

# Outils informatiques d'analyse de corpus

---

**Jean-Yves Antoine**

LI - Université Rabelais de Tours  
Jean-Yves.Antoine@univ-tours.fr

Master SdL — M1 L&R



# Outils informatiques d'analyse de corpus

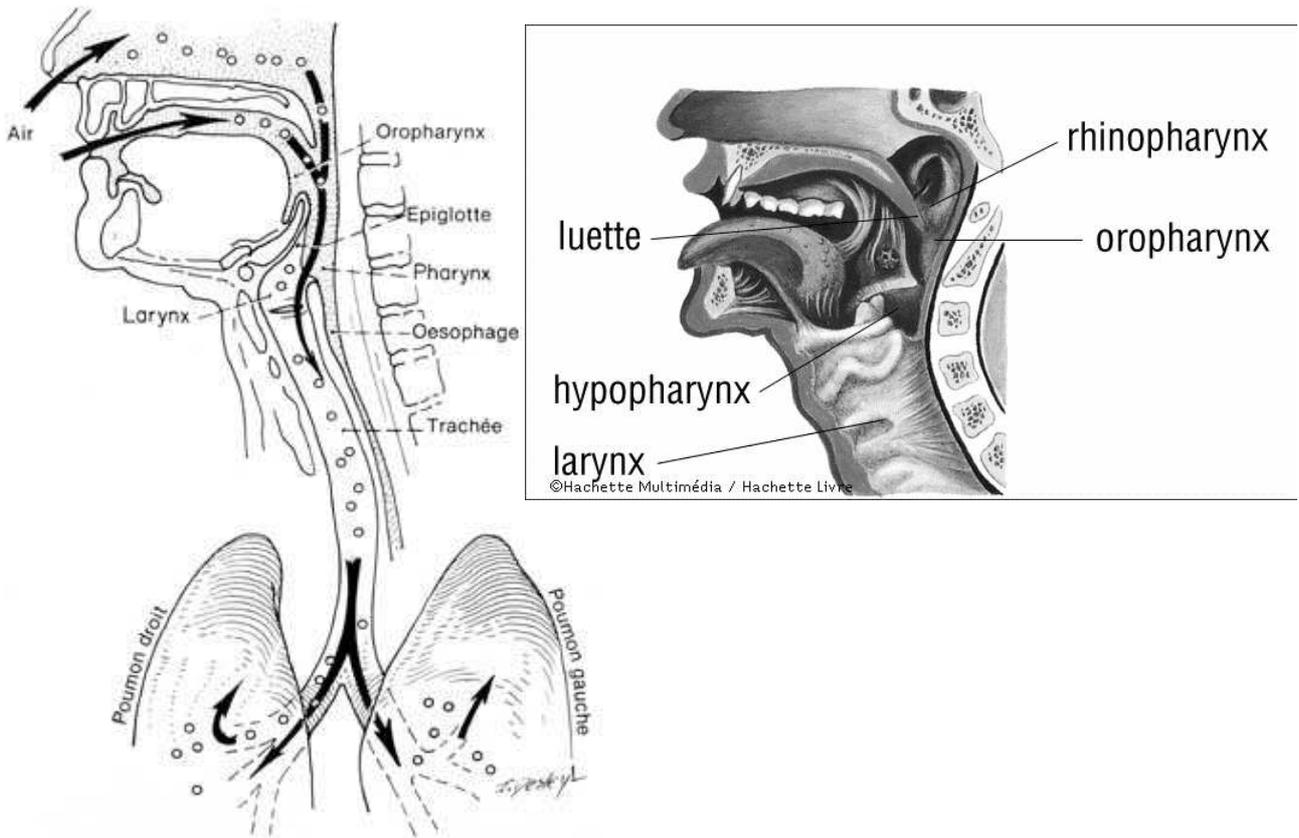
---

**PAROLE**

Master SdL — M1 L&R

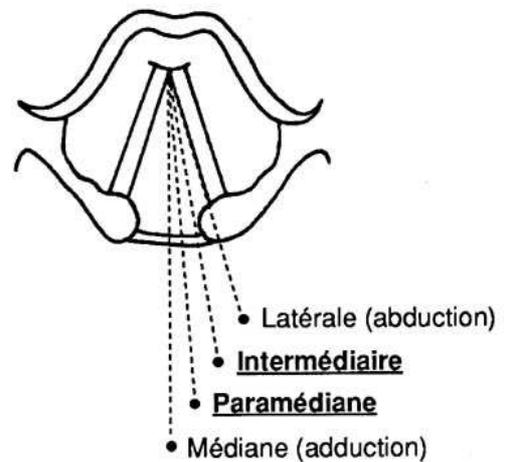
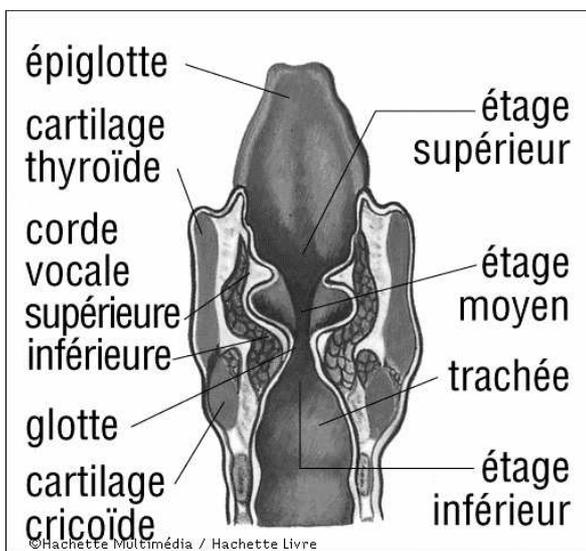


# Production : appareil phonatoire (*rappel*)

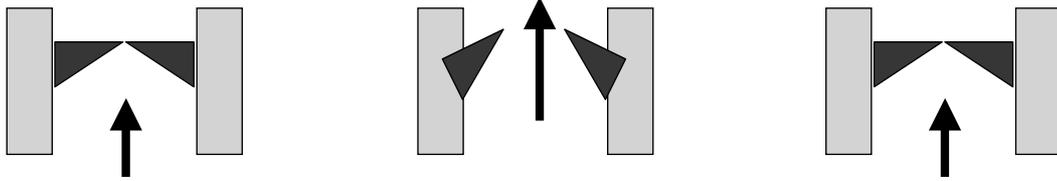


# Production : appareil phonatoire (*rappel*)

- Cordes vocales



# Production : voisement (*rappel*)



- **voisement** : son très énergétique (son **sonores** vs son sourd)
- **fréquence fondamentale** : fréq<sup>ce</sup> de vibration des cordes vocales
- **variabilité de la fréquence fondamentale**
  - hommes 60 - 100 Hz
  - enfants / femmes 200-300 Hz
  - au cours de la production : « programmation » prosodique

*tu manges du poisson* → affirmation  
*tu manges du poisson* ↗ question

# Production : articulation (*rappel*)

- **cordes vocales** : son voisé / non voisé
- **cavité nasale** : son nasal / non nasal
- **langue**
  - haut (fermé) / bas (ouvert)
  - avant / arrière
- **lieu d'articulation**
  - lèvres : son labial
  - dents : son dental, labio-dental
  - palais : son alvéolaire (avant), palatal, vélaire (arrière)

# Phonétique (rappel)

- Classification acoustique des sons de parole

– alphabet phonétique international : API [fɔ̃nɜ̃m]

i	<i>ami</i>	ɛ̃	<i>pain</i>	p	<i>pas</i>	l	<i>bal</i>
e	<i>ôté</i>	ɑ̃	<i>dent</i>	t	<i>tas</i>	ʀ	<i>bar</i>
ɛ	<i>belle</i>	ɔ̃	<i>bon</i>	k	<i>cas</i>	m	<i>rame</i>
a	<i>alanas</i>	œ̃	<i>brun</i>	b	<i>bas</i>	n	<i>panne</i>
ɑ	<i>bâton</i>			d	<i>dos</i>	ɲ	<i>agneau</i>
ɔ	<i>corps</i>	j	<i>briller</i>	g	<i>gars</i>		
o	<i>auditeur</i>	w	<i>oui</i>	f	<i>fa</i>		
u	<i>loup</i>	ɥ	<i>huile</i>	s	<i>sa</i>	h	<i>hop</i>
y	<i>utile</i>			ʃ	<i>chat</i>		
ø	<i>deux</i>			v	<i>va</i>		
œ	<i>peur</i>			z	<i>zéro</i>		
ɶ	<i>petite</i>			ʒ	<i>jars</i>		

# Phonétique (rappel)

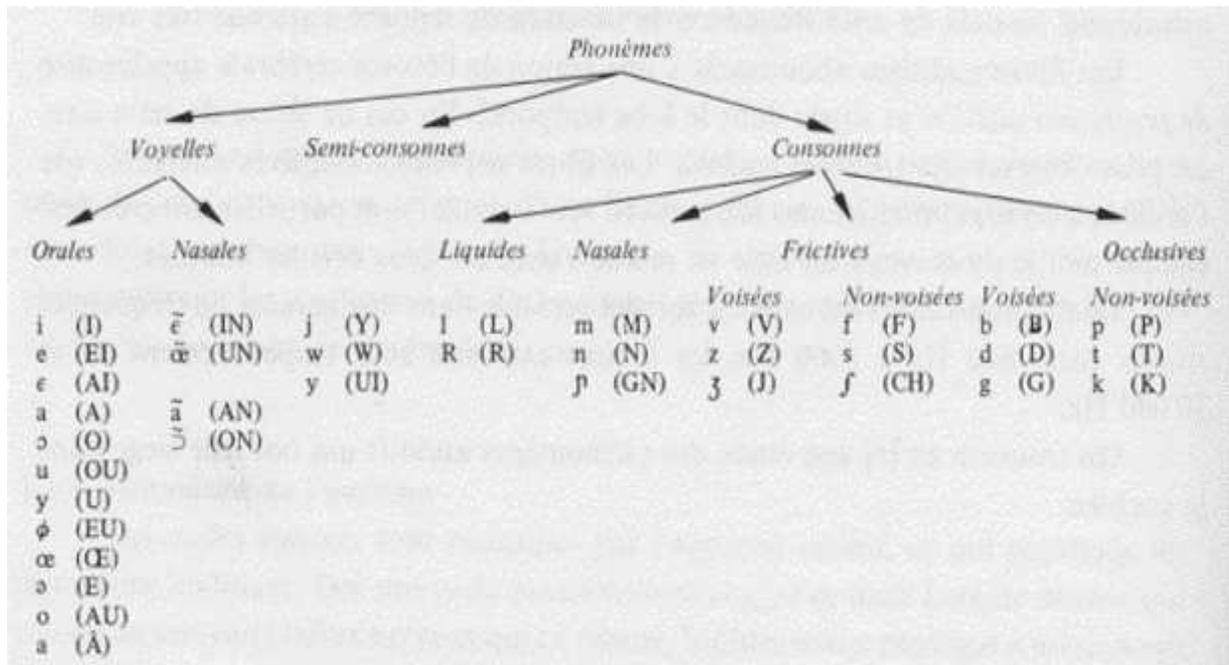
- Classification

i	<i>ami</i>	ɛ̃	<i>pain</i>	p	<i>pas</i>	l	<i>bal</i>
e	<i>ôté</i>	ɑ̃	<i>dent</i>	t	<i>tas</i>	ʀ	<i>bar</i>
ɛ	<i>belle</i>	ɔ̃	<i>bon</i>	k	<i>cas</i>	m	<i>rame</i>
a	<i>alanas</i>	œ̃	<i>brun</i>	b	<i>bas</i>	n	<i>panne</i>
ɑ	<i>bâton</i>			d	<i>dos</i>	ɲ	<i>agneau</i>
ɔ	<i>corps</i>	j	<i>briller</i>	g	<i>gars</i>		
o	<i>auditeur</i>	w	<i>oui</i>	f	<i>fa</i>		
u	<i>loup</i>	ɥ	<i>huile</i>	s	<i>sa</i>	h	<i>hop</i>
y	<i>utile</i>			ʃ	<i>chat</i>		
ø	<i>deux</i>			v	<i>va</i>		
œ	<i>peur</i>			z	<i>zéro</i>		
ɶ	<i>petite</i>			ʒ	<i>jars</i>		

**voyelles**                      **semi-voyelles**                      **consonnes**                      **liquides**

# Phonétique (*rappel*)

- Récapitulatif



Outils Linguistiques — Master L&R — © J.Y. Antoine — 9

# Phonétique : voyelles (*rappel*)

- Petite typologie des langues [N. Vallée, ICP, Grenoble]

- étude sur 317 langues

- **nombre de voyelles**

24 voyelles	1 langue
2 voyelles	2 langues
5 voyelles	23 %
6 voyelles	13 %

- **voyelles orales**

[ i ]	99% des langues
[ a ]	98 % des langues
[ u ]	94 % des langues
[ o ]	44 % des langues
[ e ]	40 % des langues

- **voyelles nasales** : langues (22%) utilisant plus de 9 voyelles
 

[ ã ]	20 % des langues
-------	------------------

- **français** variable suivant la région: *brun* vs. *brin*

Outils Linguistiques — Master L&R — © J.Y. Antoine — 10

# Phonétique (*rappel*)

- **Spécificités idiomatiques**

possibilités articulatoires communes où piochent les langues

- [x]

*c'hoar*, *Crac'h*  
*nach*, *doch*  
*bajo*, *mejor*, *Jorge*



- [h]

*dilhad*  
*caballo*



- [ŋ]

*song*  
*klang*  
*parking*



- [ɛ̃]

*den*



# Phonologie : phonèmes (*rappel*)

- **Interface entre phonétique et linguistique**

- **Phonèmes**

- plus petite unité phonique fonctionnelle *i.e.* distinctive d'un point de vue sémantique
- un phonème = plusieurs réalisations articulatoires (**allophones**)
- notation API entre barres inclinées     /fɔ̃nɛm/

- **Allophones**

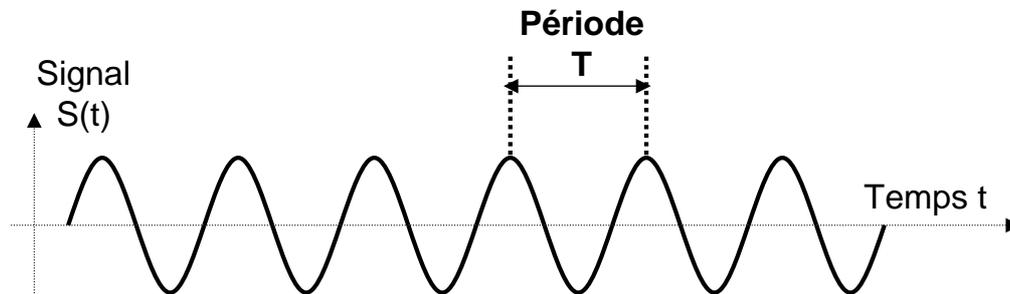
- roulement du [R] dans le français provençal
- phénomènes de **co-articulation**, assimilation, réduction

médecin	/medsɛ̃/	⇒	[medəsɛ̃]	[metsɛ̃]
il y a	/ilija/	⇒	[ilja]	[ja]

# Analyse du signal de parole

## Son périodique « pur » (sinusoïdal)

- Reproduction du même signal suivant un période **T** constante

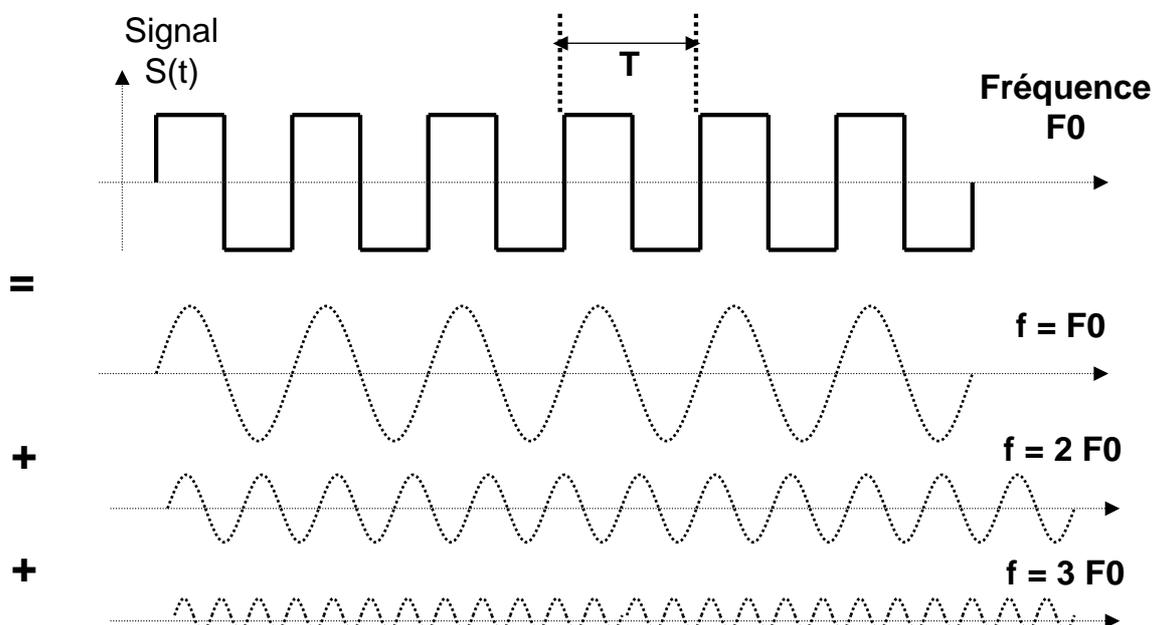


- Une seule composante fréquentielle : **fréquence fondamentale F0**
- **Fréquence** : nombre de périodes par seconde (Hertz)  $F = 1/T$ 
  - ✓  $T = 1 \text{ s}$                       alors  $F = 1 \text{ Hz}$
  - ✓  $T = 0,01 \text{ s}$                     alors  $F = 100 \text{ Hz}$
  - ✓  $F = 1000 \text{ Hz}$                   alors  $T = 1 \text{ ms}$

# Analyse du signal de parole

## Son périodique quelconque

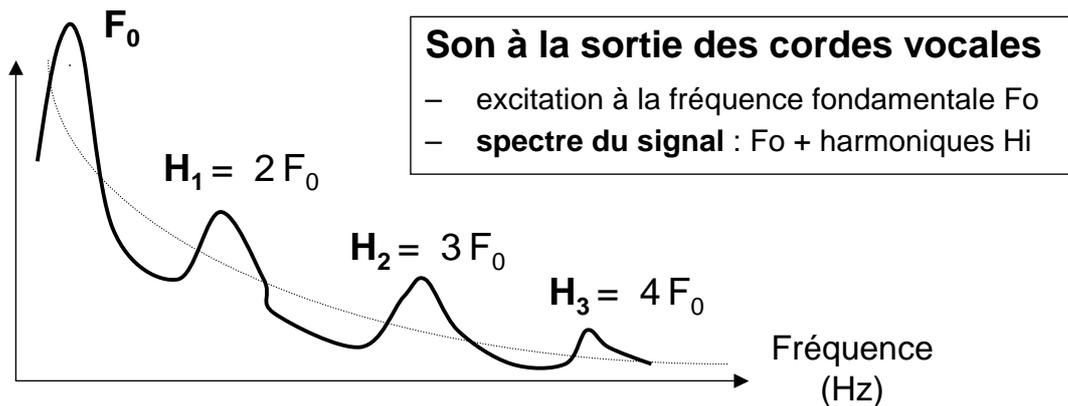
Correspond à la somme de signaux sinusoïdaux de fréquences multiples de la fréquence fondamentale et d'amplitude décroissante (séries de **Fourier**)



# Analyse du signal de parole

## Son périodique quelconque

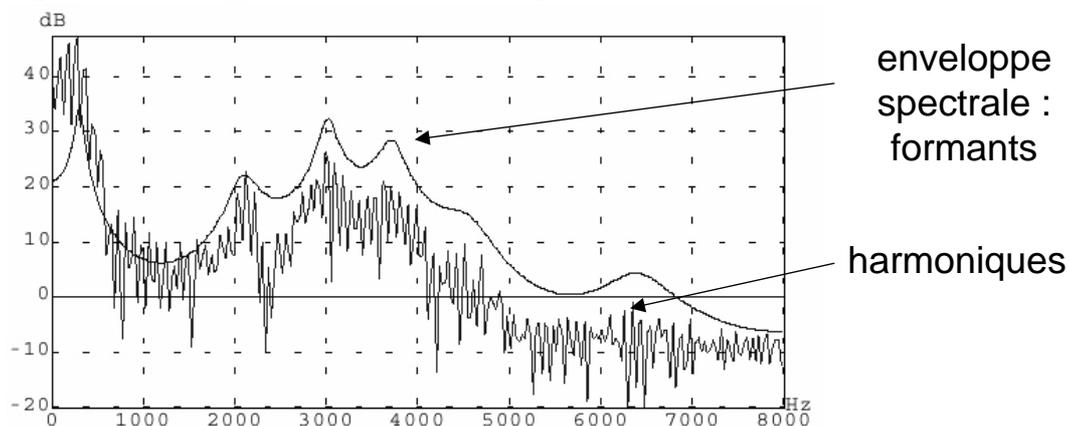
- Plusieurs composantes fréquentielles correspondant à la fréquence de chaque sinusoïde de l'analyse de Fourier
- Fréquences multiples de la fondamentale : **harmoniques**
- **Spectre** d'un signal à un instant donné : représentation de l'importance des différentes composantes fréquentielles



Outils Linguistiques — Master L&R — © J.Y. Antoine — 15

## Analyse du signal de parole : formants

- **Articulation** :
  - lieux de constriction : cavités de résonance
  - fréquences de résonance dues aux articulateurs : **formants**
  - **spectre du signal** :  $F_0$  + harmoniques + formants  $F_i$

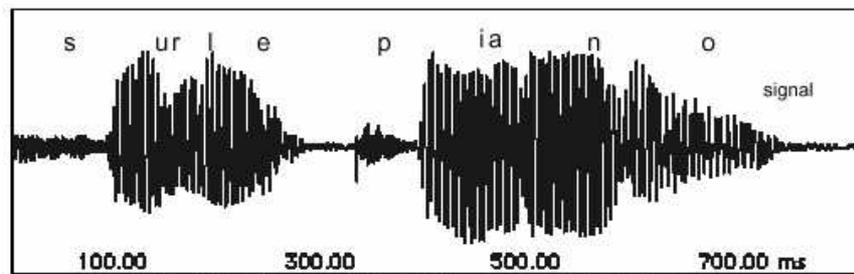


**Structure formantique caractéristique de chaque son**

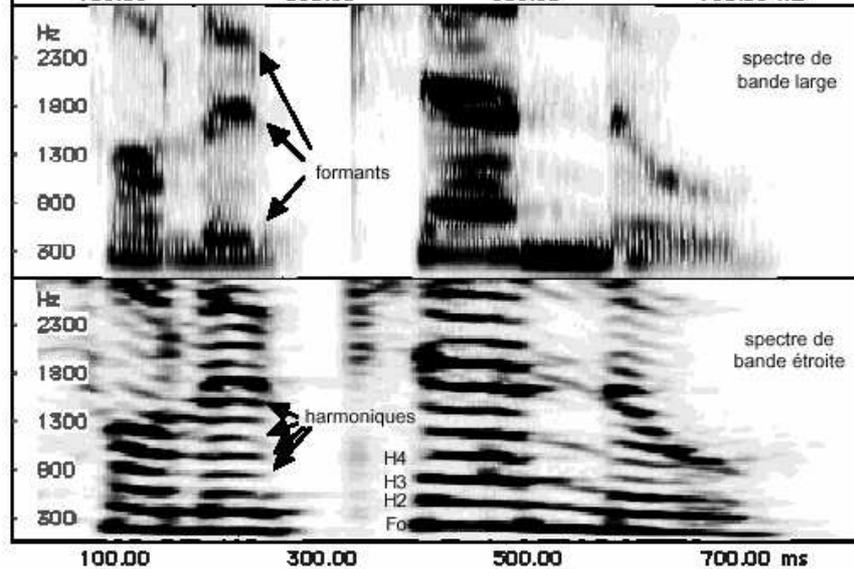
Outils Linguistiques — Master L&R — © J.Y. Antoine — 16

# Analyse du signal de parole

Énergie



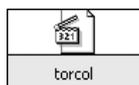
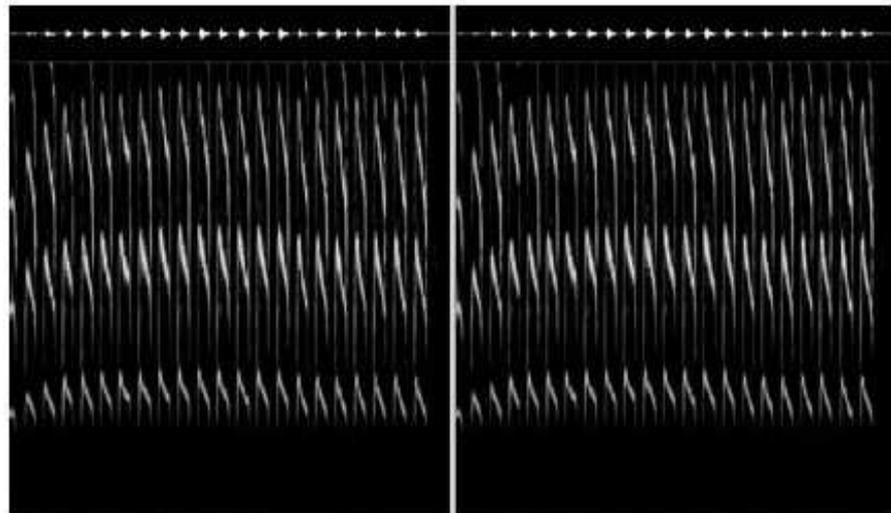
Spectrogramme



d'après [Keller 97]

# Son non articulé

Chant du *Torcol fourmilier* (*Jynx Torquilla*)

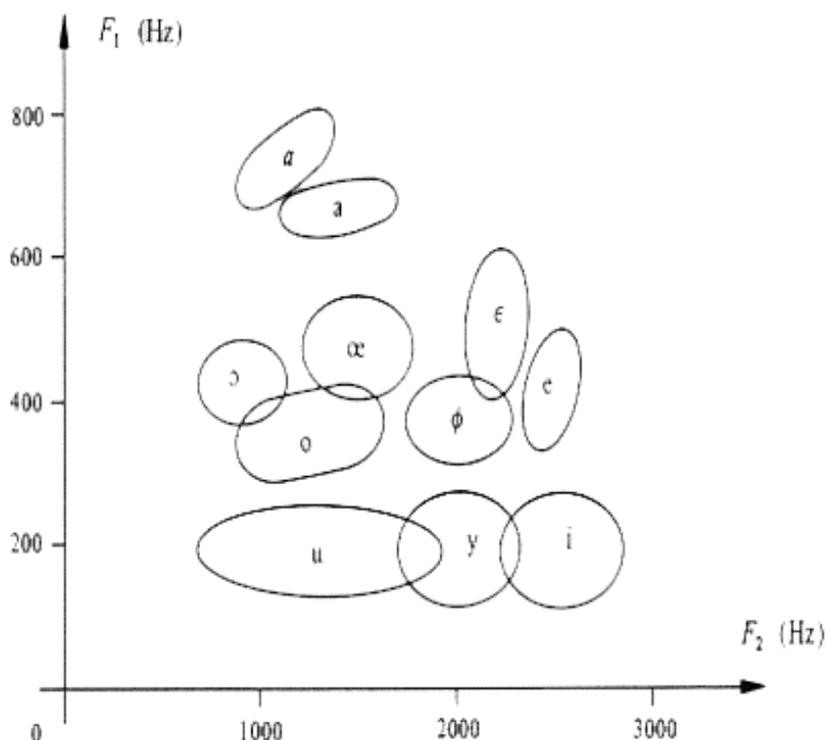


# Signal de parole : voyelles

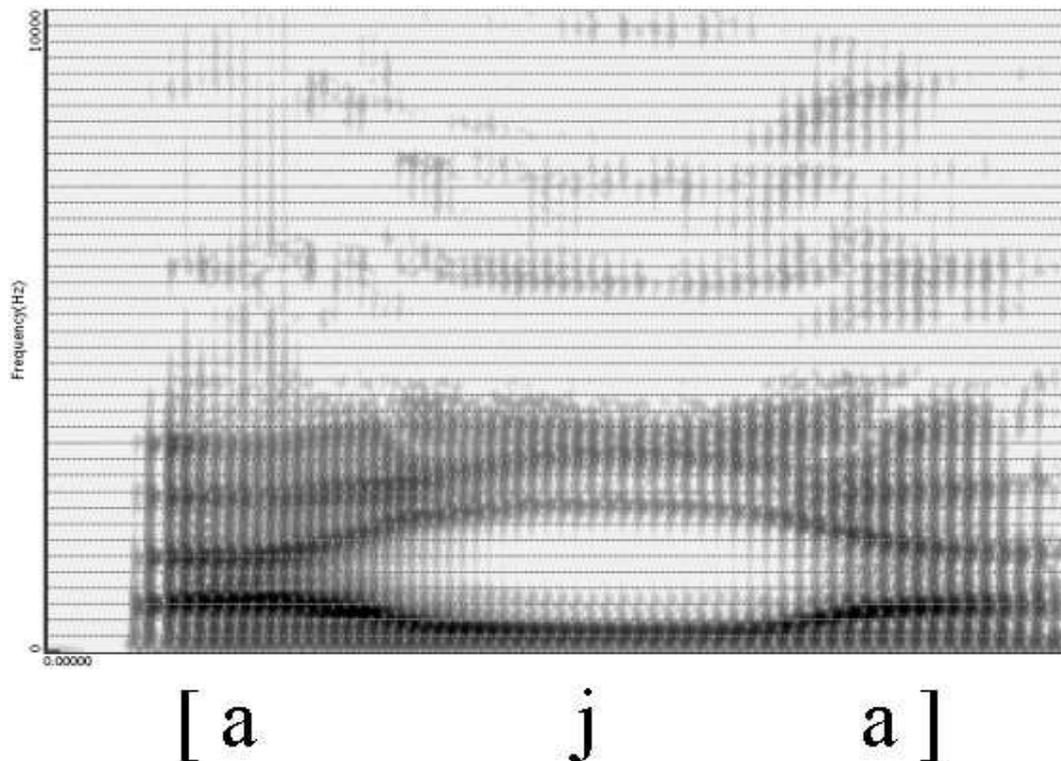
**Voyelles** : voisement + conduit vocal ouvert (sons très énergétiques)

- **voyelles orales** : français : [ i e ɛ a ɔ o u y ø œ ə ]
- **voyelles nasales** : toute voyelle orale peut être nasalisée  
français : [ ẽ ã õ œ̃ ]
- **position de la langue** : antérieure / moyenne / postérieure  
*aperture* : fermée (haut) / ouverte (bas)  
↳ **formants F1 et F2**

# Signal de parole : voyelles



# Signal de parole : voyelles



# Signal de parole : consonnes

**Consonne** : constriction lors de la production (son moins énergétique)

- **Nasales** français : [ m n ɲ ]
- **Occlusives (ou plosives)**
  - **son dynamique** : phase d'occlusion (tenue) suivi d'un relâchement brutal de l'air
  - **voisées** : son basses fréquence émis pendant le tenue  
français : [ b d g ]
  - **non voisées** : silence pendant la tenue  
français : [ p t k ]
  - **lieu d'occlusion** palais [ k g ], dents [ t d ], lèvres [ p b ]

# Signal de parole : consonnes

---

- **Fricatives (ou constrictives)**
  - pas d'occlusion mais une constriction au lieu d'articulation
  - **turbulence** : signal non périodique
  - **voisées** : excitation périodique + turbulence  
français : [ v z ʒ ]
  - **non voisées** : turbulence seule  
français : [ f s ʃ ]
  - **lieu d'occlusion** palais [ ʃ ʒ ], dents [ s z ], lèvres [ f v ]

# Signal de parole : « hybrides »

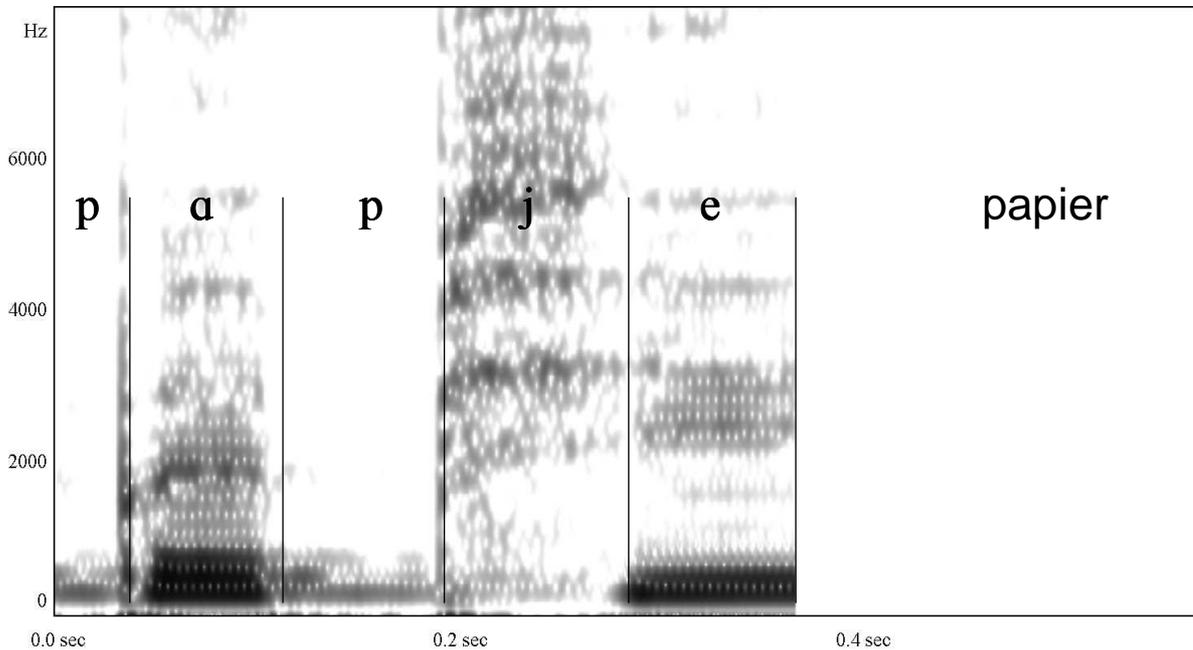
---

voisement + constrictions non complète

- **Semi-voyelles** ou ... semi-consonnes (*glides*)  
français : [ j w ɥ ]
- **Liquides** français : [ l r ]
- **Nasales** constriction assez forte mais dérivation nasale  
français : [ m n ɲ ]

# Lecture de spectrogramme

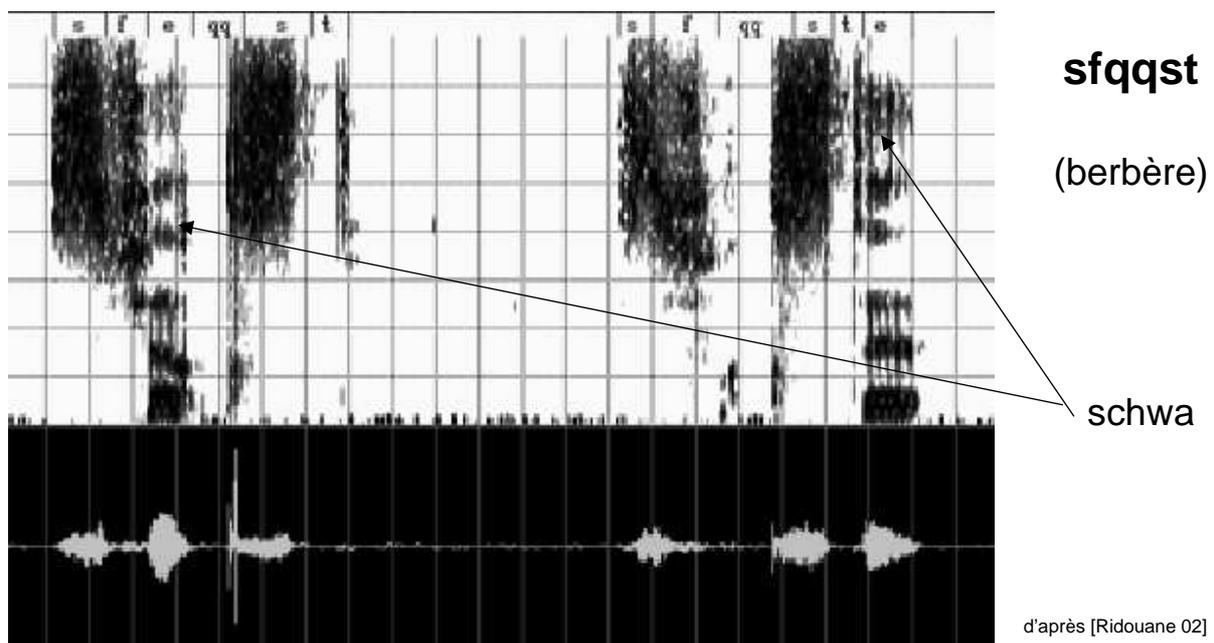
- occlusive : tenue + barre de plosion
- voyelle : structure formantique énergétique
- semi-voyelle : formants moins nets + bruit haute fréquence (constriction)



Outils Linguistiques — Master L&R — © J.Y. Antoine — 25

# Lecture de spectrogramme

- fricative : bruit haute fréquence (constriction)



d'après [Ridouane 02]

Outils Linguistiques — Master L&R — © J.Y. Antoine — 26

# Au delà des classes phonétiques...

## Autres mécanismes spécifiques à certains idiomes

- **Durée du phonème** (20 % des langues)



*kado*  
*kaado*

*coin*  
*carte*

- **Consonne roulée**



*pero*  
*perro*

*mais*  
*chien*

↪ Distinction porteuse de sens en espagnol, contrairement au roulement du [R] dans le français provençal.

- **Variation du pitch** (langues tonales : Asie, Afrique)



*ma*

ton élevé

*maman*



forte montée

*imbécile*

faible montée

*cheval*

forte descente

*réprimander*

## Prosodie

- **Rythme**

– durée des phones et des silences

- **Mélodie**

– variations de la fréquence fondamentale (*pitch*)

– exemples :

montée en fin de question intonative

variations en fin de mot, de syntagmes

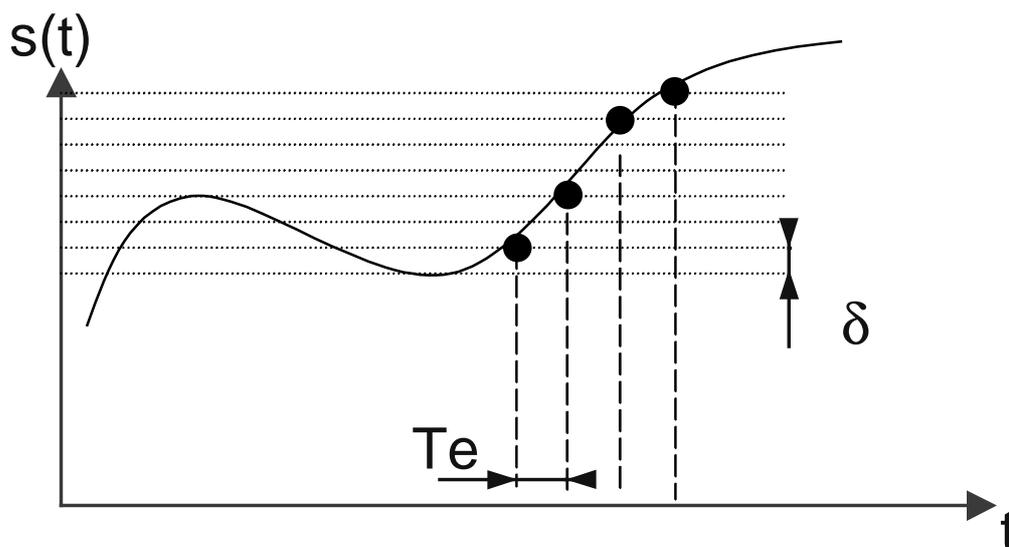
- **Intensité**

– marques d'accentuations (syllabes fortes ou faibles)

- **Quelle unité prosodique ?**

Difficile à définir car plusieurs niveaux de significations (morpho-syntaxique, sémantique, pragmatique) : micro et macro-prosodie

# Numérisation du son



- fréquence d'échantillonnage  $f_e = 1 / T_e$  (1 Hz = 1 échantillon / sec)
- quantification : numérisation sur b bits  $\Rightarrow 2^b$  niveaux de pas de quantification  $\delta$

# Numérisation du son

## • Débit binaire

$$D = f_e \times b$$

Exemple :  $D = 8 \text{ KHz} \times 8 \text{ bits} = 64 \text{ kb/s} = 8 \text{ kO/s}$

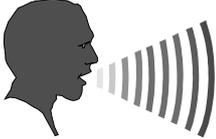
## • Détérioration du signal

- **théorème de Shannon**: l'échantillonnage n'induit aucune perte d'information si  $f_e$  est supérieure ou égale à 2 fois les plus hautes composantes fréquentielles du signal.
- **bruit de quantification** dépendant de  $\delta$  (rapport signal / bruit)

↪ **compromis entre limitation du débit binaire et minimisation de la détérioration du signal** (perte d'information)

# Numérisation du son

---

Application	signal utile	$f_e$	b
	0-3400Hz	8 KHz	8 bits
	0-5000Hz	6 kHz à 16 kHz	8 bits à 16 bits
	0-20 KHz	44.1 kHz (CD Audio) 48 kHz (DAT)	16 bits

# Formats de fichiers sonores

---

## Deux types de formats

- **Formats bruts** (*raw*) — définissent une seule norme de codage : le fichier ne contient que le signal  
exemple            .SND                            (1 canal sur 8 bits)
- **Formats auto-décrits** (*self describing*) — plusieurs normes : signal + entête décrivant ses caractéristiques (fe, type de codage, nombre de canaux...)  
exemples            .AIFF, . WAV, .AU            (compression optionnelle)  
                          .AIDC, .MP3            (compression)

## Formats hybrides

- Multimédia : son + vidéo, son + MIDI  
exemple            .MOV, .QT                    (Quicktime)

# Formats de fichiers sonores

---

- **AIFF (Audio Interchange File Format )**
  - Apple (historiquement)
  - jusqu'à 44,1 Khz en 16 bits
  - format compressé : .AIFF-C ou .AIFC
- **Wave (.Wav )**
  - format natif Microsoft Windows
  - jusqu'à 44,1 Khz en 16 bits
  - compression : ADPCM/IMA 4:1
- **μLaw (.AU )**
  - SUN (historiquement pour UNIX)
  - 8 bits
  - dualité audio limitée
- **QuickTime (.mov, .qt )**
  - rarement utilisé en traitement de la parole

Outils Linguistiques — Master L&R — © J.Y. Antoine — 33

---

# Compression

## CODECs

pilotes de COmpression / DECompression

- **PCM (Pulse Code Modulation)**
  - signal sur 8 ou 16 bits,  $f_e$  entre 8 et 44,1kHz
  - taux de compression 2,5:1
- **ADPCM / IMA (Adaptive Delta Pulse Code Modulation)**
  - mon normalisé : Microsoft, Creative Labs, IMA
  - compression 16 bits vers 4 bits : taux de 4:1
- **MP3**
  - norme MPEG (*Moving Picture Experts Group*)
  - excellent codage perceptif (psycho-acoustique) + suppression des signaux redondants + extraction des fréquences peu audibles
  - taux de compression 8:1
  - **supporté par de nombreux éditeurs de signal**

Outils Linguistiques — Master L&R — © J.Y. Antoine — 34

# Editeurs de signal de parole

---

- **Nombreux éditeurs de signal gratuits ou payants**
- **SFSWin (Speech Filling System)**
  - Gratuitiel (freeware)
  - Mark Huckwale, University College, London
  - <http://www.phon.ucl.ac.uk/resource/sfs>
  - Utilisé dans cet enseignement
- **UNICE (Vecsys)**
- **Winpitch**
- **GoldWave**
  - Shareware
  - GoldWave Inc.
  - Plus adapté au traitement de signal qu'à l'édition proprement dite.

---

## Bibliographie

---

- **Huang X., Acero A., Hon H-W.** (2001) Spoken Language Processing : a guide to theory, algorithm and system development. Prentice Hall, Upper Saddle River, NJ. (chap. 2)
- **Boite R., Boulard H., Dutoit T., Hancq J., Leich H.** (2000) Traitement de la parole. Coll. Electricité. Presses Polytechniques et Universitaires Romandes. Lausanne, Suisse (chap 1)