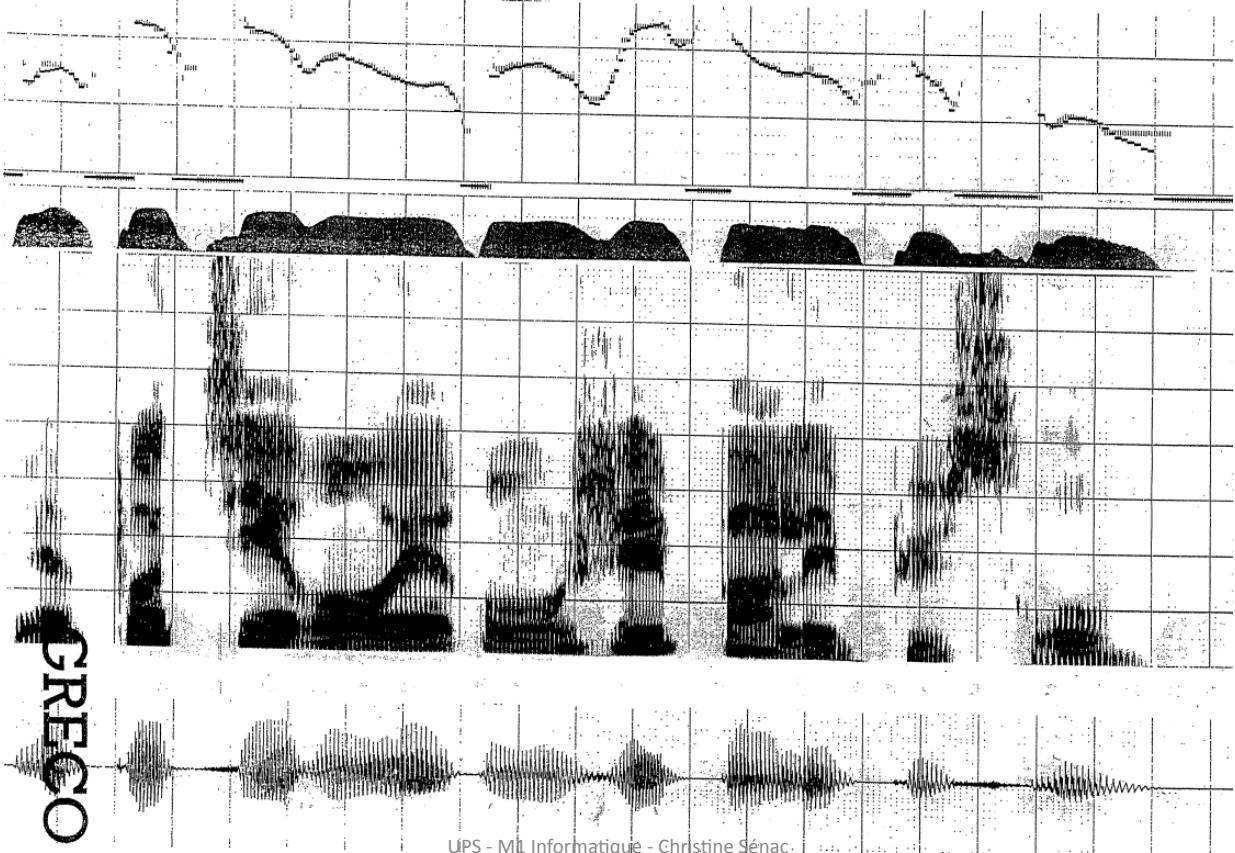
## 1.7 Lecture de spectrogramme: un exemple

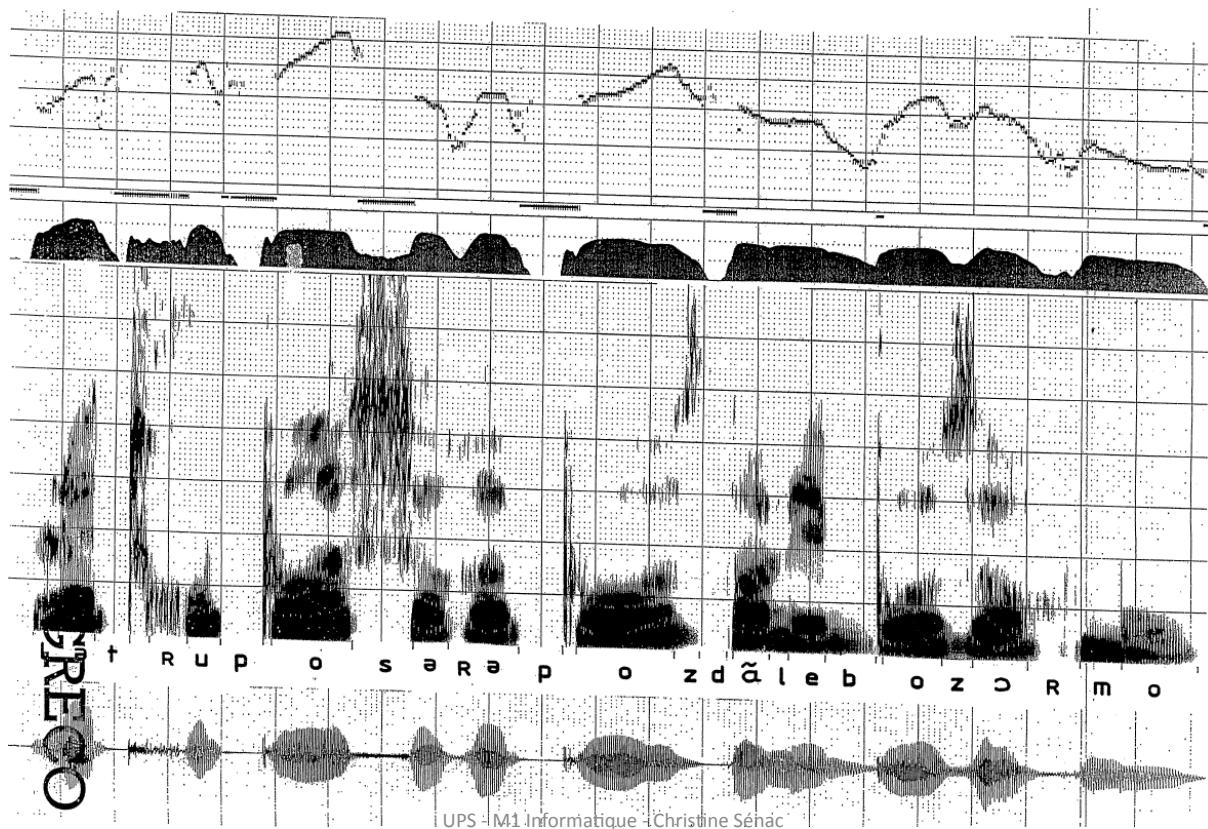


**Amplitude:** Les voyelles dominent les consonnes en amplitude.

**Taux de passages à zéro:** les fricatives dominent les autres sons. L'explosion de [t] donne lieu à un pic de paz. Le [R] apparaît comme fricatif.

UPS - M1 Informatique - Christine Sénac





I UPS IM1 Informatique - Christine SéDAC