

Paramétrisation et codage des signaux de parole

Olivier Gillet et Gaël Mahé

Université René Descartes - Paris V

Octobre 2006

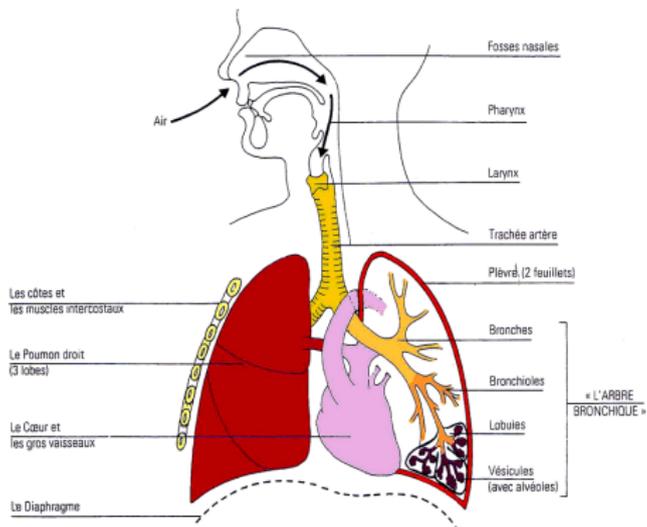
Plan

- 1 Production de la parole**
 - L'appareil respiratoire
 - Les sources vocales
 - Conduit vocal et nasal
 - Phonétique
- 2 Description des signaux de parole**
 - Propriétés des signaux
 - Spectrogramme
 - Formants
 - Analyse prédictive linéaire (LPC)
 - Coefficients cepstraux
- 3 Application au codage de la parole**

Plan

- 1 Production de la parole**
 - L'appareil respiratoire
 - Les sources vocales
 - Conduit vocal et nasal
 - Phonétique
- 2 Description des signaux de parole**
 - Propriétés des signaux
 - Spectrogramme
 - Formants
 - Analyse prédictive linéaire (LPC)
 - Coefficients cepstraux
- 3 Application au codage de la parole**

L'appareil respiratoire



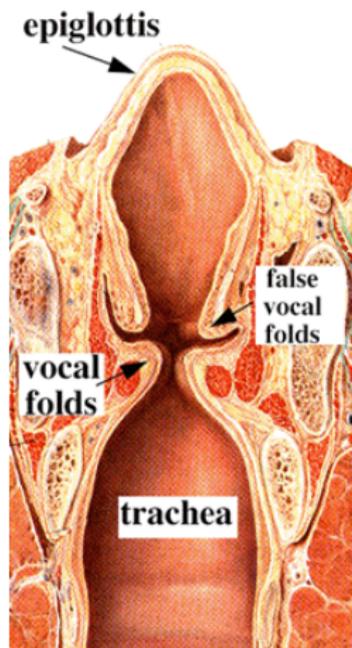
Rôle de "soufflerie"

- Apporte l'énergie nécessaire à la production de la parole.
- Lors de la phonation, l'expiration devient active.

Le larynx

Plusieurs fonctions

- Contrôler le flux d'air pendant la respiration.
- Protéger les voies respiratoires.
- Produire une source sonore.



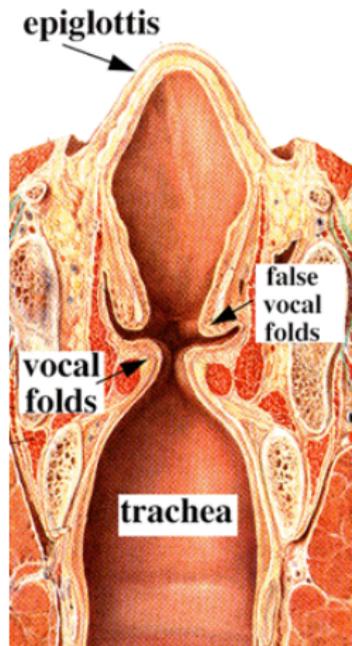
Le larynx

Plusieurs fonctions

- Contrôler le flux d'air pendant la respiration.
- Protéger les voies respiratoires.
- Produire une source sonore.

Cordes vocales

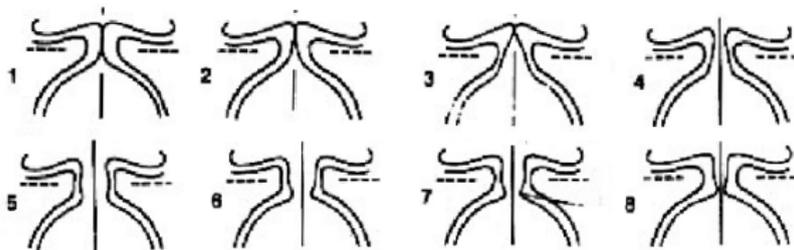
- vibrant rapprochées : sons voisés.
- vibrant écartées : murmure.
- Ecartées : bruit généré dans le conduit vocal.



Mouvement des cordes vocales

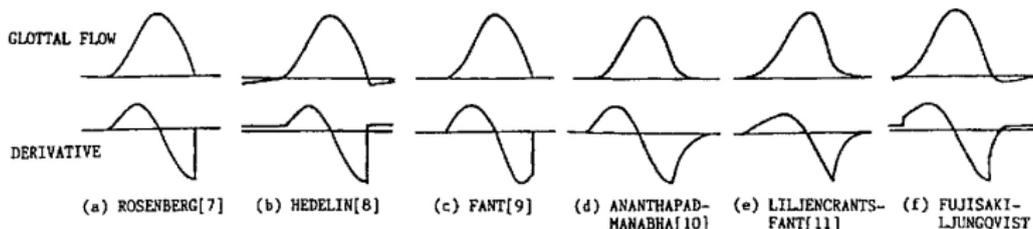


Cordes vocales serrées : sons voisés



Sons voisés : périodiques

- Les cordes vocales se touchent et s'écartent.
- Train périodique d'impulsions (effet "élastique").



Cordes vocales écartées : sons non-voisés

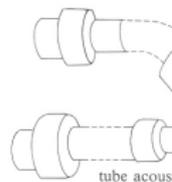
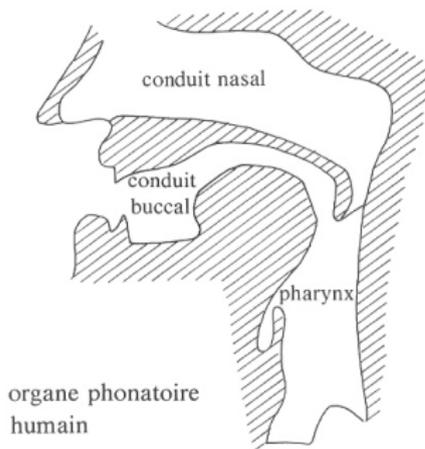
Sons non voisés : aperiodiques

- Exemple 1 : Bruits fricatifs ("s"). Obstruction du conduit vocal, créant des turbulences.
- Exemple 2 : Bruits d'explosion ("p"). Ouverture brusque du conduit vocal.
- Exemple 3 : Claquements de langue, bruits de bouche.

Anatomie et rôle

Les cavités

- 1 Conduit vocal.
- 2 Conduit nasal.



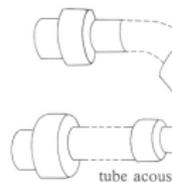
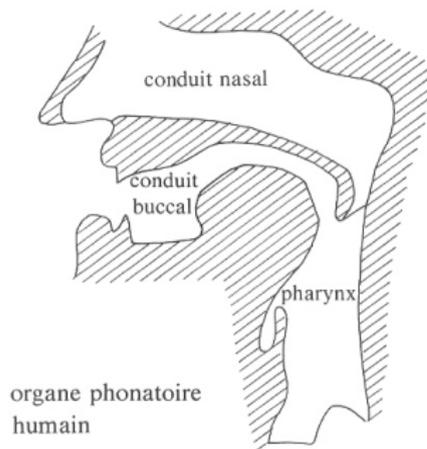
Anatomie et rôle

Les cavités

- 1 Conduit vocal.
- 2 Conduit nasal.

Les articulateurs

- 1 La langue.
- 2 La mâchoire
- 3 Les lèvres.



tube acous

Anatomie et rôle

Les cavités

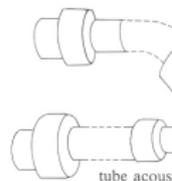
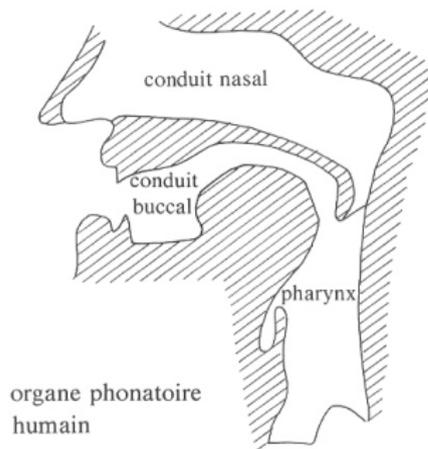
- 1 Conduit vocal.
- 2 Conduit nasal.

Les articulateurs

- 1 La langue.
- 2 La mâchoire
- 3 Les lèvres.

Leur rôle ?

Rôle de filtre.



Classification des sons

4 traits distinctifs des voyelles

- Nasal / Non-nasal. Exemple : /ã/ dans rente, /a/ dans rat.
- Ouvert / Fermé. Exemple : /a/ dans pas, /i/ dans pie.
- Antérieur / Postérieur. Exemple : /i/ dans vie, /u/ dans vous.
- Arrondi / Non-arrondi. Exemple : /o/ dans porte, /e/ dans mais.

Classification des sons

4 traits distinctifs des voyelles

- Nasal / Non-nasal. Exemple : /ã/ dans rente, /a/ dans rat.
- Ouvert / Fermé. Exemple : /a/ dans pas, /i/ dans pie.
- Antérieur / Postérieur. Exemple : /i/ dans vie, /u/ dans vous.
- Arrondi / Non-arrondi. Exemple : /o/ dans porte, /e/ dans mais.

3 traits distinctifs des consonnes

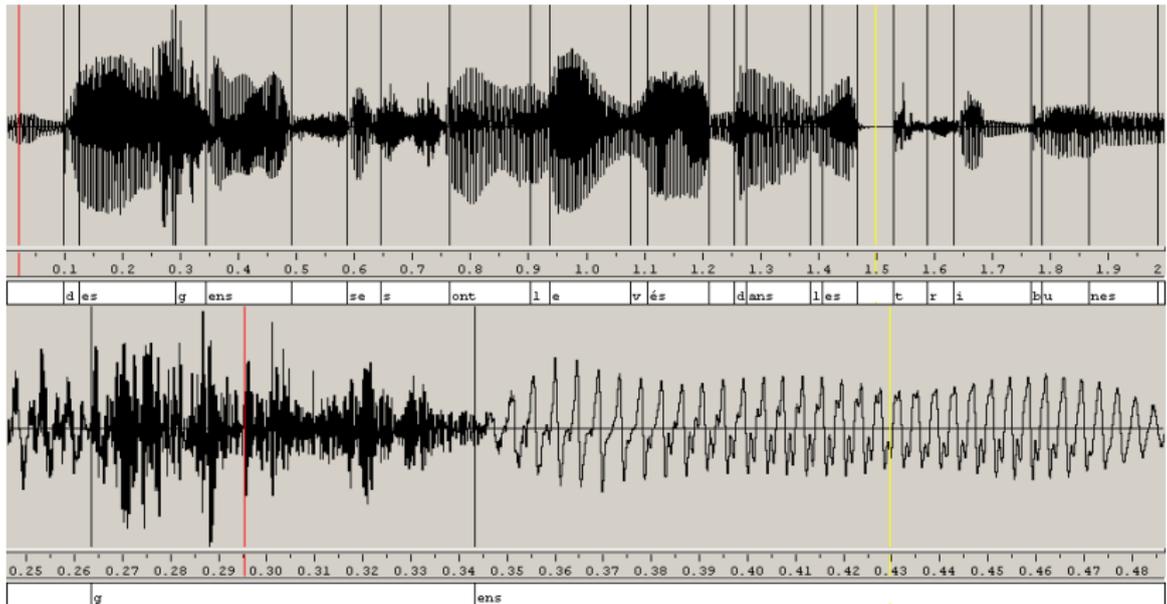
- Voisement / Non-voisement. Exemple : /v/ dans voie, /f/ dans foie.
- Mode d'articulation (occlusif, fricatif, nasal, liquide).
- Lieu d'articulation.

Plan

- 1 Production de la parole**
 - L'appareil respiratoire
 - Les sources vocales
 - Conduit vocal et nasal
 - Phonétique
- 2 Description des signaux de parole**
 - Propriétés des signaux
 - Spectrogramme
 - Formants
 - Analyse prédictive linéaire (LPC)
 - Coefficients cepstraux
- 3 Application au codage de la parole**

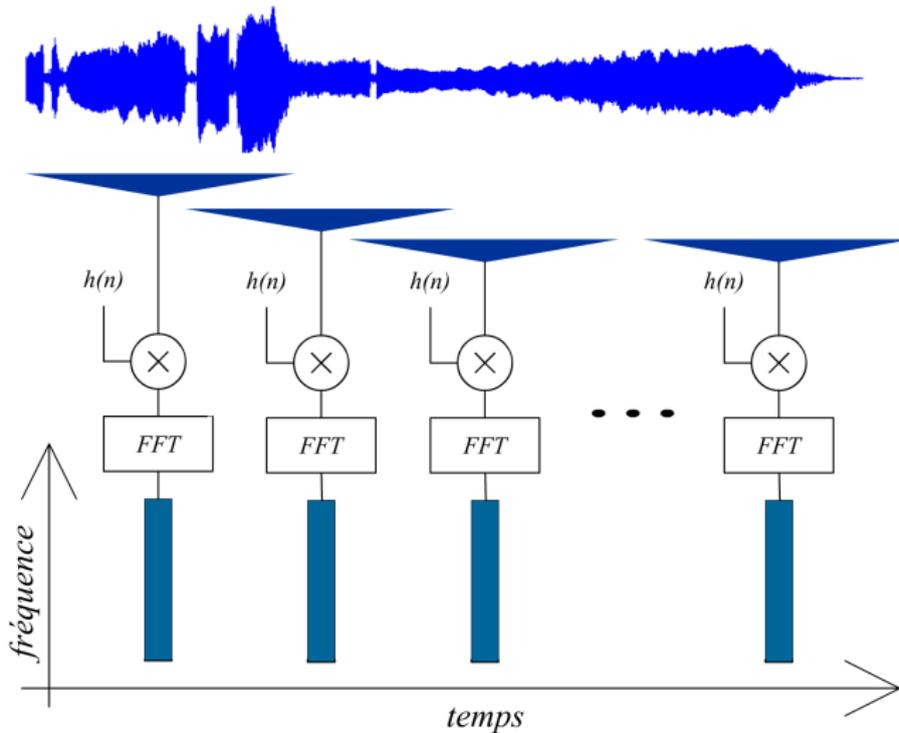
Stationnarité locale

Un signal de parole peut être considéré localement stationnaire, sur de courtes fenêtres de quelques dizaines de millisecondes.

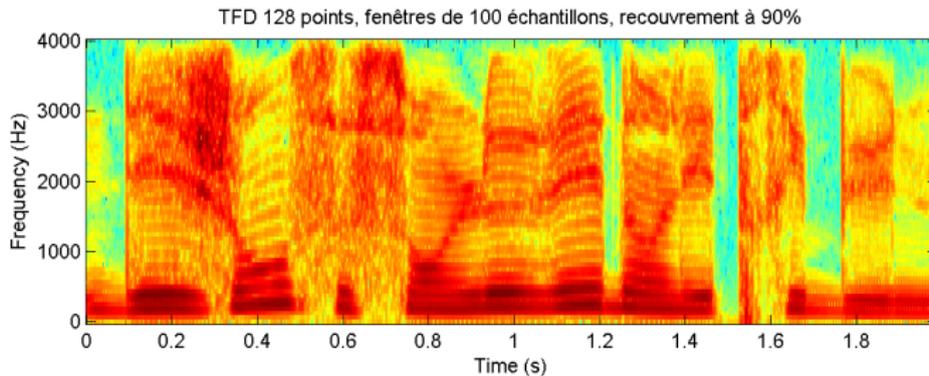
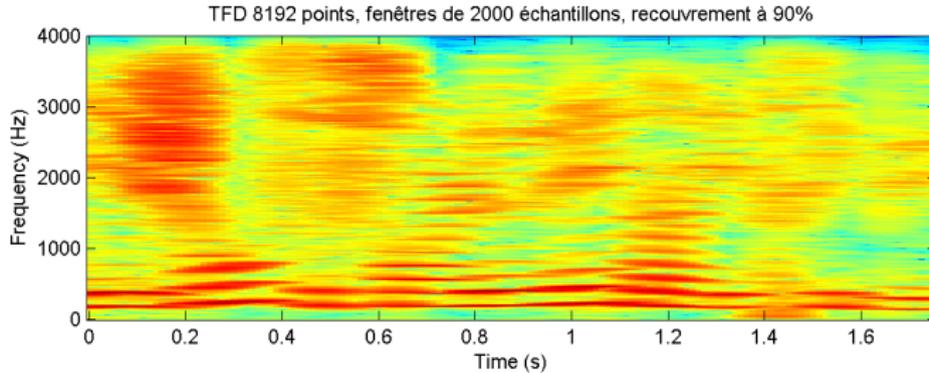


Spectrogramme

Transformée de Fourier à Court Terme

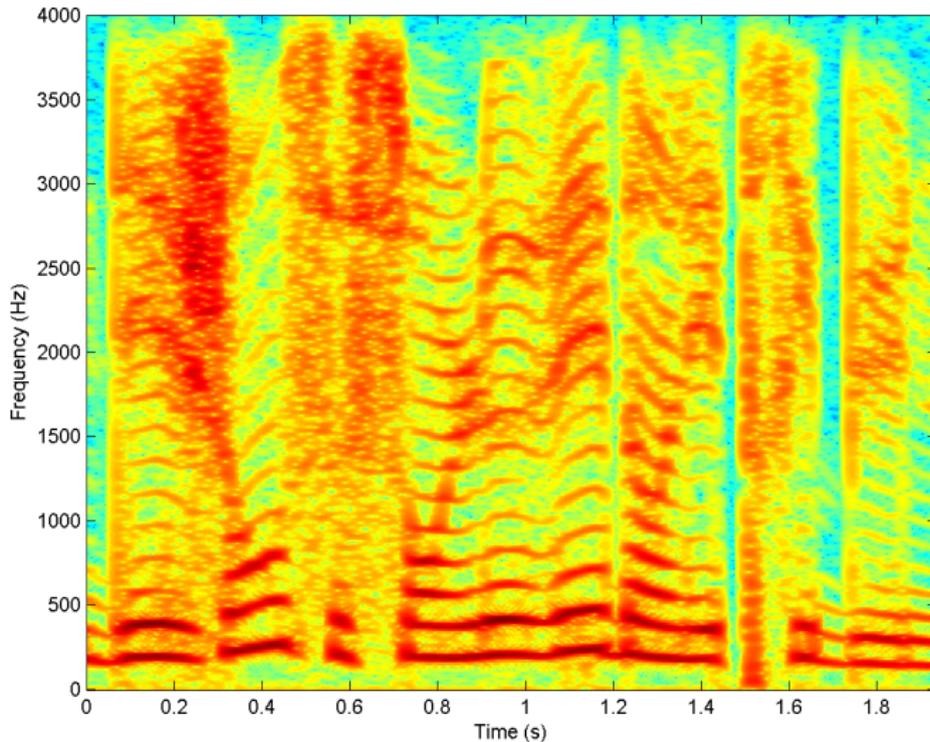


Attention aux paramètres !



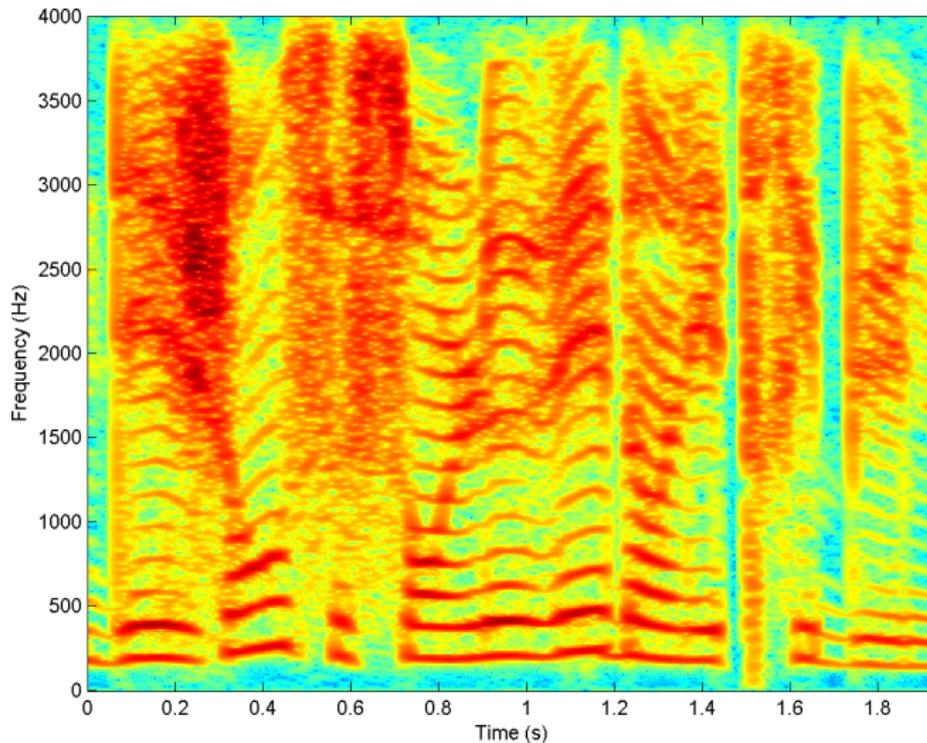
Spectrogramme

Compromis résolution temporelle / fréquentielle



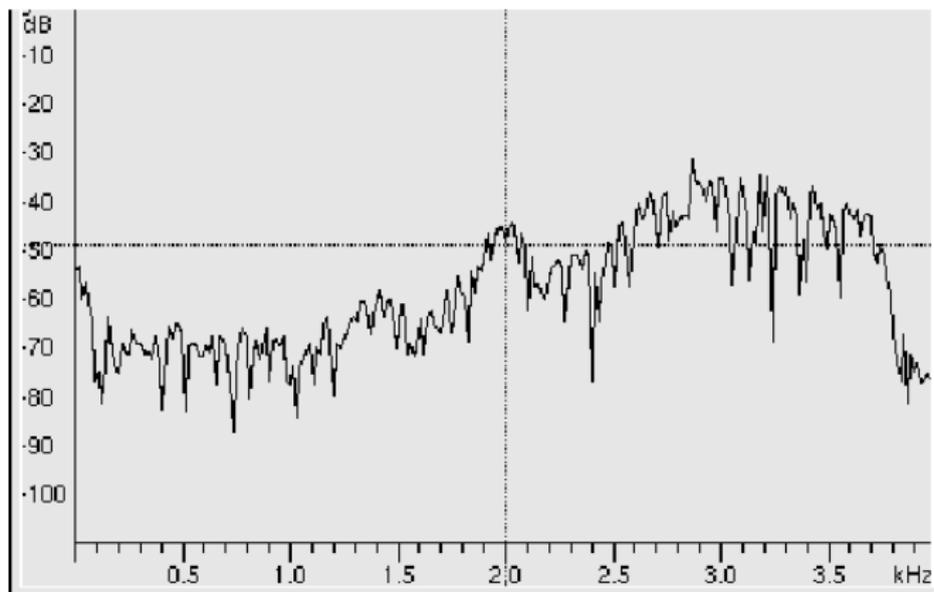
Spectrogramme

Effet du filtre de pré-accentuation

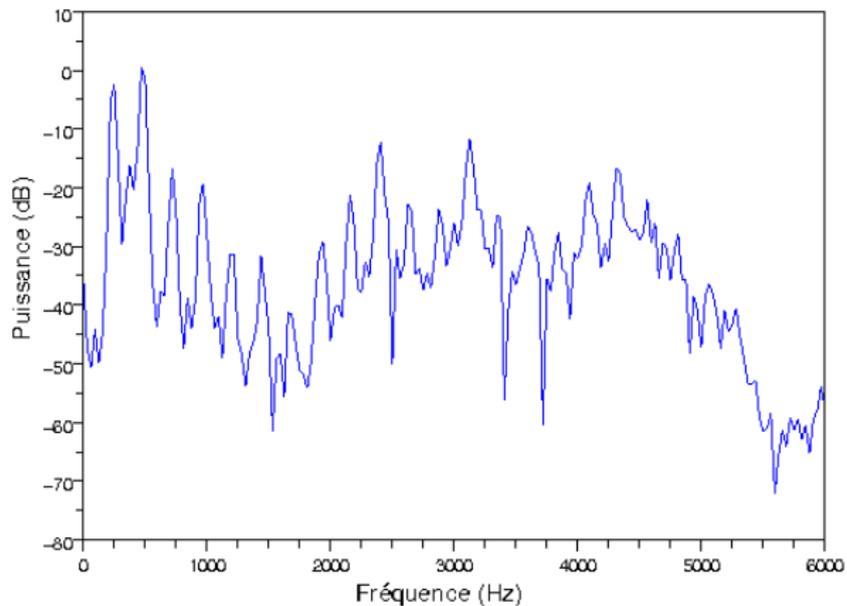


Formants

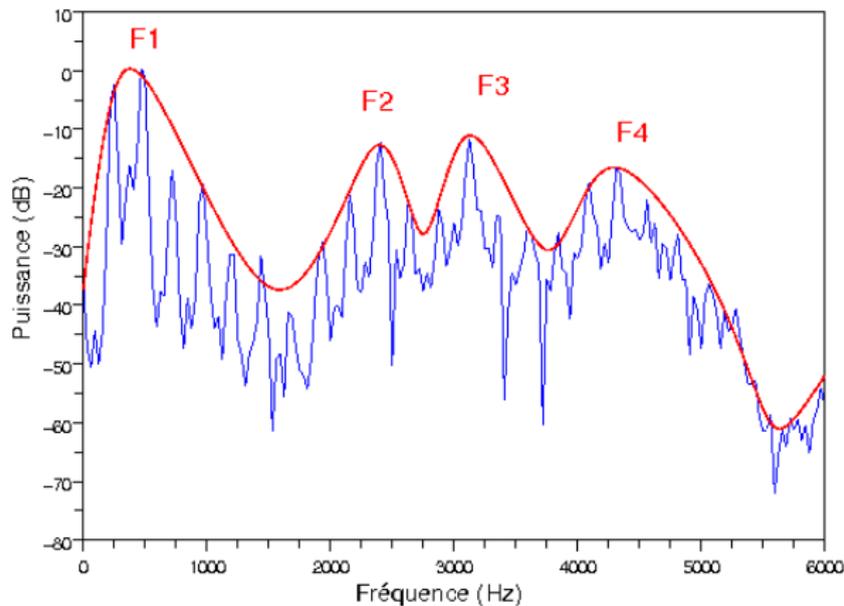
Spectre d'une consonne non-voisée "ch"



Spectre d'une voyelle "i"



Spectre d'une voyelle "i"



Formants

Voyelles : "a e i o u"

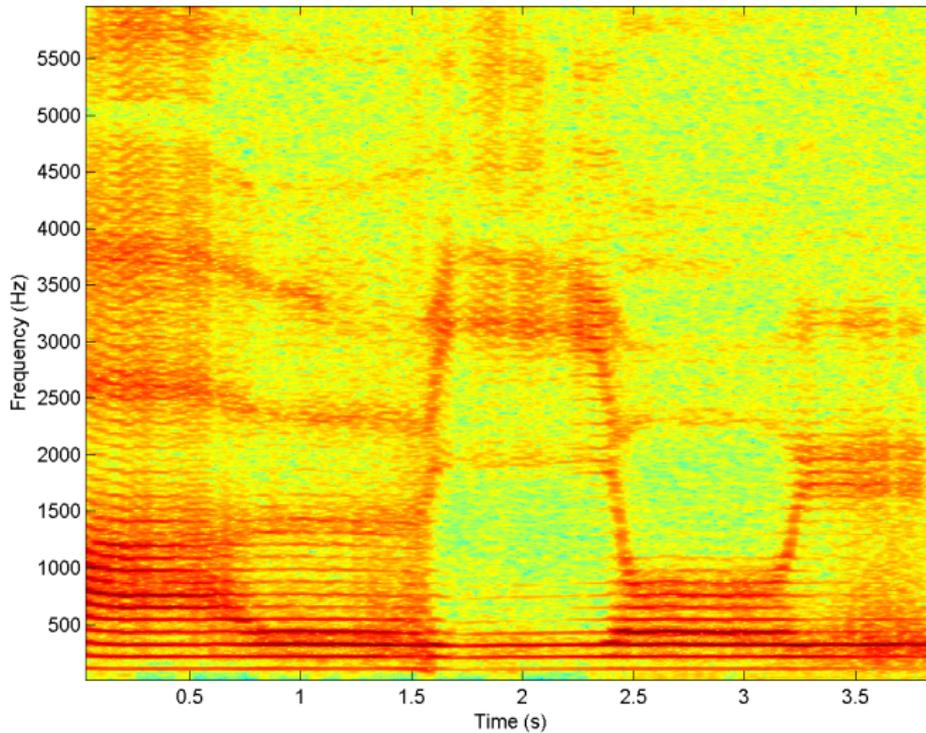
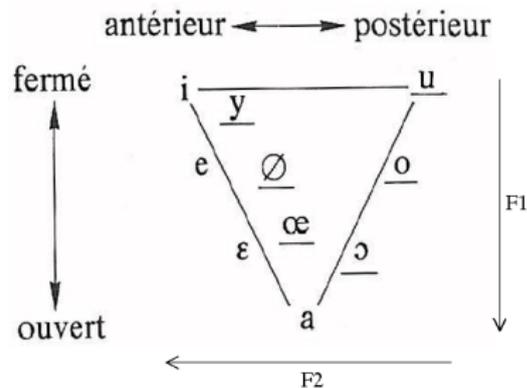
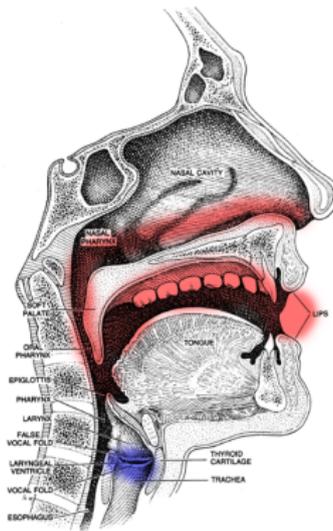
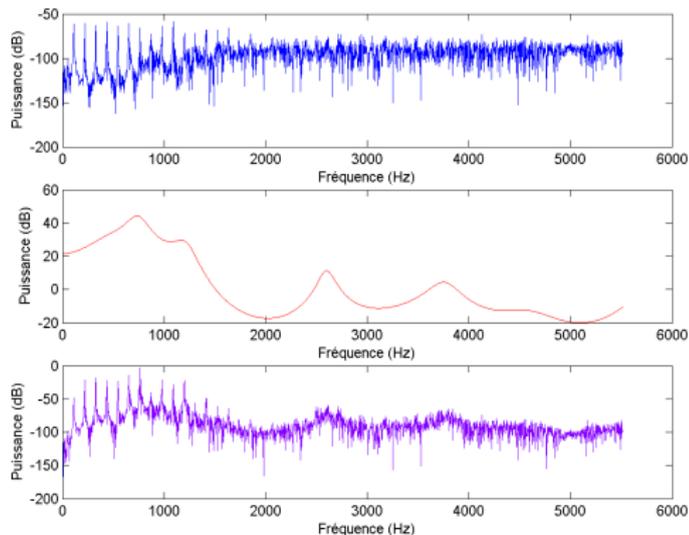


Tableau des formants (voix d'homme)

Voyelle	F_1	F_2
i	250 Hz	2250 Hz
e	420 Hz	2050 Hz
ɛ	590 Hz	1770 Hz
a	760 Hz	1450 Hz
u	290 Hz	750 Hz
o	360 Hz	770 Hz
y	250 Hz	1750 Hz



Interprétation (voyelle a)



Explication

Le train d'impulsions produit par les cordes vocales est filtré par le conduit vocal, dont les résonances correspondent aux formants.

Formants et harmoniques

Formants

- Maxima de l'enveloppe spectrale, "Bosses".
- Correspondent aux fréquences de résonance du conduit vocal.
- Typiques de chaque phonème (consonne ou voyelle)

Formants et harmoniques

Formants

- Maxima de l'enveloppe spectrale, "Bosses".
- Correspondent aux fréquences de résonance du conduit vocal.
- Typiques de chaque phonème (consonne ou voyelle)

Harmoniques

- N'existent que pour les phonèmes voisés.
- Raies régulièrement espacées dans le spectre.
- Ecart entre deux raies = hauteur de la voix

Objectif de l'analyse prédictive linéaire

Rappel...

On a vu que les signaux de parole pouvaient être décrits comme la modification d'une **source** par un **filtre**.

Objectif de l'analyse prédictive linéaire

Rappel...

On a vu que les signaux de parole pouvaient être décrits comme la modification d'une **source** par un **filtre**.

Il serait intéressant de pouvoir extraire...

- La partie correspondant à la source (train d'impulsion dans le cas des sons voisés, ou bruit).
- La forme du filtre et donc les formants.

Objectif de l'analyse prédictive linéaire

Rappel...

On a vu que les signaux de parole pouvaient être décrits comme la modification d'une **source** par un **filtre**.

Il serait intéressant de pouvoir extraire...

- La partie correspondant à la source (train d'impulsion dans le cas des sons voisés, ou bruit).
- La forme du filtre et donc les formants.

Intérêt

- Reconnaissance : la position des formants et le caractère voisé/non-voisé permettent d'identifier le phonème.
- Codage : la position des formants change lentement ; et les signaux de la source se représentent très facilement.

Analyse prédictive linéaire

Principe

Modèle source-filtre discret : $x(n) = e(n) - \sum_{k=1}^P a_k x(n-k)$

Conduit vocal modélisé par le filtre tout-pôles :

$$H(z) = \frac{1}{1 + \sum_{k=1}^P a_k z^{-k}}$$

tel que $X(Z) = H(z)E(z)$.

On cherche les coefficients de H qui minimisent l'énergie de e .

Analyse prédictive linéaire

Principe

Modèle source-filtre discret : $x(n) = e(n) - \sum_{k=1}^P a_k x(n-k)$

Conduit vocal modélisé par le filtre tout-pôles :

$$H(z) = \frac{1}{1 + \sum_{k=1}^P a_k z^{-k}}$$

tel que $X(Z) = H(z)E(z)$.

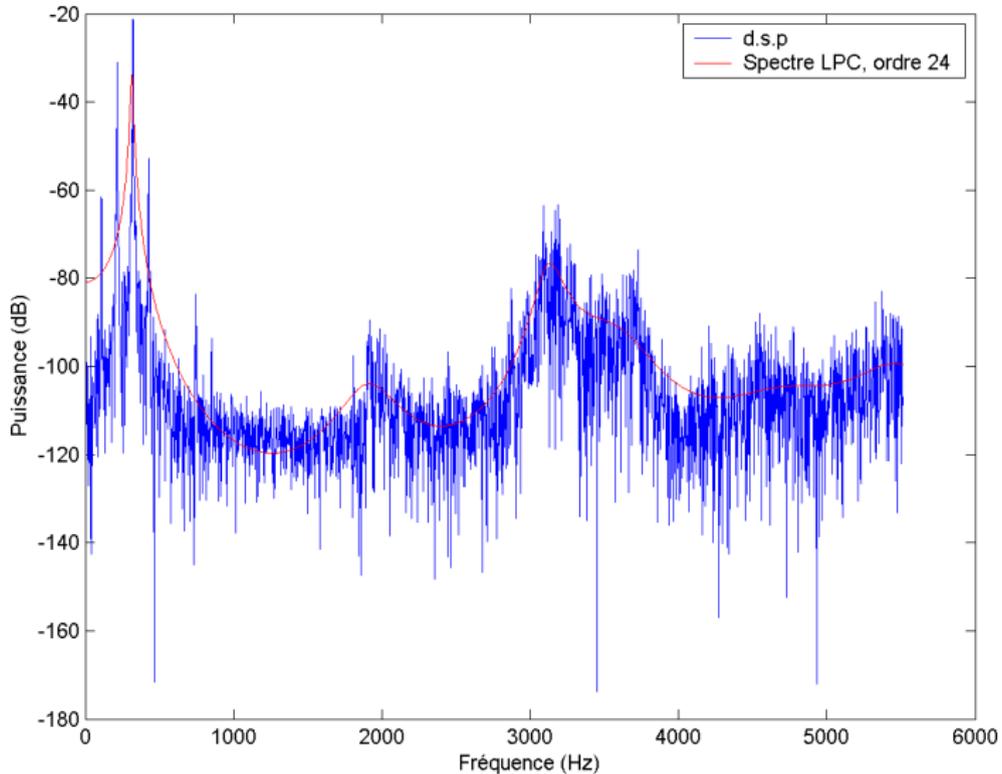
On cherche les coefficients de H qui minimisent l'énergie de e .

Spectre LPC

La réponse en fréquence du filtre $H(z)$ est une version "lissée" de la densité spectrale de puissance de e .

Analyse prédictive linéaire (LPC)

Exemple (voyelle "u")



Cepstre

Motivation

Comment extraire simplement la source et le filtre ?
 x signal observé, e source, h filtre.

Cepstre

Motivation

Comment extraire simplement la source et le filtre ?
 x signal observé, e source, h filtre.

Etape 1 : Simplifier la relation entre e , x et h

- Domaine temporel : $x(n) = e(n) * h(n)$.
- Domaine fréquentiel : $X(f) = E(f)H(f)$.
- Cepstre : $\mathcal{C}_e = \text{TF}^{-1}[\log |E(f)|]$
 → Domaine quéfrentiel : $\mathcal{C}_x = \mathcal{C}_e + \mathcal{C}_h$

Cepstre (2)

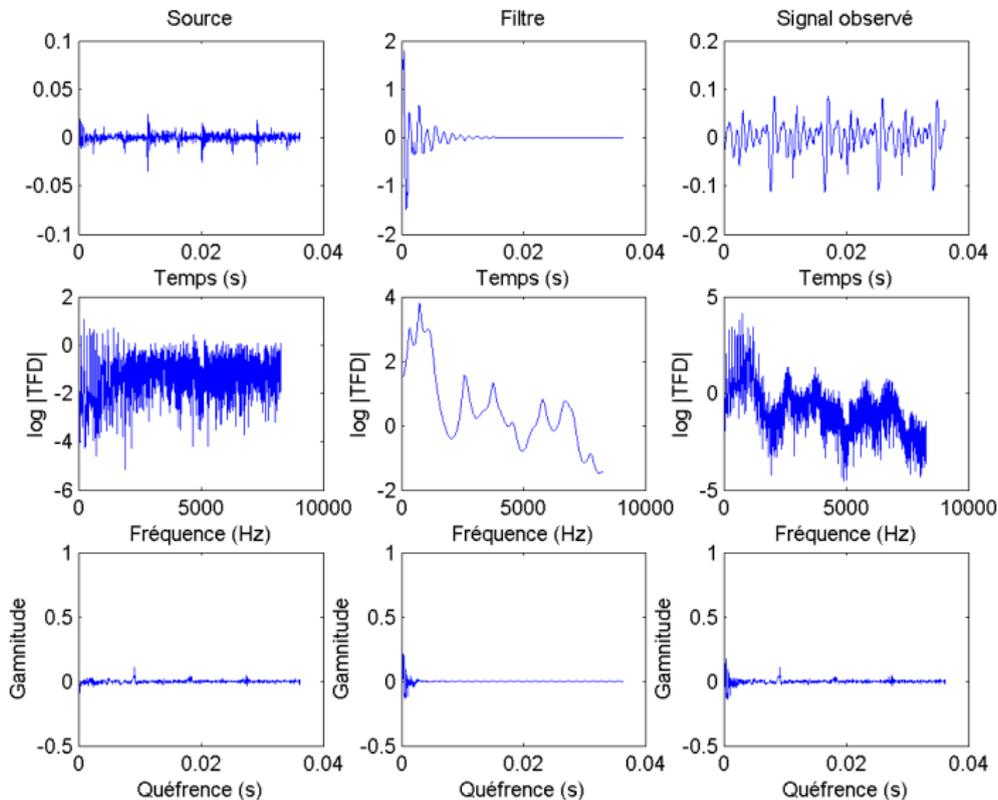
Etape 2 : disjointre les deux composantes

Propriétés du cepstre

- $\log |H(f)|$ a des variations lentes \rightarrow Seuls les premiers coefficients de sa TF^{-1} sont importants.
- $\log |E(f)|$ est périodique dans le cas des sons voisés \rightarrow sa TF^{-1} exhibera un pic important à la période correspondante.

Ainsi, les premiers coefficients du cepstre décrivent la forme du filtre ; tandis que ses pics suivants caractérisent la source.

Un exemple



MFCC

Définition

MFCC : Mel Frequencx Cepstrum Coefficients.

Motivation : Le cepstre traite de façon égale chaque fréquence. Or, l'oreille perçoit l'échelle des fréquence de façon déformée ; et en groupant des fréquences adjacentes (bandes critiques).

MFCC

Définition

MFCC : Mel Frequencx Cepstrum Coefficients.

Motivation : Le cepstre traite de façon égale chaque fréquence. Or, l'oreille perçoit l'échelle des fréquence de façon déformée ; et en groupant des fréquences adjacentes (bandes critiques).

Calcul des MFCC, en 3 étapes

- 1 Module de la TFD de e .
- 2 Calcul de l'énergie dans chaque bande critique → 24 coefficients fréquentiels $X_0...X_{23}$.
- 3 TFD inverse de X .

On se limite souvent aux 12 premiers coefficients.

Plan

- 1 Production de la parole**
 - L'appareil respiratoire
 - Les sources vocales
 - Conduit vocal et nasal
 - Phonétique
- 2 Description des signaux de parole**
 - Propriétés des signaux
 - Spectrogramme
 - Formants
 - Analyse prédictive linéaire (LPC)
 - Coefficients cepstraux
- 3 Application au codage de la parole**

Principe des codeurs par analyse/synthèse

Idée générale

- Analyse source-filtre du signal par trames de qqms ms
- Transmission des paramètres de la source et du filtre.
- Le décodeur resynthétise le signal à partir de ce modèle.

Principe des codeurs par analyse/synthèse

Idée générale

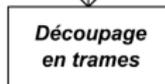
- Analyse source-filtre du signal par trames de qqs ms
- Transmission des paramètres de la source et du filtre.
- Le décodeur resynthétise le signal à partir de ce modèle.

Codage de la source

- Approche 1 : on la quantifie avec très peu de bits (une partie de l'information étant dans le filtre, la source est moins sensible à la quantification).
- Approche 2 : on la paramétrise de façon très simple et on transmet les paramètres.

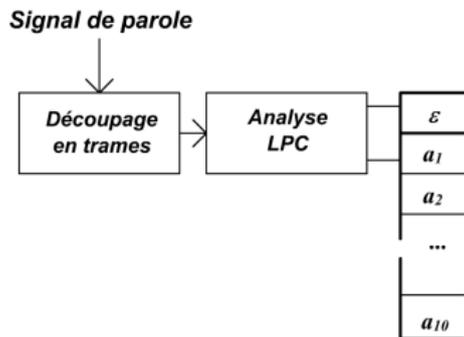
Codeur LPC10

Signal de parole



*Découpage
en trames*

Codeur LPC10



Codeur LPC10

Signal de parole

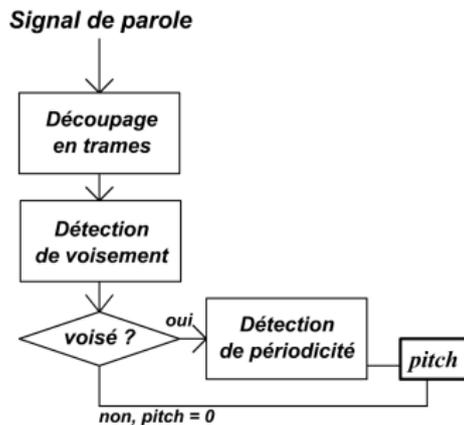


*Découpage
en trames*

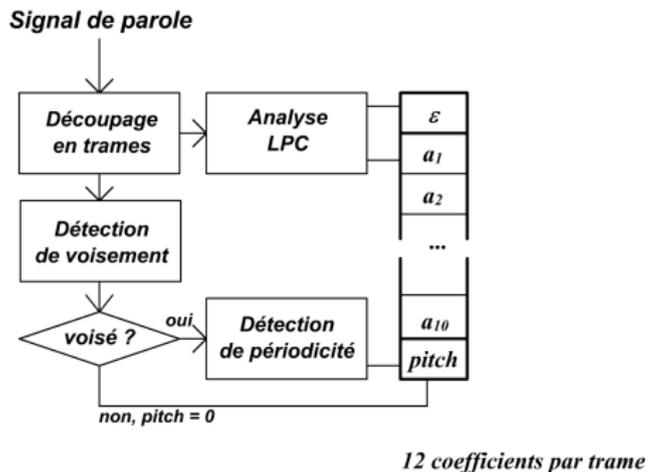


*Détection
de voisement*

Codeur LPC10

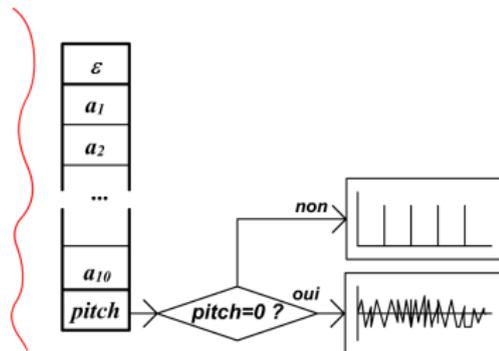


Codeur LPC10

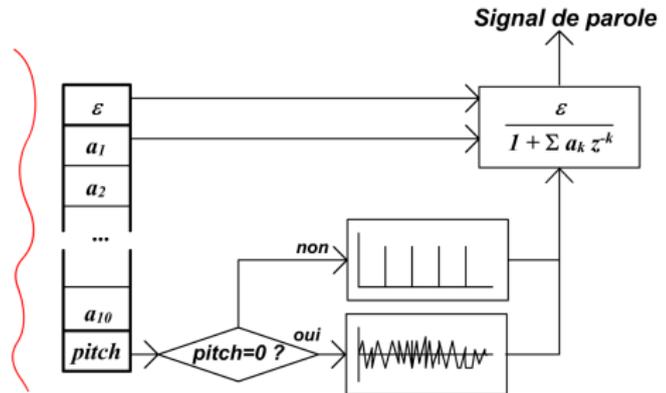


Codage

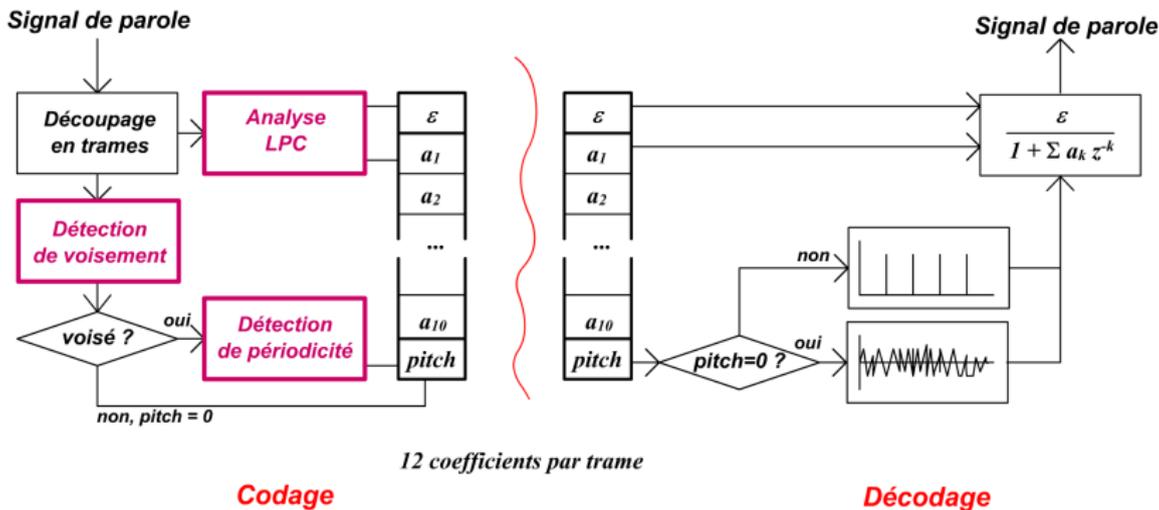
Codeur LPC10



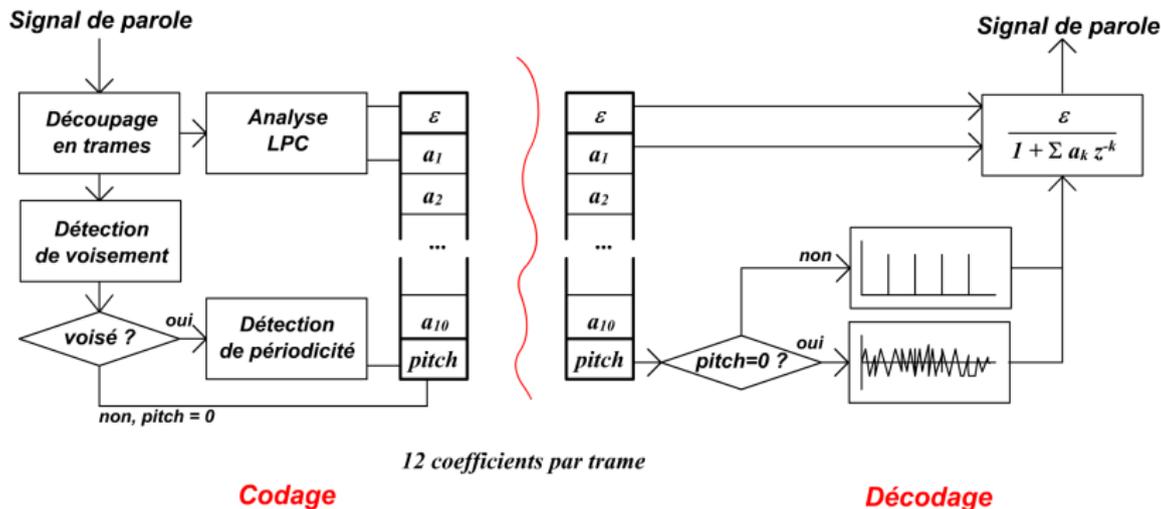
Codeur LPC10



Codeur LPC10



Codeur LPC10



Codeur LPC10 (2)

