

# CHAPITRE2

## RECONNAISSANCE AUTOMATIQUE DE LA PAROLE

### PARTIE 1: GÉNÉRALITÉS

Dr. Djellali Hayet  
Département Informatique  
UBMA Université

Master IATI

Semestre2, 2020

## PLAN

- × Introduction
- × Les sources de Variabilités
- × Les Difficultés
- × La Reconnaissance
- × Modalités Paroles
- × Applications du Traitement Automatique de la Parole
- × Les Sources de Variabilités
- × Typologie des Systèmes
- × Types de paroles
- × Appareil Phonatoire et Conduit vocal
- × Phonétiques

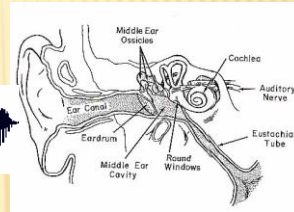
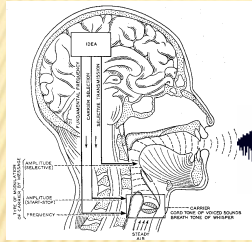
## INTRODUCTION

- ✘ La parole est le principal moyen de communication dans toute société humaine.
- ✘ L'interaction homme-machine devrait plus ou moins passer par la parole.
- ✘ La parole permet de se dégager **de toute obligation de contact physique** avec la machine.
- ✘ Cet avenir n'est cependant pas encore à notre portée..

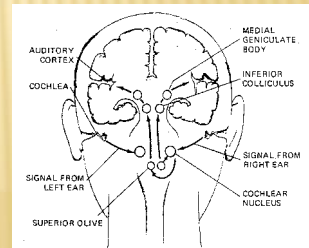
## OBJECTIFS

- ✘ La recherche en Reconnaissance Automatique de la Parole, RAP, étudie le processus humain de génération et de compréhension de la parole.
- ✘ L'étude des organes biologiques de production et de compréhension de la parole.
- ✘ Connaître les différentes taxonomies des sons observables en parole et les Variations qui peuvent y être constatées .
- ✘ Les étapes de constructions d'un système de RAP et la description de chaque module.

## COMMENT L'HOMME LE FAIT?

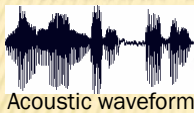


- ✘ Articulation produites
- ✘ Les ondes sonores que L'oreille transporte au cerveau Pour traitement

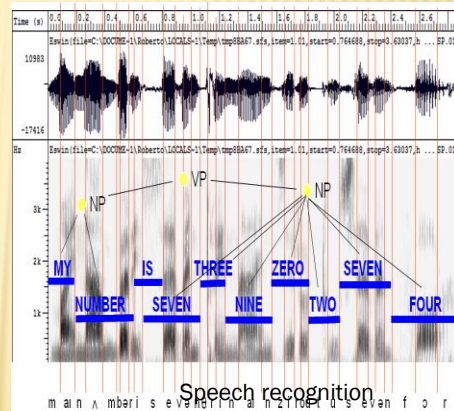


5/34

## COMMENT LES ORDINATEURS LE FONT?



Acoustic signal



- ✘ Numérisation
- ✘ L'analyse Acoustique du signal de parole
- ✘ Interpretation Linguistique

6/34

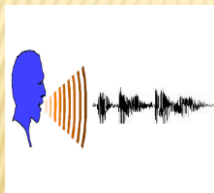
## OU EST LA DIFFICULTÉ?

- ✗ Numérisation
  - + Convertir le signal analogique en une représentation numérique
- ✗ Traitement du Signal
  - + Séparation de la parole du bruit de fond
- ✗ Phonétique
  - + Variabilité de la voix des locuteurs
- ✗ Phonologie
  - + Reconnaître et distinguer les sons (phonemes similaire)
- ✗ Lexique et syntaxe
  - + Les paramètres de la parole continue
- ✗ Syntaxe et pragmatique
  - + Interpréter les paramètres prosodiques

7/34

## INTRODUCTION

• A partir d'un signal de parole différentes informations peuvent être extraites:



Reconnaissance du  
Locuteur

QUI A PARLE

Reconnaissance de la  
Parole

QU'est ce qui a été  
dit?

Reconnaissance de la  
langue

Quelle Langue

Reconnaissance des  
émotions

Quelle l'état  
psychologique?

## RECONNAISSANCE

- ✘ L'appellation "reconnaissance de la parole" (ASR pour Automatic Speech Recognition en anglais) se réfère à plusieurs types de systèmes dont la mission est de décoder l'information portée par le signal vocal. On distingue plusieurs types de reconnaissance :
- ✘ **La reconnaissance du locuteur**, dont le but est de reconnaître la personne qui parle parmi une population de locuteurs (identificateur) ou de vérifier son identité (vérificateur).
- ✘ **La reconnaissance de parole**, dont le but est de transcrire l'information symbolique exprimée par le locuteur. ( Monolocuteur, Multilocuteur, ou Indépendant du Locuteur).
- ✘ **Reconnaissance de la Langue**  
Déterminer dans quelle langue la parole est prononcée.
- ✘ **Reconnaissance des émotions**  
Déterminer l'état psychologique du locuteur (triste, nerveux, content, peur ,...etc)

## DÉFINITION

- ✘ **La parole – faculté de communiquer la pensée par un système de sons articulés émis par les organes de la phonation.**
- ✘ **Reconnaissance Automatique de la Parole**  
*Transformation d'un signal de parole en une séquence de symboles représentative du contenu du signal.*
- ✘ **Le Niveau Acoustique**  
La parole apparaît physiquement comme une variation de la pression de l'air causée et émise par le système articulatoire.

## MODALITÉS PAROLES

- ✗ Pourquoi ?
- ✗ Transmission de la voix est majoritaire
- ✗ Problèmes spécifiques car liés fondamentalement à l'humain
- ✗ Applications très nombreuses pour les télécommunications
- ✗ Très importante en:
  - + recherche d'information (Indexation)
  - + gestion des document

## LES APPLICATIONS EN RAP

- ✗ **Les Serveurs Vocaux**

Ces systèmes traitent des vocabulaires limités (  $\leq 1000$  mots ).  
 Une difficulté de tels systèmes : L'application doit être indépendante du locuteur et du canal de transmission ( téléphone par exemple ). On distingue deux orientations principales dans ce style d'application :

  - + **Les systèmes à commandes vocales** : le système doit reconnaître des mots isolés qui sont interprétés comme des commandes élémentaires. Il est recommandé de choisir un vocabulaire composé de mots.
  - + **Les systèmes de compréhension** : l'utilisateur peut poser une question plus complexe en utilisant des mots clés dans une phrase. On doit adjoindre au système un module de compréhension pour donner un sens à la commande afin de fournir une réponse. La syntaxe de la phrase est généralement très rigide et le mode d'élocution contraint.
- ✗ **Les systèmes de dictée vocale** : Le vocabulaire peut être de l'ordre de 20.000 à 50.000 mots et plus, la syntaxe étant celle du langage naturel. Ces systèmes sont très souvent doivent s'adapter au locuteur.

## LES SOURCES DE VARIABILITÉS

- Les facteurs intra-locuteurs  
co-articulation, variation dans la prononciation, etc.
- Les facteurs inter-locuteurs  
physiologie, age, sexe, psychologie, familiarité avec l'application, etc.
- L'environnement  
bruit, micro, canal de transmission, présence d'autres locuteurs, etc.

## LA VARIABILITÉ INTRA-LOCUTEUR

- ✘ La variabilité intra-locuteur identifie les différences dans le signal produit par une même personne. Cette variation peut résulter de l'état physique ou moral du locuteur.
- ✘ Une maladie des voies respiratoires peut ainsi dégrader la qualité du signal de parole.
- ✘ L'humeur ou l'émotion du locuteur peut influencer son rythme d'élocution.

## LES SOURCES DE VARIABILITÉS - VARIABILITÉ INTER LOCUTEUR

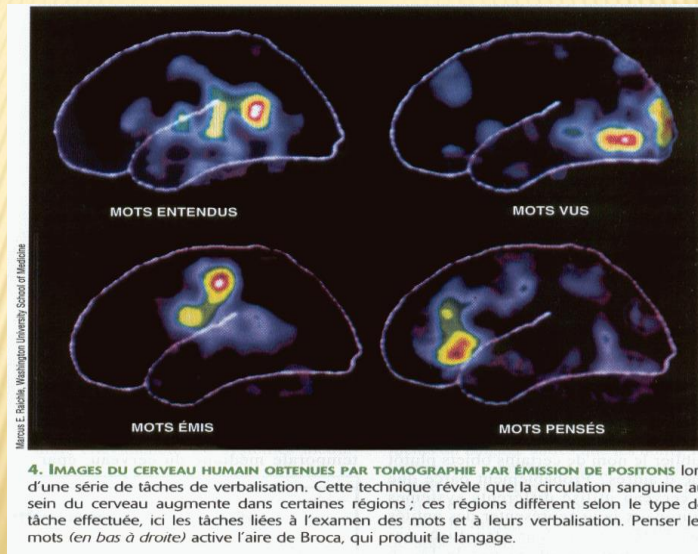
- ✘ Les différences inter-locuteurs sont dues à la nature physiologique. La parole est principalement produite grâce aux cordes vocales qui génèrent un son à une fréquence de base, le fondamental. Cette fréquence de base sera différente d'un genre à l'autre (homme - femme).
- ✘ Une voix d'homme étant plus grave qu'une voix de femme. Ce son est ensuite transformé par l'intermédiaire du conduit vocal, délimité à ses extrémités par le larynx et les lèvres. Cette transformation permet de générer des sons différents (voyelles - consonnes).
- ✘ Le conduit vocal est de forme et de longueur variables selon les individus et, plus généralement, selon le genre et l'âge.
- ✘ Le conduit vocal féminin adulte est, en moyenne, d'une longueur inférieure de 15% à celui d'un conduit vocal masculin adulte.
- ✘ Le conduit vocal d'un enfant en bas âge est bien sûr inférieur en longueur à celui d'un adulte. Un même phonème pourra avoir des réalisations acoustiques très différentes.
- ✘ les différences de prononciation qui existent au sein d'une même langue et qui constituent les accents régionaux

## LES SOURCES DE VARIABILITÉS

- ✘ La variabilité due à l'environnement peut provoquer une dégradation du signal de parole sans que le locuteur ait modifié son mode d'élocution.
- ✘ Les moyens de transport peuvent entraîner des déformations du signal.
- ✘ Le bruit ambiant peut ainsi provoquer une déformation du signal de parole en obligeant le locuteur à accentuer son effort vocal.

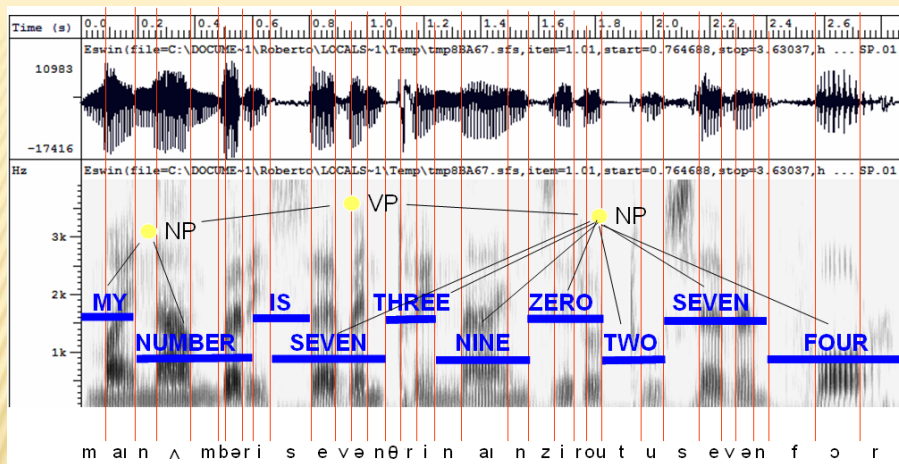


## LE CERVEAU ET LA VERBALISATION [BON, 2004]



## TYPOLOGIE DES SYSTÈMES

- Type de parole
- Taille du vocabulaire
- Niveau de dépendance par rapport aux locuteurs
- Environnement d'utilisation
- Profil des utilisateurs potentiels



- ✗ Identifier un phonème
- ✗ Identify un mot
- ✗ Identifier la structure de la phrase /le sens
- ✗ Interpreter les paramètres prosodiques (pitch)

19/34

## TYPE DE PAROLE

- Mots isolés
- Mots connectés
- Détection de mots clés
- Parole contrainte
- Parole continue
- Parole spontanée

## TAILLE DU VOCABULAIRE

- Quelques mots (5 – 50)
- Petit vocabulaire (50 – 500)
- Vocabulaire moyen (500 – 5000)
- Grand vocabulaire (5000 – 50000)
- Très grand vocabulaire (> 50000)

## ENVIRONNEMENT D'UTILISATION

- Parole large-bande  
(ordinateur, etc.)
- Environnement calme (bureau + micro-casque)
- Parole bande-étroite avec distorsion  
(téléphone, etc.)
- Bruit de fond

## NIVEAUX D'ANALYSE PROPRIÉTÉS ACOUSTIQUES ET LINGUISTIQUES

- ✗ acoustico-phonétique
  - + présence des sons d'une langue
- ✗ Phonotactique
  - + fréquence et enchaînement des sons
- ✗ Prosodique
  - + intonation, rythme
- ✗ Lexical
  - + mots possibles d'une langue
- ✗ Syntaxique
  - + enchaînement possible de mots dans la langue
- ✗ Sémantique
  - + sens de l'enchaînement des mots
- ✗ Pragmatique
  - + contexte

## EXEMPLE D'APPLICATIONS (1)

- ✗ Dictée vocale
  - Parole continue
  - Grand vocabulaire
  - Bureau+micro-casque
  - Adaptation au locuteur
  - Utilisateurs d'ordinateurs

## EXEMPLE D'APPLICATIONS (1)

### ✘ Service téléphonique

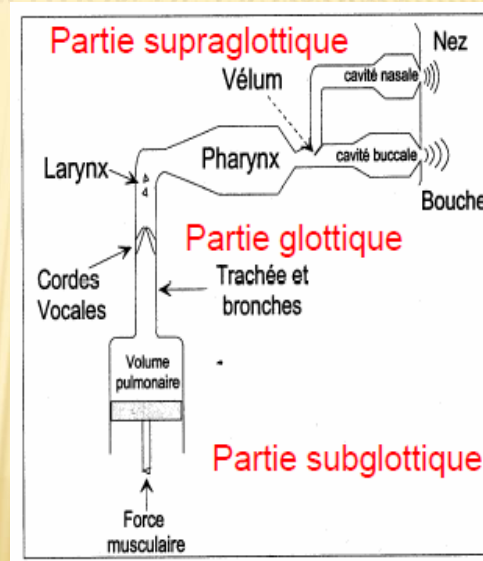
- Détection de mots clés
- Quelques mots
- Indépendant du locuteur
- Parole téléphonique
- Grand public

## L'APPAREIL PHONATOIRE HUMAIN

- ✘ La production de la parole est assurée par plusieurs organes successifs:
- ✘ Les poumons sont indispensables dans ce processus puisqu'ils assurent la génération de l'air sous pression. Cet air, expulsé, traverse alors les cordes vocales qui entrent ou non en action pour produire un voisement qui correspond à la fréquence fondamentale qui est le timbre de la voix. Cette fréquence fondamentale produite est propagée dans l'ensemble du conduit vocal qui est de forme et de volume variable.
- ✘ Plusieurs organes concourent à ces possibles modifications qui permettent de produire des sons différents. Parmi ces organes se trouve la langue, acteur principal des modifications qui peut agir par constriction ou occlusion du conduit vocal.
- ✘ Les dents et les lèvres agissent également par occlusion ou constriction.
- ✘ Le conduit vocal est constitué du seul conduit buccal. La luette et son prolongement vers le palais, le vélum, assurent normalement la fermeture du conduit nasal pendant la production de parole.
- ✘ Le conduit nasal peut, dans certains cas, être connecté au conduit vocal. Cette connexion permet de générer des sons supplémentaires en modifiant le volume de la caisse de résonance normalement constituée par le seul conduit buccal.

## PRODUCTION ET PERCEPTION DE LA PAROLE

- ✘ Anatomie
- ✘ Drygaglo 'Traitement de la parole', EPFL Lausanne(2005).
- ✘ Il n'y a pas à proprement parlé d'organes spécifiques destinés à la production de la parole.
- ✘ La fréquence de vibration des cordes vocales est appelée
- ✘ **fréquence du fondamental (pitch)**
- ✘ **hommes: 80-200 Hz,**
- ✘ **femmes: 150-450 Hz,**
- ✘ **enfants: 200-600 Hz**



## L'APPAREIL PHONATOIRE PARTIE SUBGLOTTIQUE

La production de la parole est assurée par plusieurs organes successifs:

- ✘ Poumons, trachée : Elle fournit l'énergie nécessaire à la phonation en insufflant l'air vers la partie glottique.
- ✘ Les poumons sont indispensables dans ce processus puisqu'ils assurent la génération de l'air sous pression. Cet air, expulsé, traverse alors les cordes vocales.

## **PARTIE GLOTTIQUE OU LARYNX**

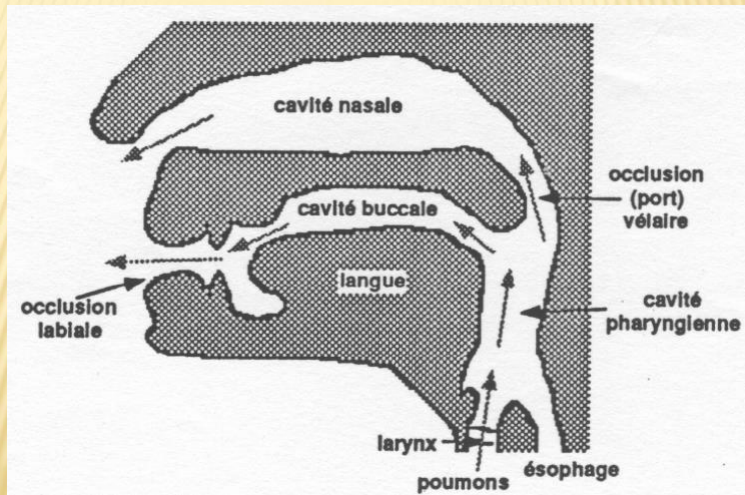
- ensemble de cartilages, ligaments et muscles
- cordes vocales: replis tendus horizontalement qui, sous l'effet des muscles, jouent un rôle de valve vis-à-vis de l'air des poumons; elles libèrent un flux d'air vers la partie supraglottique.

Cet air, expulsé, traverse alors les cordes vocales qui entrent ou non en action pour produire un voisement(sons quasi périodique).

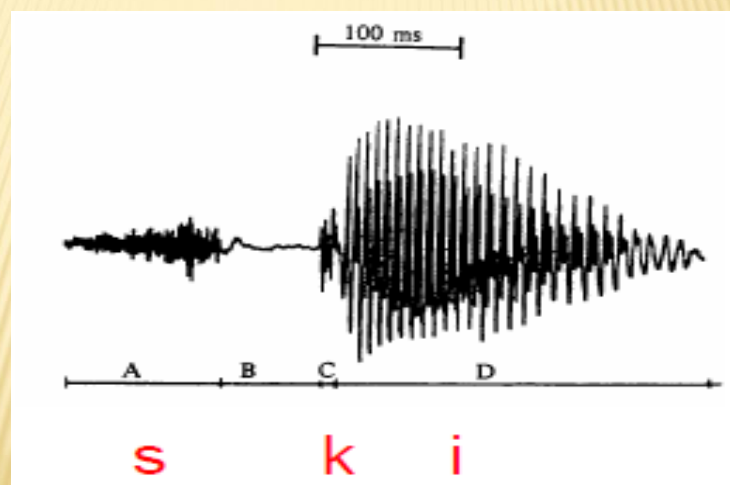
## **PARTIE SUPRAGLOTTIQUE OU CONDUIT VOCAL**

- Cavités Orales (pharyngienne et buccale) à géométrie variable en fonction des articulateurs:
  - Langue, mâchoire inférieure, lèvres
- Cavités nasales à géométrie fixe peuvent être couplées aux cavités orales par abaissement du voile du palais

# ANATOMIE



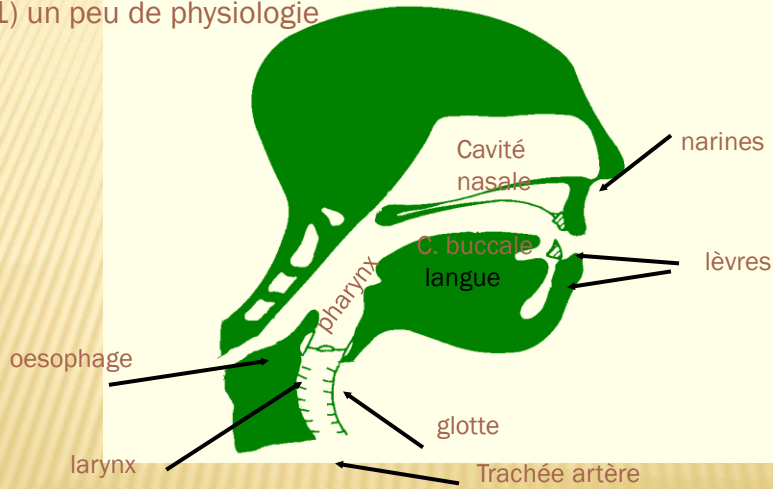
# EXEMPLE





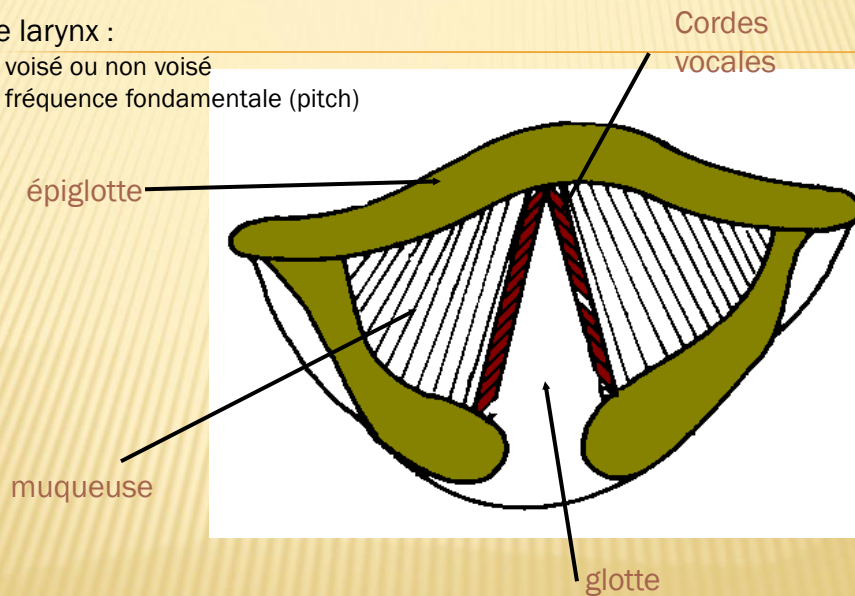
# I-PRODUCTION NATURELLE DE LA PAROLE

## 1) un peu de physiologie



## Le larynx :

- voisé ou non voisé
- fréquence fondamentale (pitch)



## CONDUIT VOCAL NON STATIONARITÉ DE LA PAROLE

- ✘ Le signal vocal n'est pas un signal stationnaire: le conduit vocal se déforme d'une façon continue et les paramètres du modèle sont donc variables dans le temps.
- ✘ Toutefois, les déformations sont suffisamment lentes pour que la modélisation puisse être quasi constants pendant des intervalles de temps de l'ordre de **10 ms**.

## COMMENT ? LE SIGNAL DE PAROLE (AUDIO)

- ✘ La parole se distingue des autres sons par ses caractéristiques acoustiques
- ✘ Les sons de parole sont produits par deux processus Différents
  1. **Vibration des cordes vocales**  
Source de voisement
  2. **Turbulence créée par l'air**  
s'écoulant rapidement dans une constriction du conduit vocal lors de relâchement d'une **occlusion** du conduit vocal
- ✘ c'est une Source de bruit

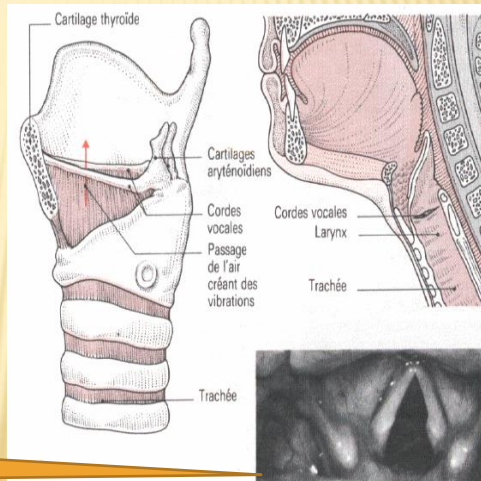
## LES CORDES VOCALES

Les Cordes Vocales sont deux replis situées à la partie supérieure du larynx, tendus entre le cartilage thyroïde en avant et les aryténoïdes en arrière. Elles sont unies et fixes en avant. Mobiles en arrière formant un V ouvert.

Elles se joignent pour vibrer à l'émission des sons.

**Pour la production de la voix**, les cordes vocales sont rapprochées l'une contre l'autre par des muscles qui agissent sur les cartilages aryténoïdiens.

Cordes Vocales au repos



### Les cordes vocales en action (1)



## LA PRODUCTION DE LA PAROLE

✘ La production de la parole est contrôlée linguistiquement à la fois au niveau de l'**excitation (source)** et des **résonances (conduit)**.

### ✘ Sources

- ✓ Source quasi-périodique: cordes vocales en vibration
- ✓ Source de bruit: écoulements turbulents produits par une obstruction du passage de l'air dans le conduit vocal

### ✘ Résonances

Les articulateurs façonnent dans le conduit vocal des cavités dont les fréquences de résonance déterminent la forme **du spectre**.

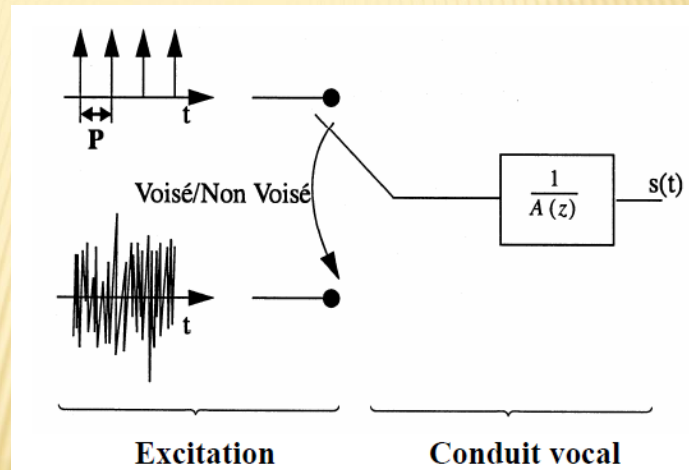
## L'APPAREIL PHONATOIRE

✘ **Le fonctionnement de l'appareil phonatoire explique les caractéristiques du signal vocal**

✘ Les observations précédentes conduisent à modéliser la phonation par le filtre  $1/A(z)$  ci-dessous dont l'entrée est soit :

1. une suite d'impulsions périodiques (cas voisé)
2. un bruit blanc (cas non voisé)

## MODÈLE SOURCE FILTRE



## COMPLEXITÉ DU SIGNAL DE PAROLE

- ✘ Signal résultant des interactions entre production, et compréhension
- ✘ Continuité du signal, coarticulation (pas de coupure de mots)
- ✘ Variabilité (inter- et intra-locuteurs, conditions acoustiques)

## TAXONOMIE DES SONS

- ✘ La taxonomie des sons est définie de deux manières, grâce à la phonétique et à la phonologie.
- ✘ La phonétique peut être considérée comme véritablement descriptive, associant chaque son de la langue à un symbole et à une classe.
- ✘ La phonologie s'intéresse à la description des interdépendances entre sons et au codage effectif des mots du langage lors du processus d'oralisation.
- ✘ La phonologie essaie donc plus d'expliquer les différences qui peuvent exister entre la transcription phonétique d'un mot du langage et la transcription phonétique exacte du mot qui est effectivement prononcé.

## PHONÉTIQUE

- ✘ Les sons produits par le système phonatoire humain peuvent être rattachés à différentes classes.  
Ces classes permettent de regrouper les sons selon leurs principales caractéristiques qui sont facilement identifiables .  
La subdivision des sons en éléments de granularités variables et la division de l'ensemble de ces sons, ou phonèmes, en classes distinctes, est à l'origine de la constitution d'alphabets phonétiques qui caractérisent des langues différentes.

## CLASSES PHONÉTIQUES

- ✘ Les différents sons de la parole sont regroupés en classes phonétiques en fonction de leurs caractéristiques principales. Ces caractéristiques représentent des différences qui sont suffisamment importantes pour qu'il soit possible de classer les différents sons visibles sur un spectrogramme selon leur classe respective en très peu de temps et sans aucune écoute de la phrase correspondante [lonchamp90].

## CLASSES PHONÉTIQUES

- Le langage parlé est constitué de phonèmes
- Il y a des phonèmes voisés et des phonèmes non voisés.
- Les phonèmes voisés sont quasi-périodiques et leur spectre présente des résonances et des anti résonances
- L'analyse du signal vocal avec la FFT impose des durées allant de 10 à 30 milliseconde.

## PHONÈMES

- ✘ **Phonème (notion phonologique)** – unité minimale du langage
- ✘ porteuse d'une signification linguistique. élément sonore du langage articulé.
- ✘ • **Allophone – réalisation**  
d'un phonème (un phonème peut avoir plusieurs allophones)
- ✘ Les phonèmes sont les éléments sonores les plus brefs qui permettent de distinguer différents mots
- ✘ Exemples [p] [b]  
pas / bas  
paie / baie  
pot / beau

## LES CLASSES DE PHONÈMES

- ✘ Consonnes exemples sons
- ✘ Occlusives
- ✘ Nasales
- ✘ Glissantes
- ✘ Liquides
- ✘ Fricatives
- ✘ Plosifs
- ✘ Voyelles (orales- fermées)



## LES VOYELLES

- ✘ Les différentes classes phonétiques présentes en français et en anglais sont :
- ✘ Les voyelles : Elles se caractérisent principalement par le voisement qui crée des formants. Ces formants, qui sont des zones fréquentielles de forte énergie, correspondent à une résonance dans le conduit vocal de la fréquence fondamentale produite par les cordes vocales.
- ✘

## LES PHONÈMES

Afin de lire à voix haute un texte , un ordinateur devra enchaîner des phonèmes, et leur donner une intonation.

- ✘ la langue française contient 37 phonèmes, la langue anglaise plus de 40 phonèmes, ...
  - Les phonèmes voisés sont quasi périodiques, au contraire des phonèmes non voisés.
  - En réalité, les phonèmes non voisés n'utilisent pas les cordes vocales, comme quand on siffle et qu'on utilise la bouche comme une cavité résonante :
    - C'est le cas de certains phonèmes, 's' et 'f' entre autres,
    - C'est aussi le cas quand on chuchote.

# LES PHONÈMES DU FRANÇAIS

symbole phonétique	exemple en langue française	classe	phonétique
a	plat		voyelles
ɑ	mât		
i	pile		
y	rue		
ɔ	bol		
o	pôt		
ə	le		
ɛ	lait		
e	blé		
ø	peu		
œ	heure		
u	roue		
ɑ̃	blanc		voyelles nasales
ɔ̃	bon		
ẽ	lin		
œ̃	brun		
ʃ	hier		semi-consonnes
ɥ	huit		
w	oui		
l	lent		liquides
R	rue		
m	masse		nasales
n	nous		
ɲ	signal		

f	fer	sourdes	fricatives
s	assis		
ʃ	chou		
v	verre	sonores	
z	Asie		
ʒ	joue		
p	passe	sourdes	occlusives
t	toux		
k	cou		
b	basse	sonores	
d	doux		
g	goût		

## PHONÈMES EN ANGLAIS

phonème API	phonème ARPABET	exemple en langue anglaise	classe phonétique
<i>i</i>	IY	beat	voyelles
<i>I</i>	IH	bit	
$\epsilon$	EH	bet	
$\text{æ}$	AE	bat	
<i>a</i>	AA	bob	
$\text{ɔ}$	AO	bought	
$\text{ʊ}$	UH	book	
<i>u</i>	UW	boot	
$\Lambda$	AH	but	

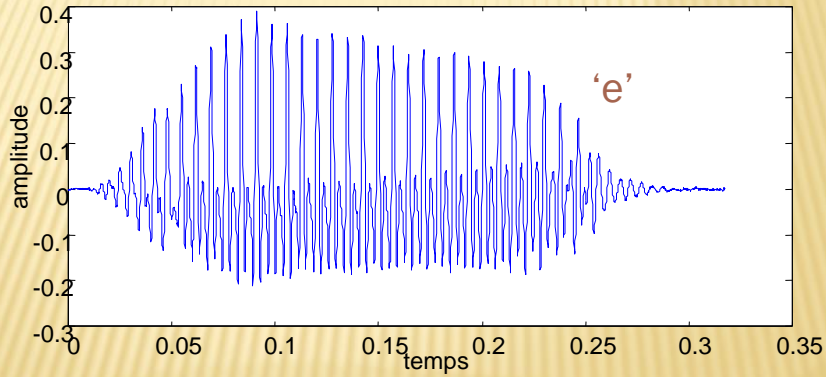
### Les cordes vocales en action (1)



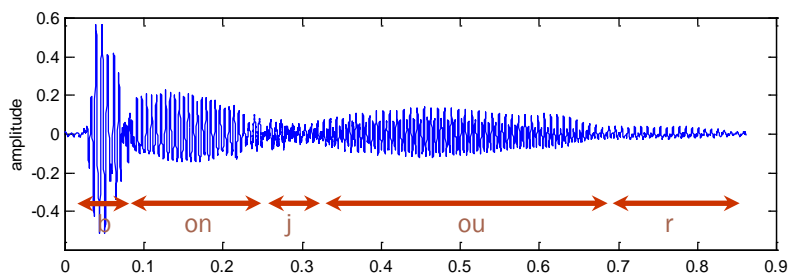
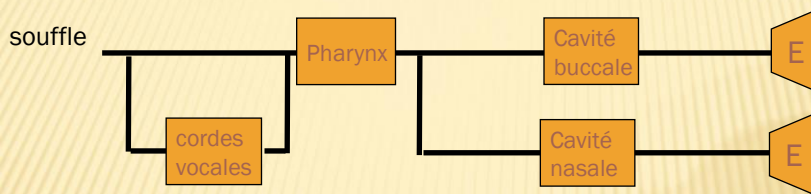
## 2) Le rôle des cordes vocales : sons voisés

Un son voisé est défini par :

- sa fréquence fondamentale (=hauteur)
- son timbre = rapport entre fondamental et harmonique



Représentation simplifiée :



#### 4) Les différents types de sons (phonèmes)

##### Les voyelles (voisées)

###### Orales

[A, E, I, O, U, OU...]



###### Nasales

[IN, UN, AN, ON]



##### Les consonnes

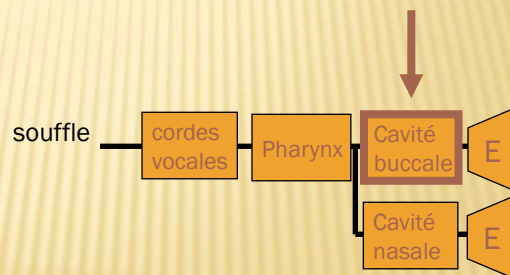
###### Liquides

[R,L]



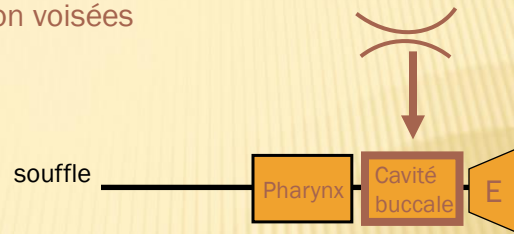
###### Nasales

[M,N,GN]



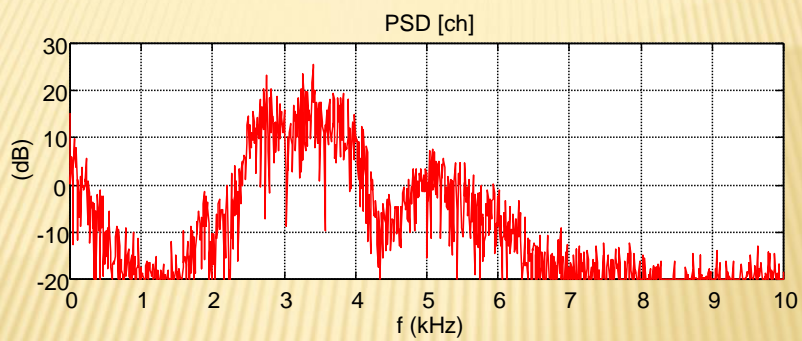
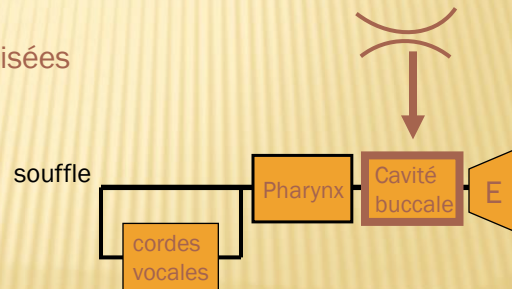
## Fricatives non voisées

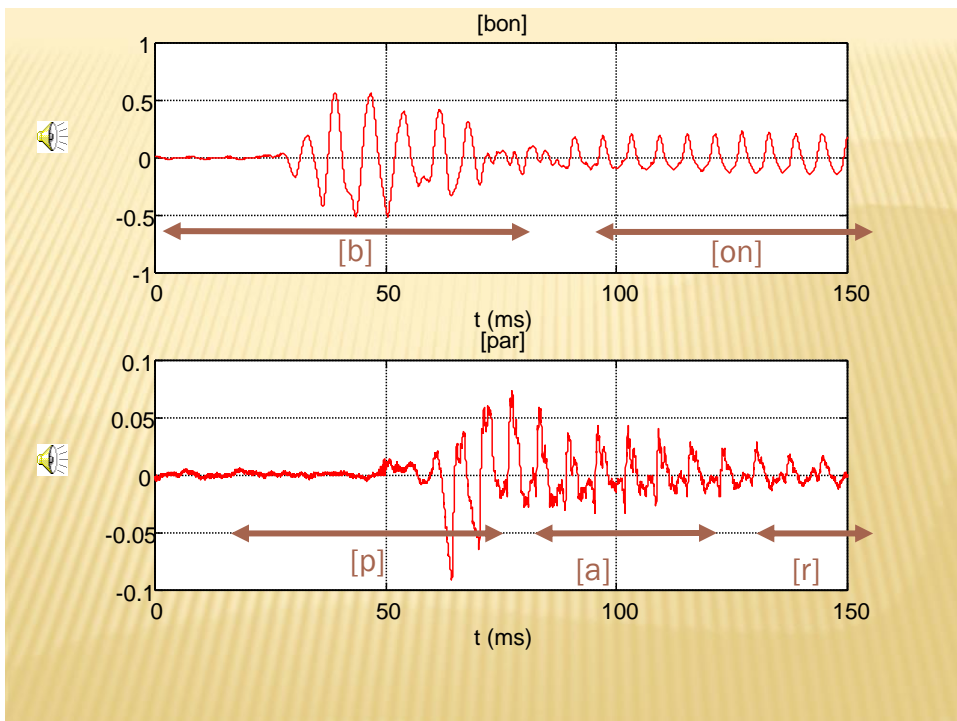
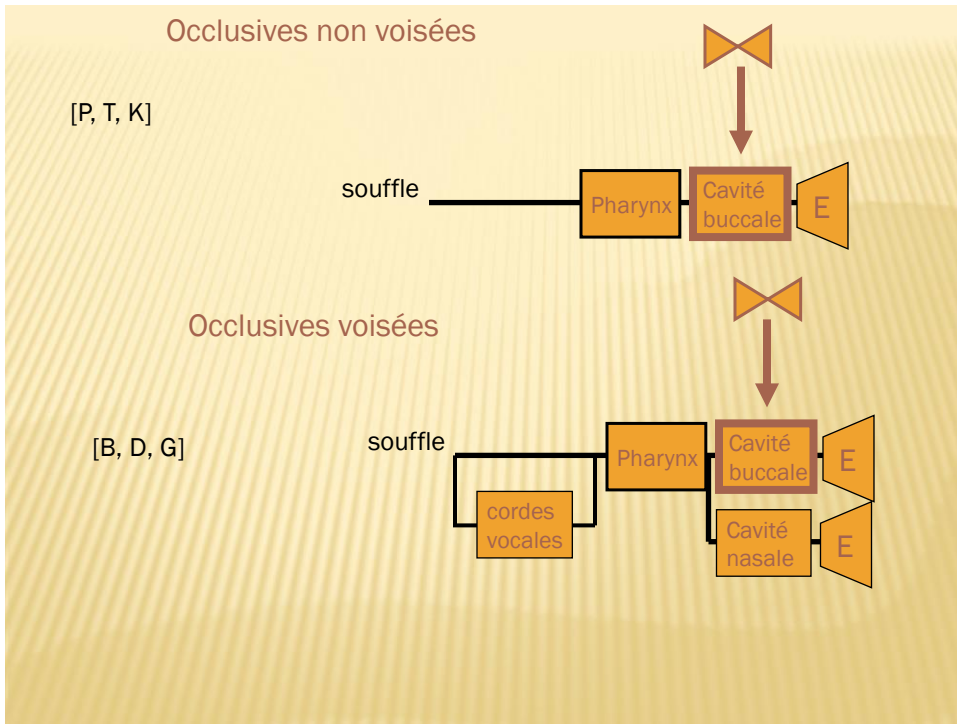
[F, S, CH]



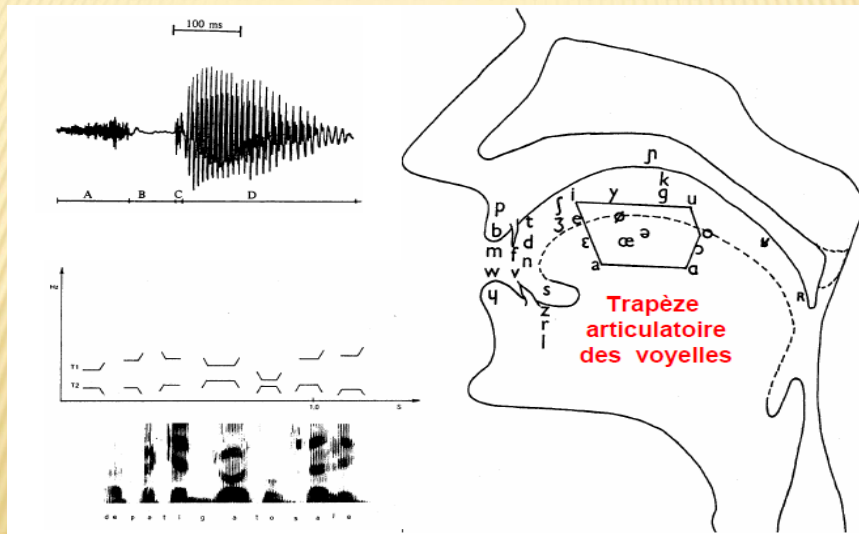
## Fricatives voisées

[V, Z, J]



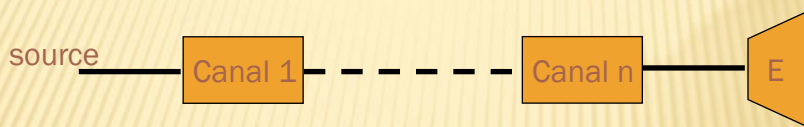


## TRAPÈZE ARTICULATOIRE DES VOYELLES



### 3) Modélisation ARMA de la parole

- représentation schématique du signal de parole :



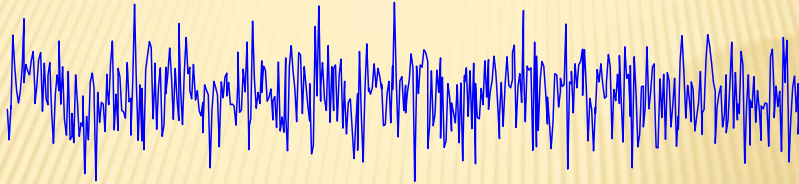
la source peut être un bruit blanc (sons non voisés ou chuchotés) ou 1 train périodique pour les sons voisés.

le conduit vocal peut être représenté par une succession de tubes acoustiques (pharynx, cavité buccale, nasale) modélisés par des résonateurs.

l'émetteur, lèvres ou narines, représente une charge acoustique pour l'onde sonore qui y arrive.



- source pour les sons non voisés



## LES OCCLUSIVES

- ✗ les occlusives : les phonèmes de cette classe se caractérisent oralement par la fermeture du conduit vocal, fermeture précédant un brusque relâchement.
- ✗ Les occlusives sont donc constituées de deux parties successives : une première partie de silence, correspondant à l'occlusion effective, et une deuxième partie d'explosion, au moment du relâchement.
- ✗ Les occlusives peuvent être voisées, à la manière des voyelles, ou sourdes, c'est à dire non voisées.
- ✗ Les occlusives voisées peuvent également être appelées occlusives sonores.

## LES FRICATIVES

- ✘ Les fricatives : dans cette classe sont regroupés les sons produits par la friction de l'air dans le conduit vocal lorsque celui-ci est rétréci au niveau des lèvres, des dents ou de la langue.
- ✘ Cette friction produit un bruit de hautes fréquences et peut être voisée ou sourde.

## LES SONNANTES

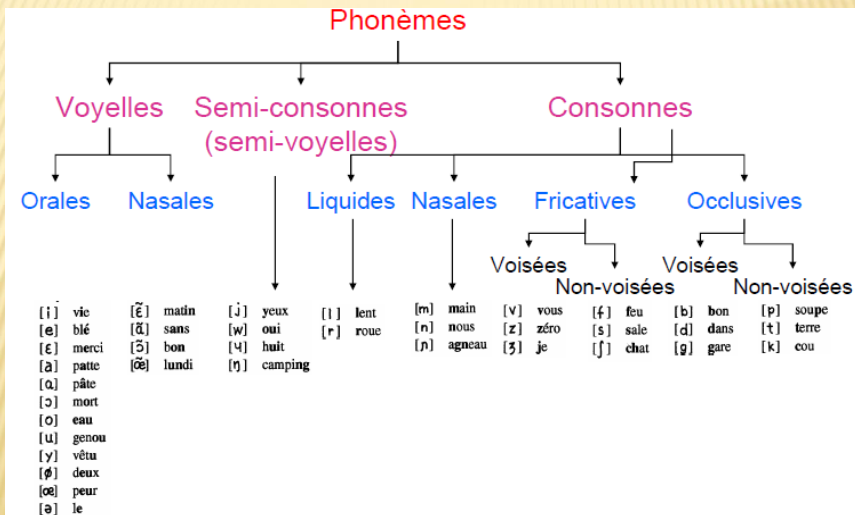
- ✘ Les sonantes : cette classe est en fait constituée du regroupement des trois sous-classes que sont:
- ✘ Les semi-consonnes, Les liquides , Les nasales.
- ✘ Les semi-consonnes (ou semi-voyelles ou glissantes) : elles ont la structure acoustique des voyelles mais ne peuvent en jouer le rôle car elles ne sont que des transitions vers d'autres voyelles.
- ✘ D'un point de vue syntaxique, une règle stricte de la langue française veut que deux voyelles ne puissent jamais se suivre. Cette règle est très largement respectée dans la construction des mots mais présente, comme toute règle, quelques exceptions.
- ✘ La classe des semi-consonnes a été créée pour pallier ces exceptions . Les semi-consonnes sont évidemment sonores.

## LES LIQUIDES ET NASALES

Les liquides : Les liquides sont très similaires aux voyelles et aux semi-consonnes mais leur durée et leur énergie sont généralement plus faibles. Elles sont sonores.

- ✗ Les nasales : les phonèmes sont formés par passage de l'air dans le conduit vocal depuis les cordes vocales. Ce passage exclut normalement toute connexion du conduit normal, le conduit buccal, avec le conduit nasal. Ce dernier peut cependant être employé, dans un nombre limité de cas puisque sa physiologie ne permet pas de créer des sons autrement qu'en modifiant le volume de la caisse de résonance qu'il constitue par l'intermédiaire de la langue, faisant occlusion dans le conduit buccal.
- ✗ Les nasales sont donc produites de la même manière que les occlusives nasales mais l'air n'est pas, cette fois, comprimé dans le conduit vocal. Le vélum est en effet abaissé pour permettre à l'air d'être expiré. Les nasales sont voisées. Il est à noter que certaines voyelles possèdent également un caractère de nasalité.

## LES PHONÈMES DE LA LANGUE FRANÇAISE

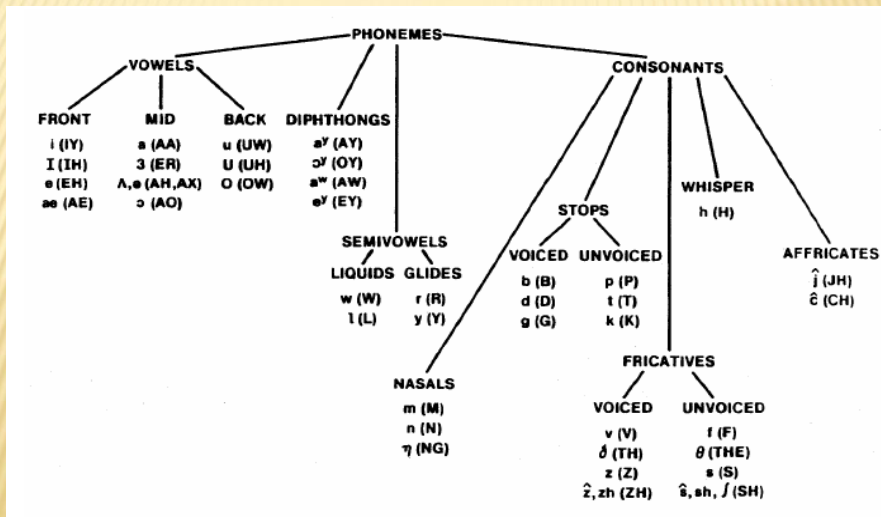


$\alpha^y$	AY	my	diphthongues
$\text{ɔ}^y$	OY	boy	
$e^y$	EY	bait	
$o^w$	OW	boat	
$\alpha^w$	AW	down	
$j$	Y	you	semi-voyelles
$w$	W	wit	liquides
$l$	L	let	
$r$	R	rent	
$m$	M	met	nasales
$n$	N	net	
$\eta$	NX	bang	
$h$	HH	hat	fricatives
$f$	F	fat	
$\theta$	TH	thin	
$s$	S	sat	
$\bar{s}$	SH	shut	
$v$	V	vat	
$\delta$	DH	that	
$z$	Z	zoo	
$\bar{z}$	ZH	azure	
$\bar{c}$	CH	church	affriquées
$\bar{j}$	JH	judge	

## LES DIPHTONGUES & LES AFFRIQUÉES

- ✘ Les diphtongues : cette classe phonétique est propre à l'anglo-américain.
- ✘ Les affriquées : cette classe est, elle aussi, propre à l'anglo-américain mais les affriquées peuvent également être observées dans le français québécois.

# LES PHONÈMES EN ANGLAIS AMÉRICAIN



## VOYELLES

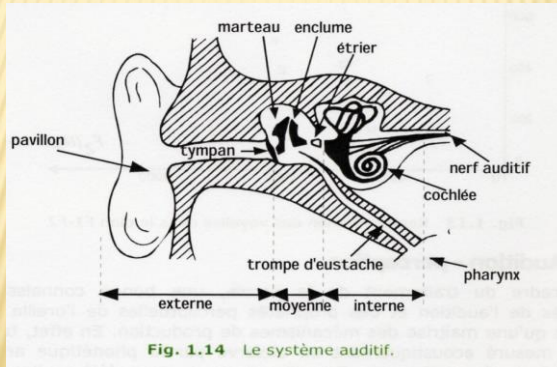
### ✘ Source d'excitation

Unique due à la vibration des cordes vocales (excitation voisée)

- ✘ **Configuration quasi stable** du conduit vocal, Elles peuvent être **orales** (sans couplage avec le conduit nasal) ou **nasales** (conduit vocal couplé avec le conduit nasal).

## L'APPAREIL AUDITIF

- ✘ La fréquence d'échantillonnage  $f_e$  doit être choisie en fonction de l'appareil auditif



## LE SPECTROGRAMME

- ✘ Le spectrogramme est un outil de visualisation utilisant la technique de la transformée de Fourier et donc du calcul de spectres. Il a commencé à être largement utilisé en 1947.
- ✘ Le spectrogramme permet de mettre en évidence les différentes composantes fréquentielles du signal à un instant donné, une transformée de Fourier rapide [aho74] étant régulièrement calculée à des intervalles de temps rapprochés.

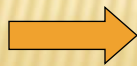
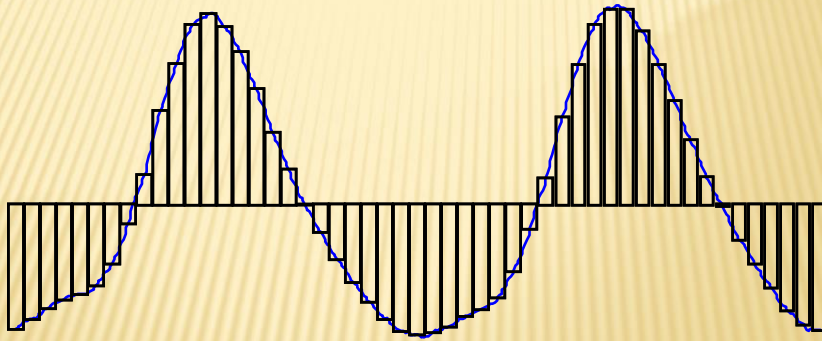
## TP1

- × Goldwave ou praat
- × Enregistrer le signal de parole
- × Fixer la fréquence d'échantillonnage
- × La durée
- × Enregistrer
- × Lecture du signal de parole
- × Wavread(fichier)
- × Visualiser: plot(signal) ou une partie?
- × Visualiser le spectrogram
- × Spectrogram(signal)

## FRÉQUENCE D'ÉCHANTILLONNAGE

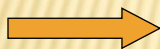
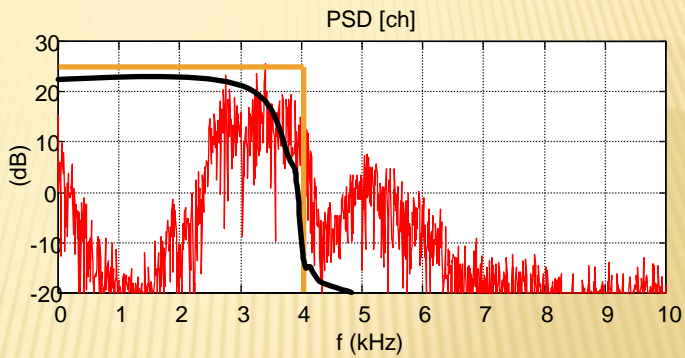
- × Theoreme de Shannon
- × La fréquence d'échantillonnage assurant un non repliement du spectre doit être supérieure à 2 fois la fréquence haute du spectre du signal analogique.
- ×  $F_{ech} = 2 * F_{max}$

# I-L'ÉCHANTILLONNAGE (SAMPLING)



Fréquence d'échantillonnage ???

## Recouvrement spectral



Théorème de Shannon :  $f_{\max} = F_e/2$

- filtre anti-repliement (*anti-aliasing*)



Initial



aliased



prefiltered