

Codage de la parole

Mines d'Alès
Janvier 2024

Plan

- I. Rappels : la conversion analogique/numérique
- II. Codage PCM
- III. Les paramètres fondamentaux d'un CP
- IV. Le codage par analyse-synthèse
- V. Exemples
- VI. Etude comparative

I. Rappels : CAN

$p(t)$

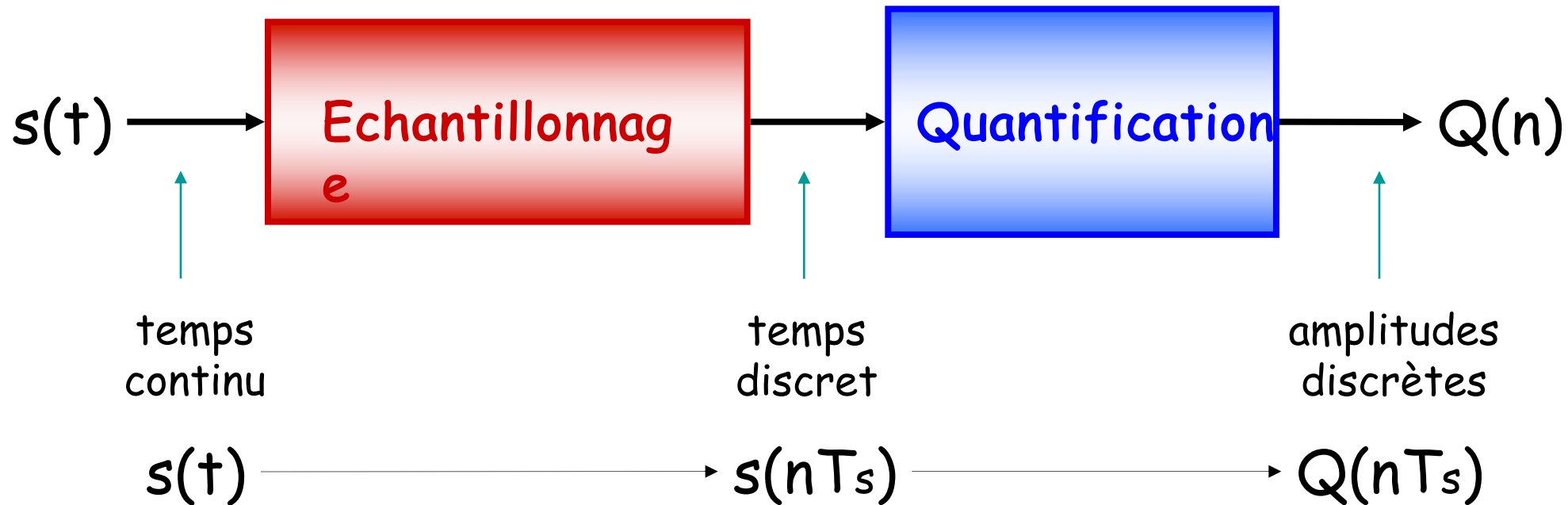


$s(t)$ \longrightarrow $s(t)$ signal analogique

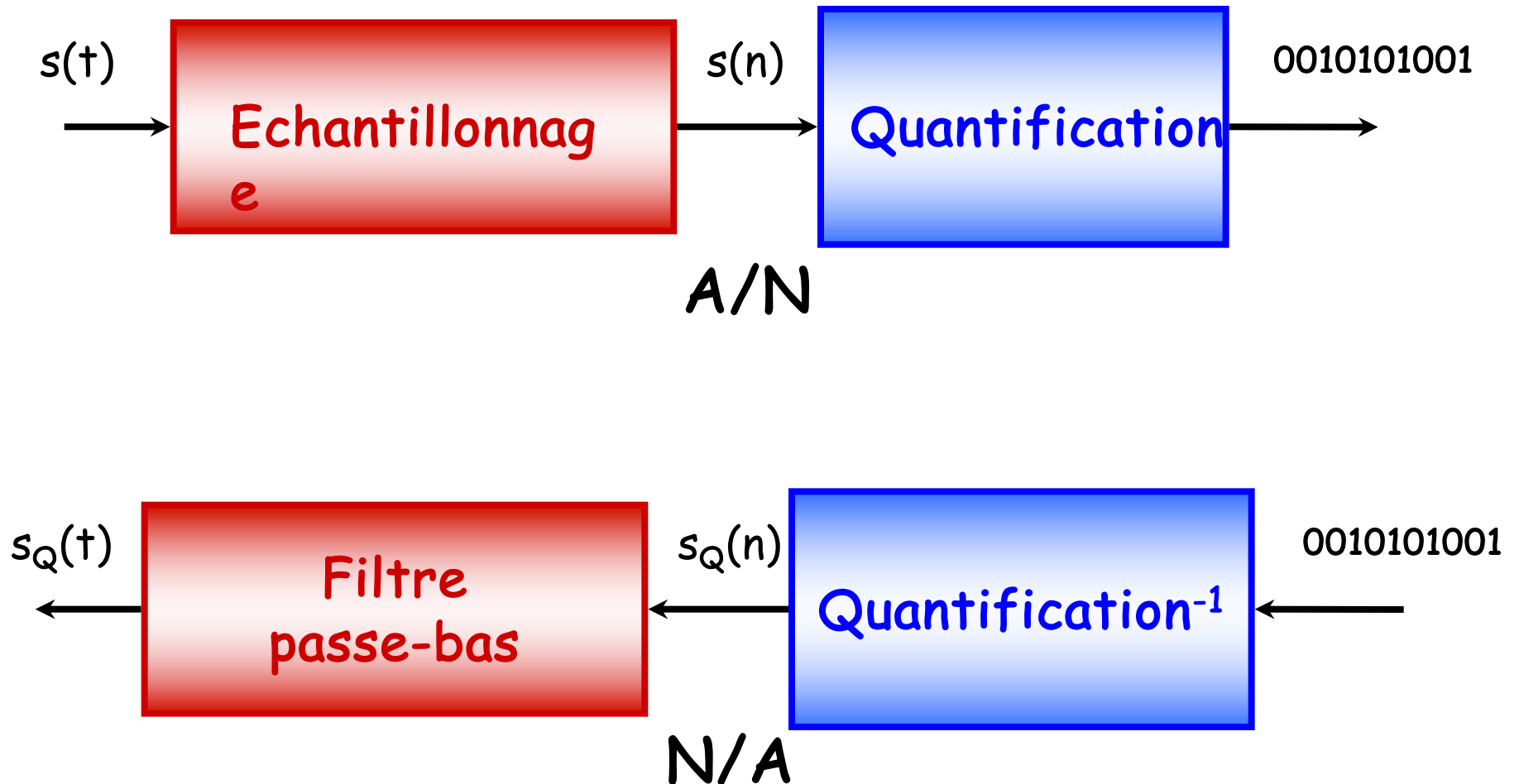
Temps continu
et
amplitudes continues

Numériser $s(t)$ signifie trouver une représentation
temps discret - amplitudes discrètes

Le convertisseur A/N



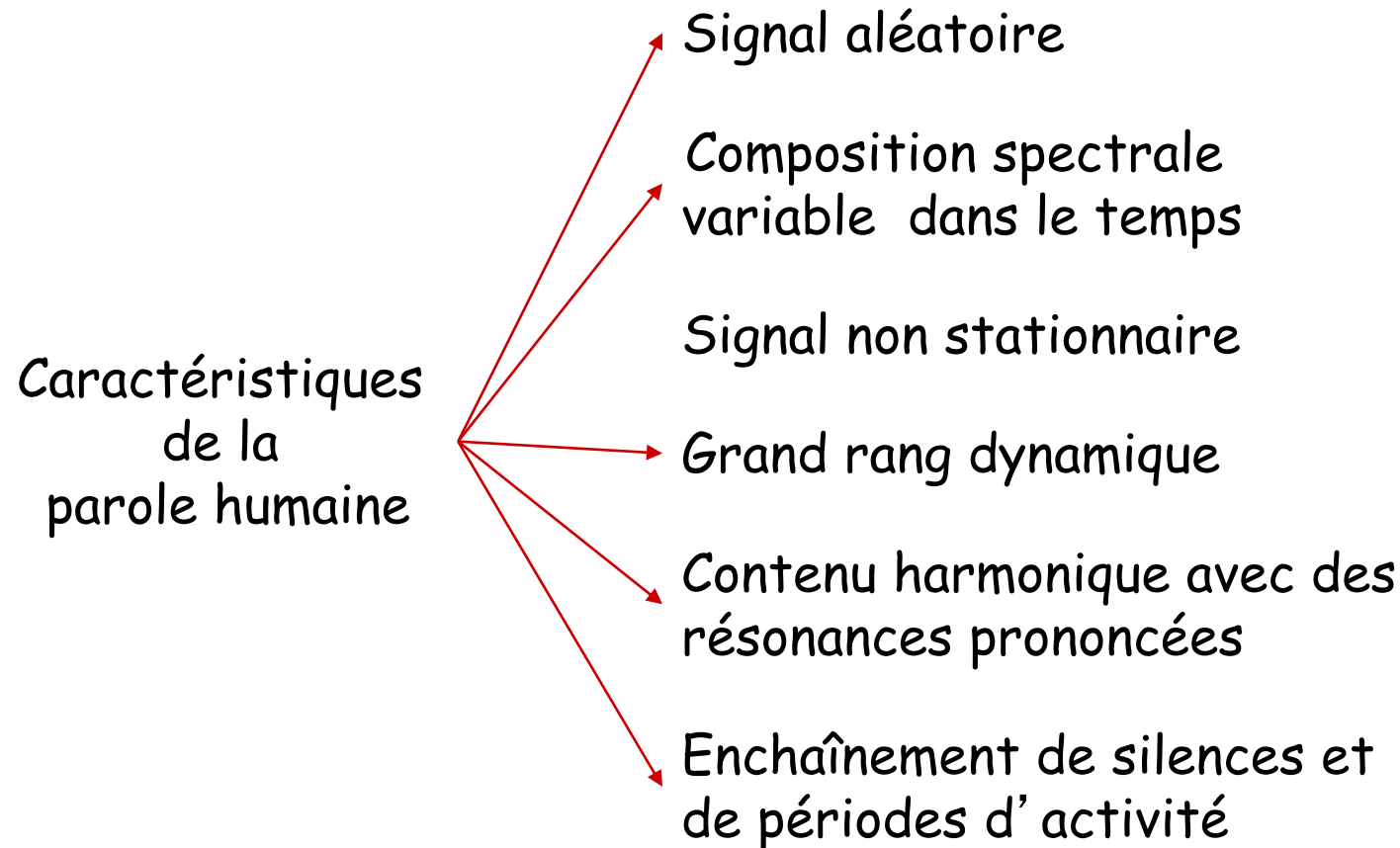
La conversion $A/N \rightarrow N/A$



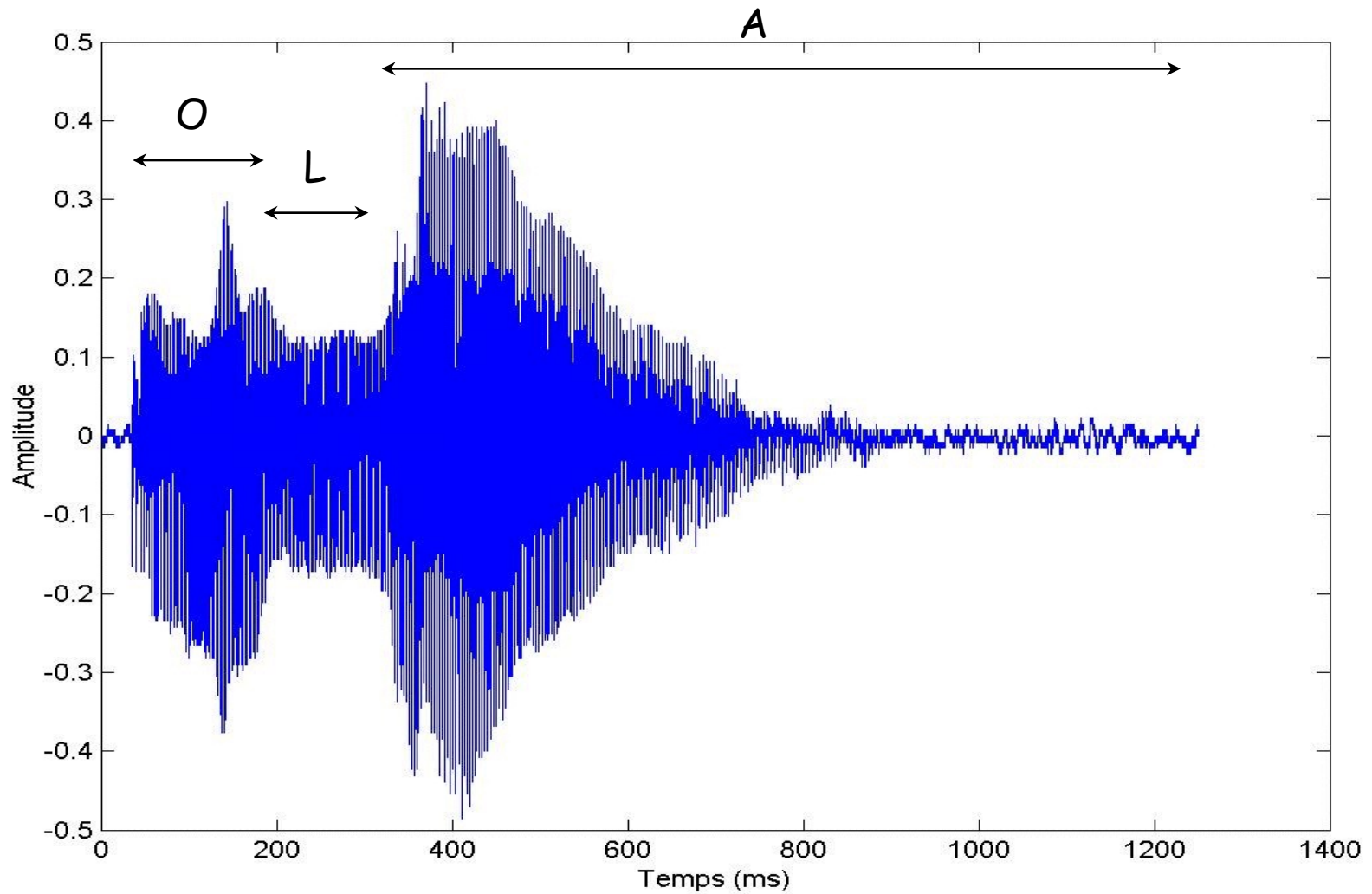
Plan (II)

- I. Conversion analogique/numérique
- II. Codage PCM**
- III. Les paramètres fondamentaux d'un CP
- IV. Le codage par analyse-synthèse
- V. Exemples
- VI. Etude comparative

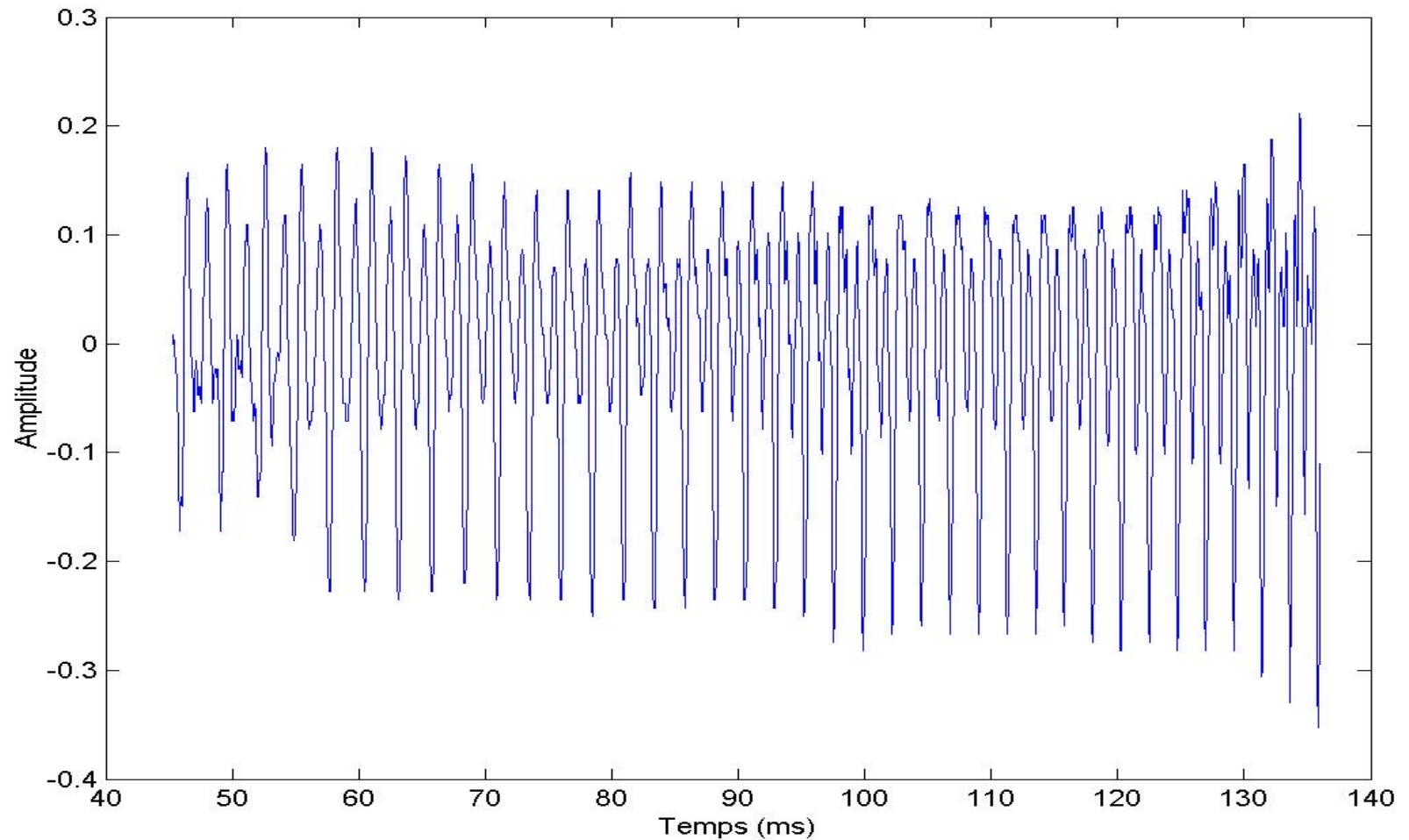
II. Le PCM et la parole humaine



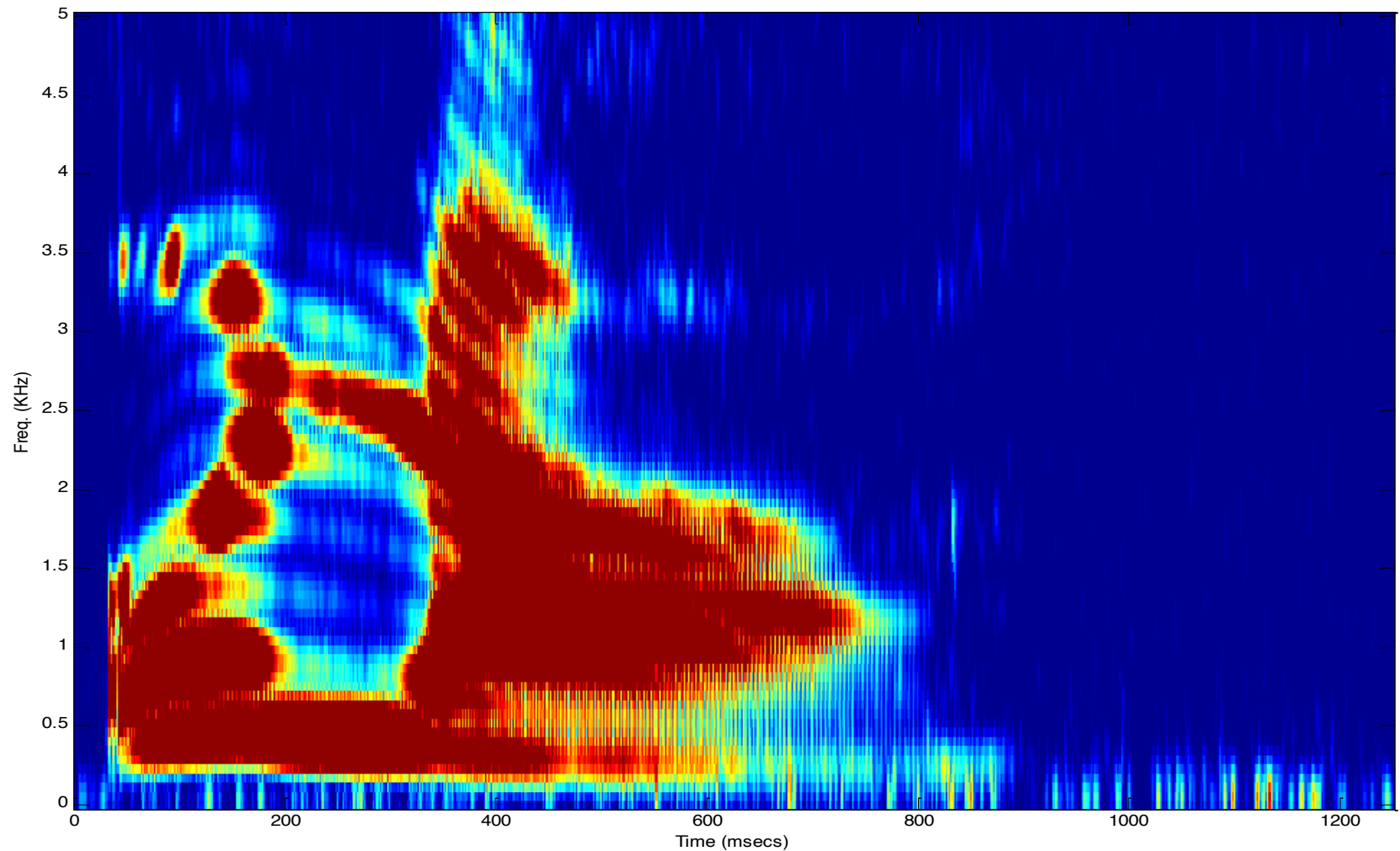
Exemple



Exemple (suite)



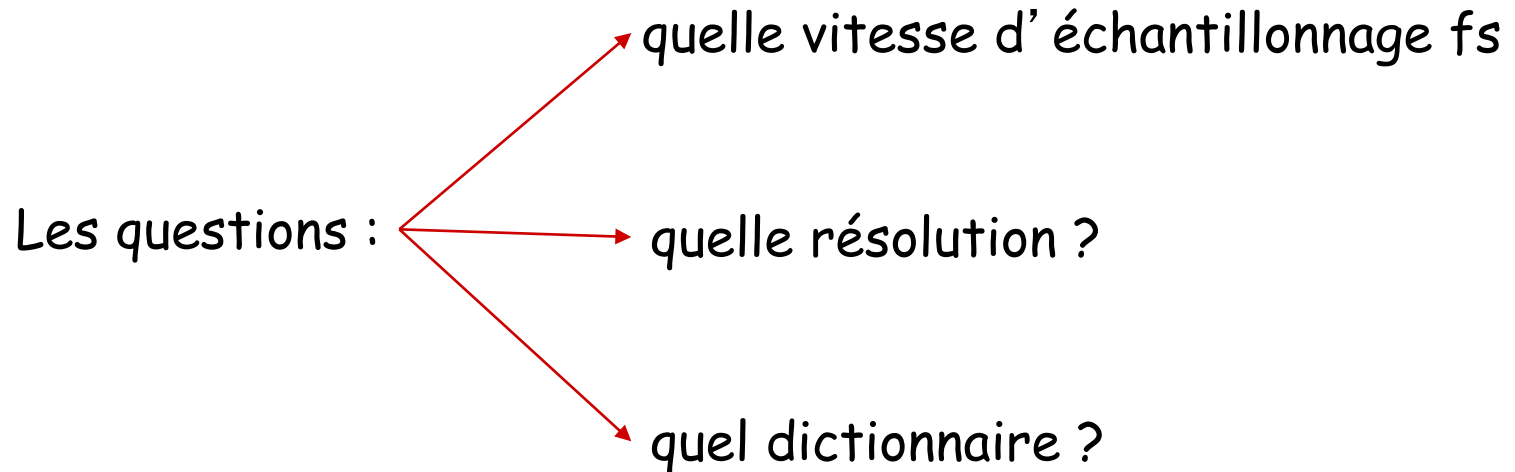
Analyse spectrale



PCM - les questions

Comment numériser
la parole humaine ?

Approche conversion A/N !!



PCM - les réponses

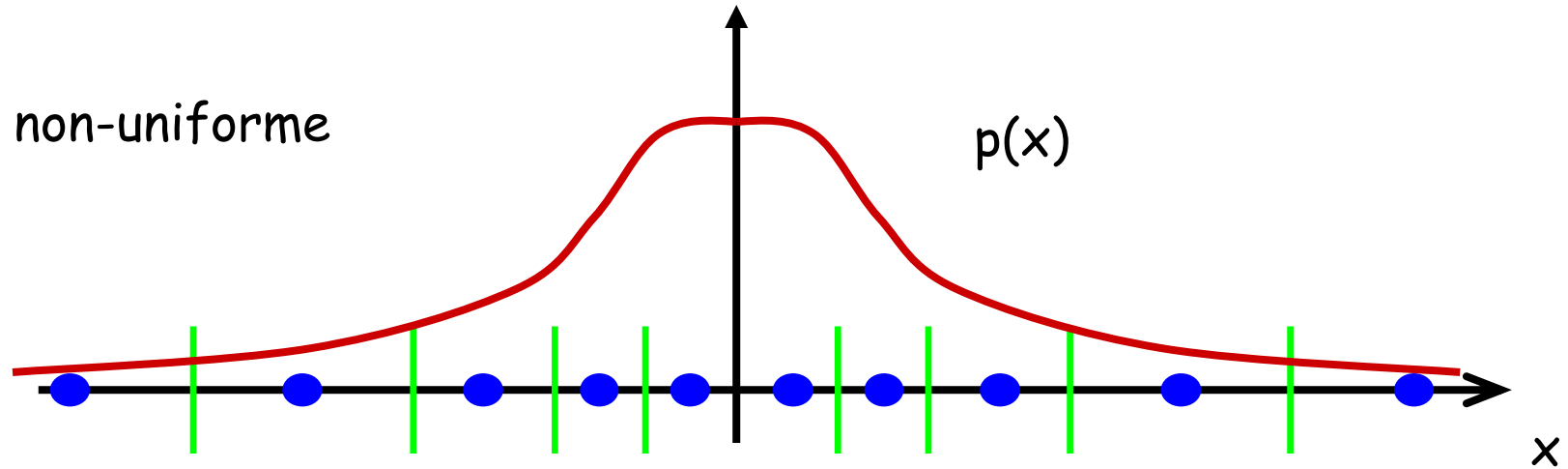
Quelle vitesse
d'échantillonnage
 f_s ? \longrightarrow B est estimé
à 4kHz $\longrightarrow f_s = 8 \text{ kéch/s}$

Quelle résolution ? \longrightarrow Estimé à 10 bits/éch

Quel dictionnaire ? $\begin{cases} \longrightarrow \text{Dictionnaire non-uniforme} \\ \longrightarrow \text{Compression logarithmique} \end{cases}$

PCM - les réponses (suite)

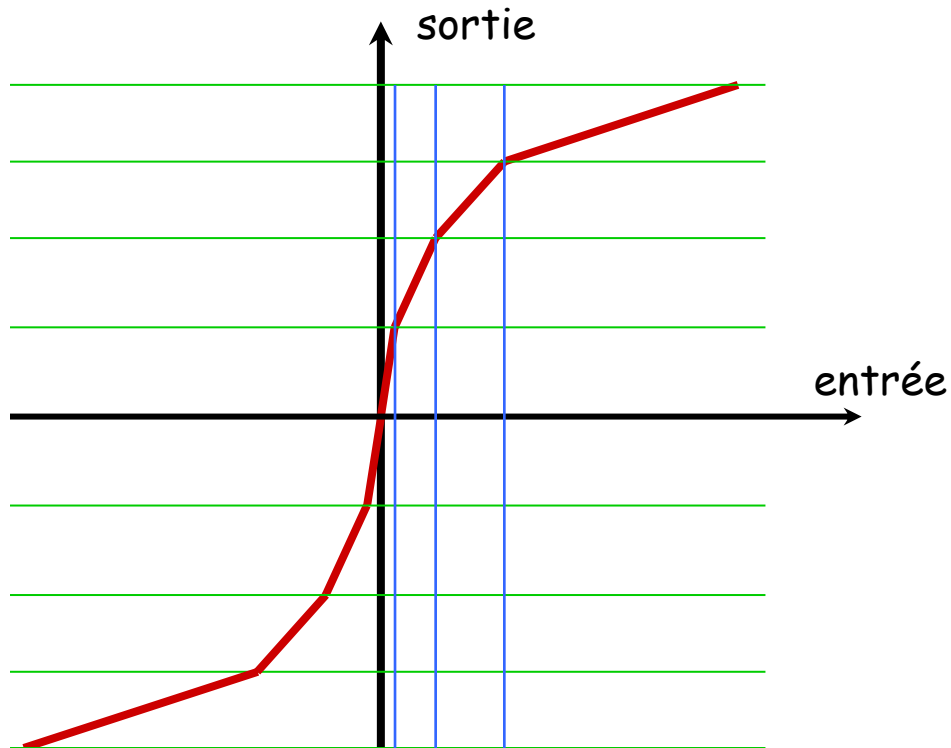
Dictionnaire non-uniforme



Compression logarithmique



PCM - un exemple - la loi μ



Loi- μ

$$x = V/V_{\max}$$

$$y = \frac{\operatorname{sgn}(x)}{\log(1+\mu)} \log(1 + \mu|x|)$$

$\mu=100$ ou 255

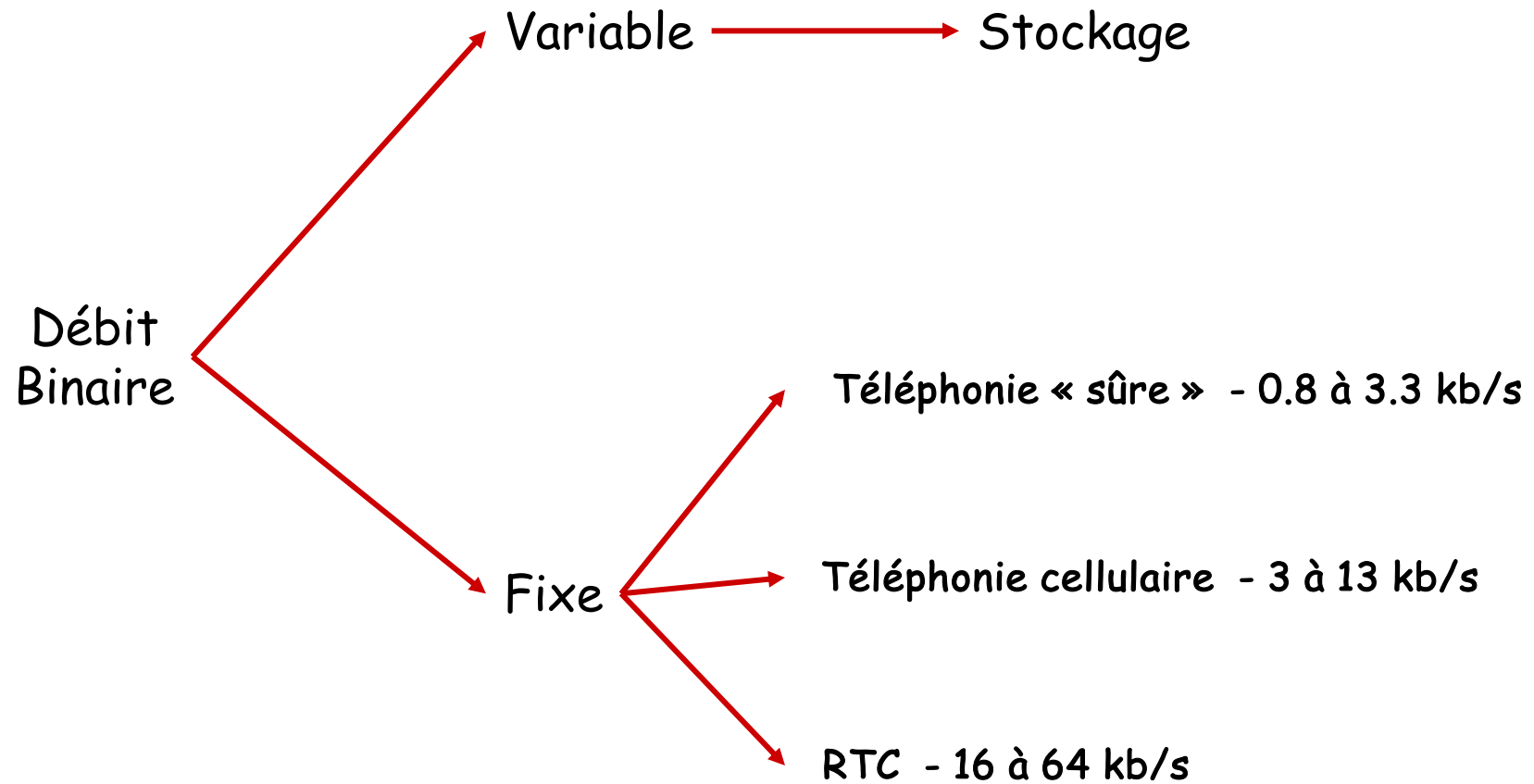
Plan (III)

- I. Conversion analogique/numérique
- II. Codage PCM
- III. Les paramètres fondamentaux d'un CP**
- IV. Le codage par analyse-synthèse
- V. Exemples
- VI. Etude comparative

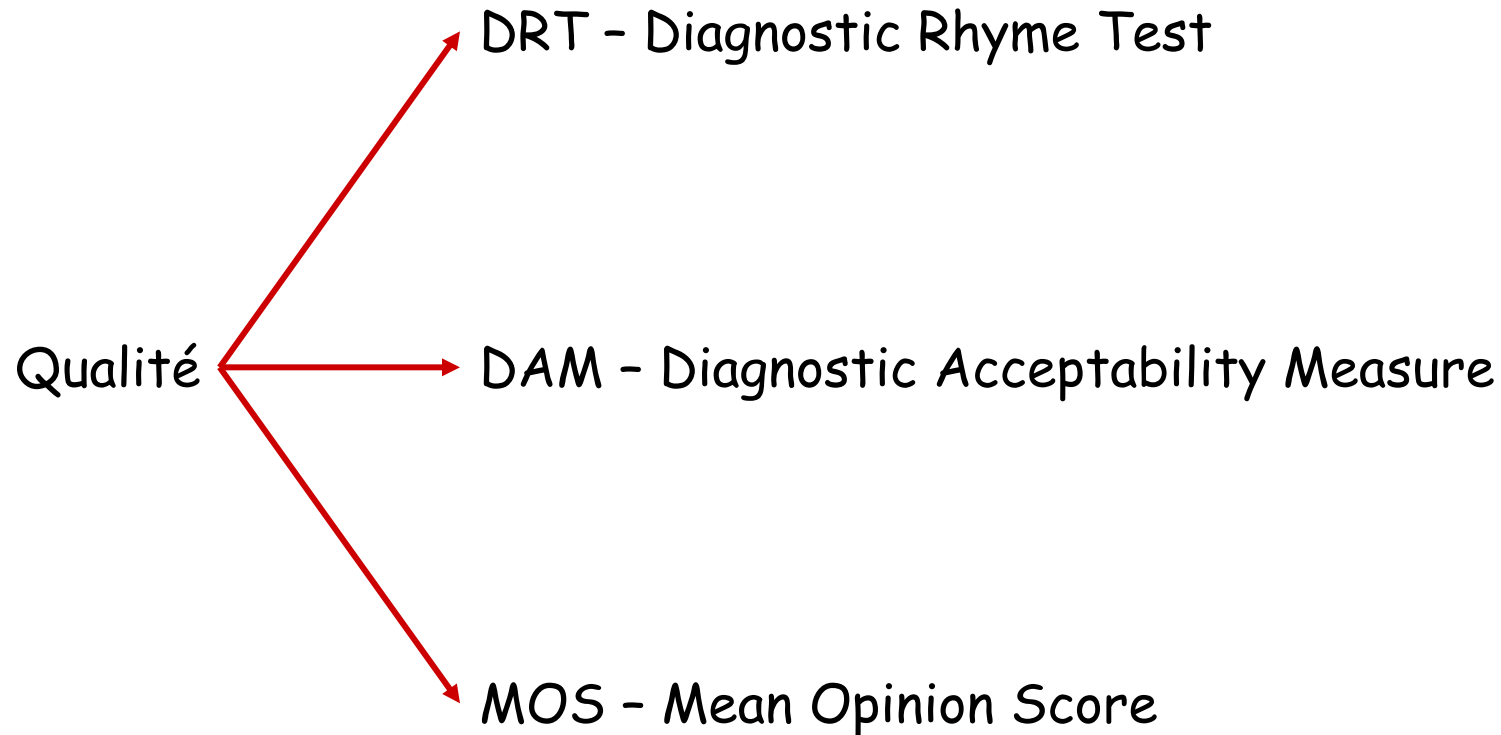
Caractéristiques d'un codeur de parole

- ✓ Débit binaire
- ✓ Qualité perceptuelle
- ✓ Complexité algorithmique
- ✓ Mémoire
- ✓ Retard de traitement
- ✓ Sensibilité aux erreurs
- ✓ Bande passante

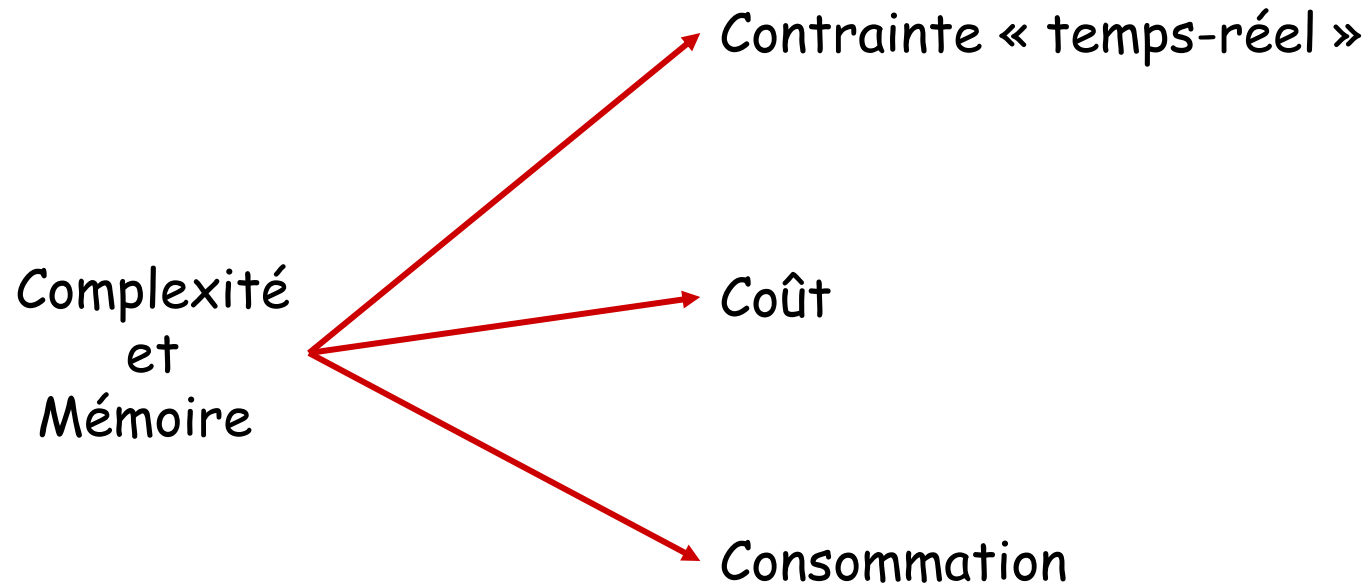
Débit binaire



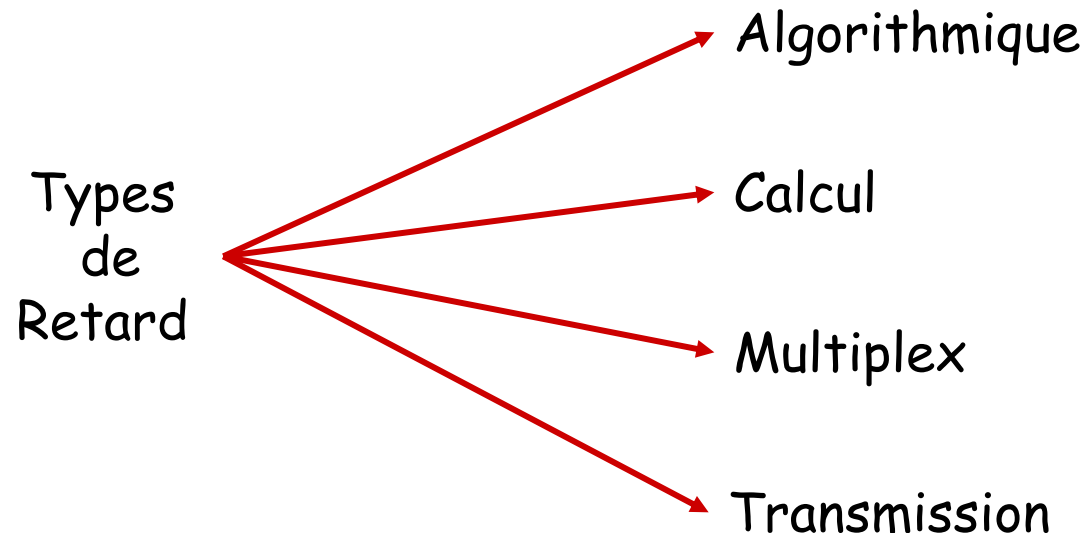
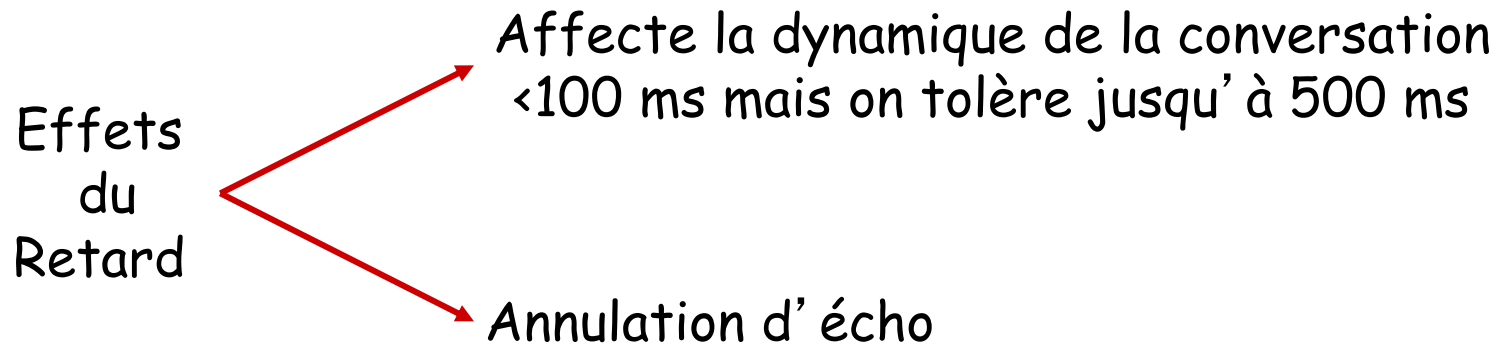
Qualité perceptuelle



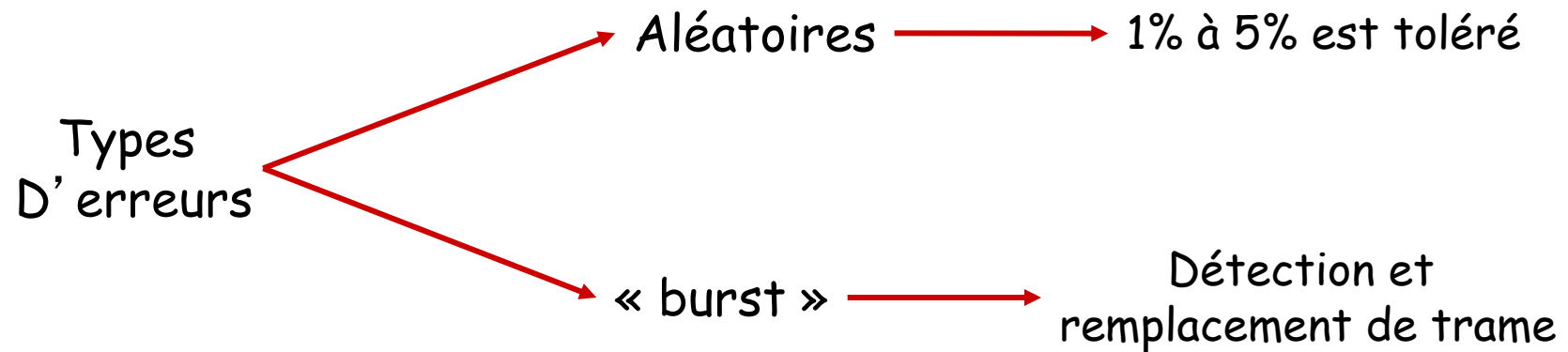
Complexité et mémoire



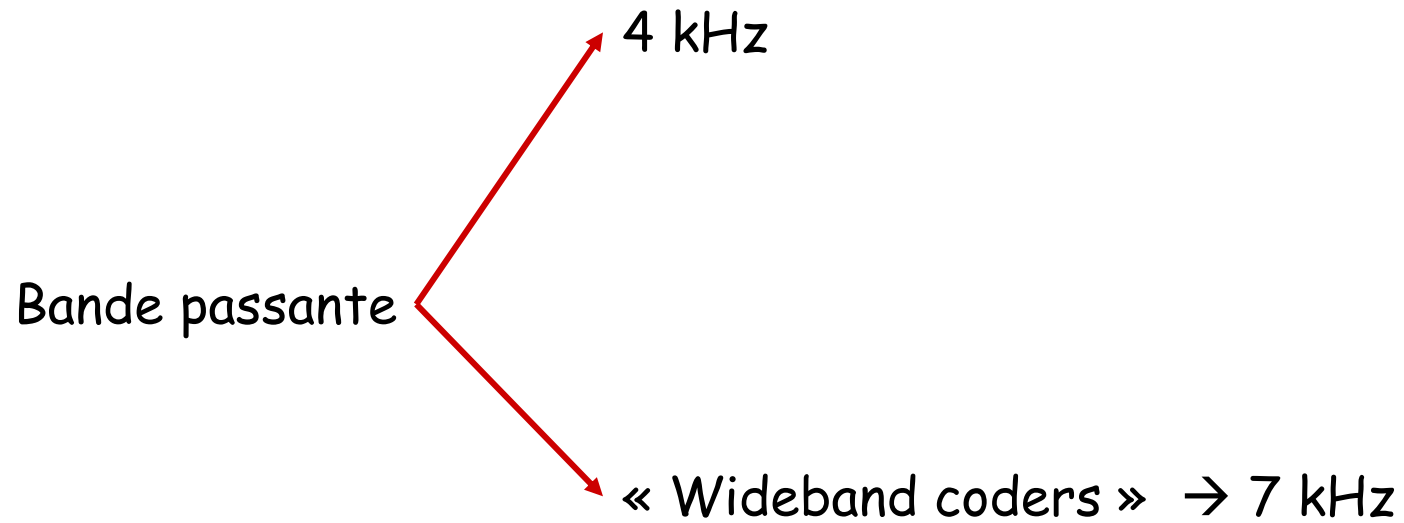
Retard de traitement



Sensibilité aux erreurs de transmission



Bande passante



Plan (IV)

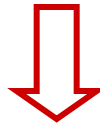
- I. Conversion analogique/numérique
- II. Codage PCM
- III. Les paramètres fondamentaux d'un CP
- IV. Le codage par analyse-synthèse**
- V. Exemples
- VI. Etude comparative

Motivation

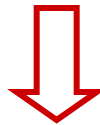
Pour les débits binaires
entre 16 et 64 kb/s



La conversion A/N est une solution
simple de codage de la parole



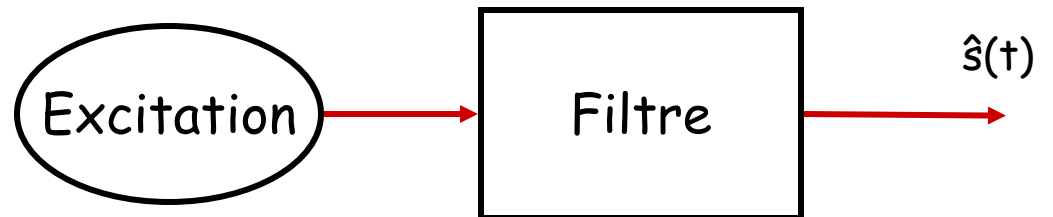
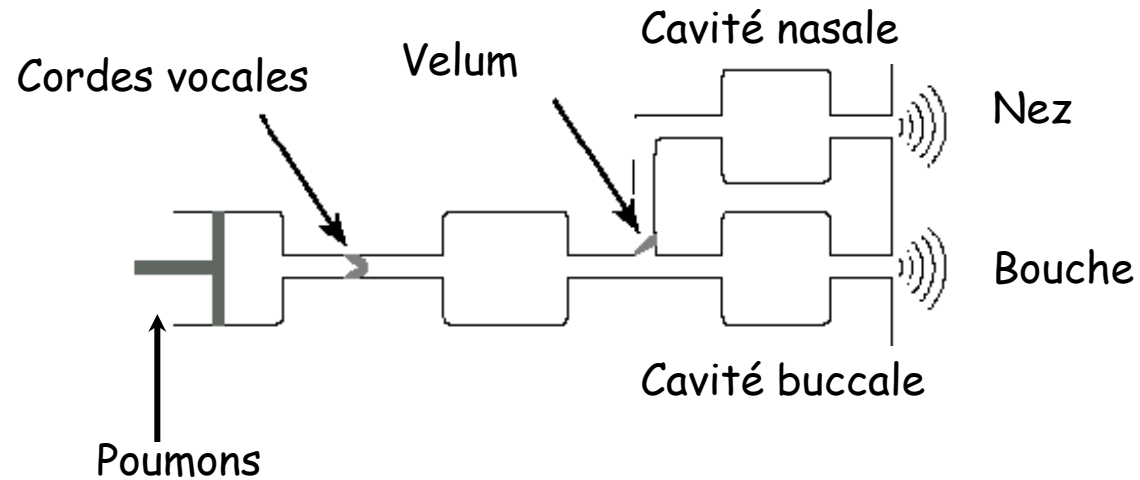
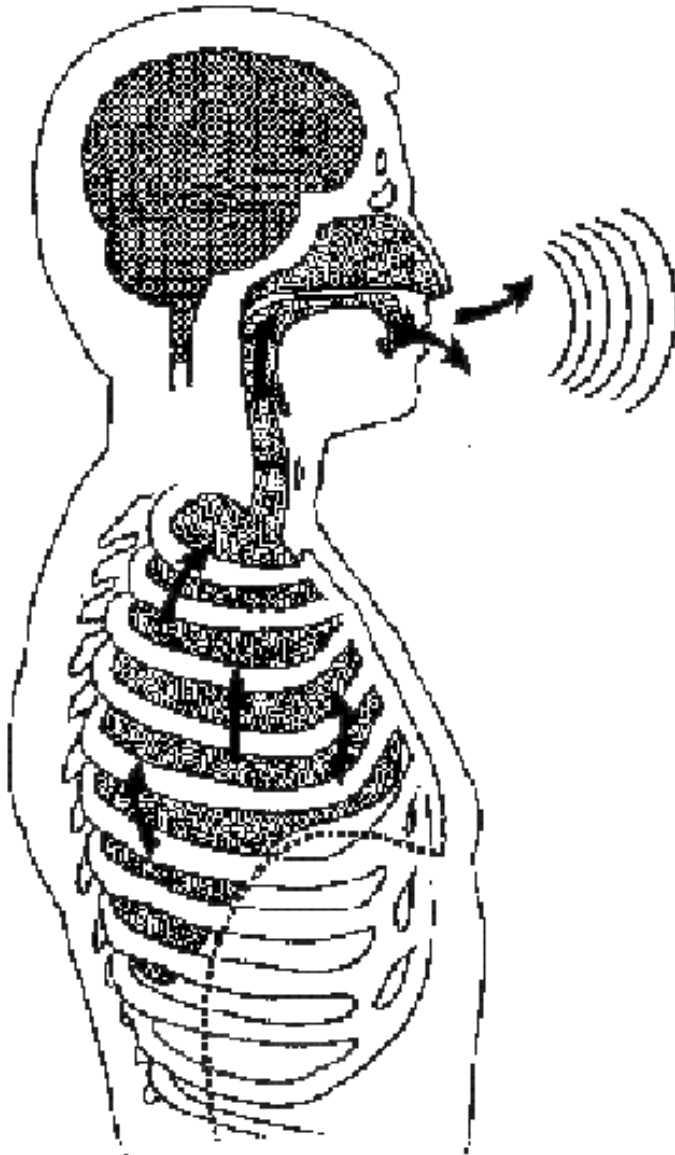
Comment faire pour réduire le débit ?



Modèle de synthèse

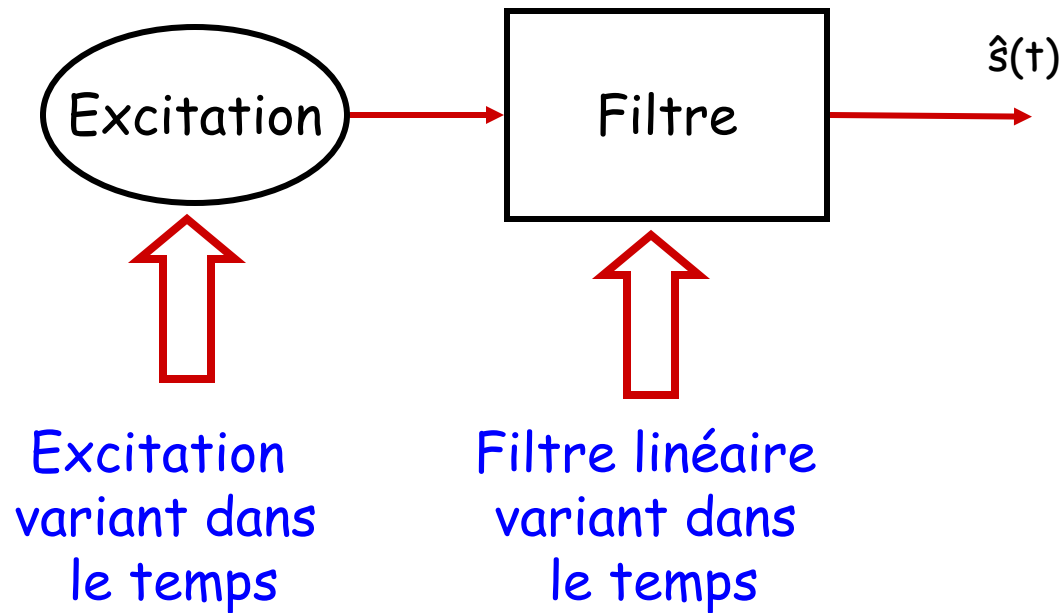
Approche « analyse/synthèse » (« vocoding »)

Imitons la nature !



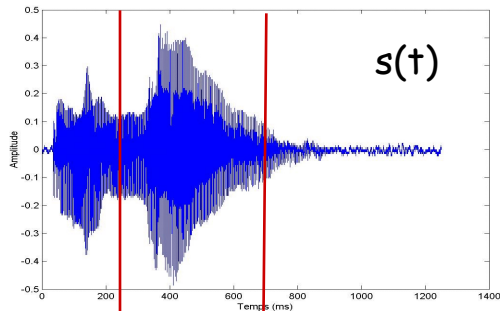
Modèle synthétique

Le modèle synthétique



Approche Analyse/Synthèse

L'approche A/S



Wt
20-35ms

Paramètres

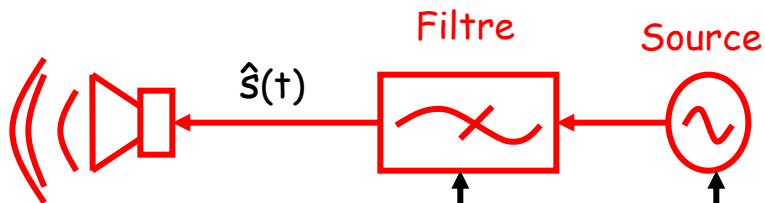
Source
Modèle

Filtre

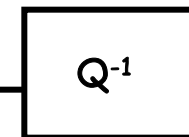
$$\hat{s}(t) \sim s(t) : |\hat{s}(t) - s(t)|^2 < \varepsilon$$



00110101001.....



Paramètres



00110101001.....



Plan (v)

- I. Conversion analogique/numérique
- II. Les paramètres fondamentaux d'un CP
- III. Codage PCM
- IV. Le codage par analyse-synthèse
- V. Exemples**
- VI. Etude comparative

Le codeur G.711

Caractéristiques

- ✓ Il est le standard international pour le codage du réseau RTC;
- ✓ Débit : 64 kbps;
- ✓ C' est un codeur PCM à 8 kéch/s;
- ✓ Résolution 8 bits par échantillon;
- ✓ Il suit les recommandations ITU-T;
- ✓ Entrée PCM uniforme ou compressée loi-A / loi μ ;
- ✓ Sortie PCM uniforme ou compressée loi-A / loi μ ;
- 1. Le code à été testé sur les plate-formes suivantes :
 - 1. *TI TMS320C54x*
 - 2. *Plate-formes RISC,*
 - 3. *MS Windows.*

Le codeur RPE-LTP du GSM

Caractéristiques

1. Développé par les Groupe Spécial Mobile et normalisé par l'ETSI;
2. Débit binaire 13 kb/s;
3. Excitation par impulsions régulières (« Régular Pulse Excited »);
4. Filtrage LPC et prédiction à long terme (« Long Term Prediction »);
5. Qualité perceptuelle comparable à celle du G.711;
6. Il supporte simultanément des services de parole et données;
7. Complexité algorithmique relativement faible

Le codeur EFR du GSM

Caractéristiques

1. Version améliorée du RPE-LTP et normalisé par l'ETSI;
2. Particularité : Débit variable !
3. Débit binaire : 12.2 kb/s;
4. Même débit global que celui du codeur RPE-LTP;
5. Excitation de type CELP algébrique; ;
6. Qualité RTC;
7. Il supporte DTX grâce à un détecteur d'activité vocale (VAD);
8. Testé et validé sur pratiquement toutes les plateformes commerciales.

Le codeur AMR du GSM

Caractéristiques

1. C' est le codeur proposé pour l' UMTS et normalisé par l' ETSI;
2. Particularité : Débit variable !
3. Le débit est adaptatif en fonction de la charge du réseau;
4. Débits binaires : 12.2, 10.2, 7.95, 6.7, 5.9, 5.15 et 4.75 kb/s;
5. Filtre LPC d' ordre 10 avec filtrage court terme (20ms et sous-trames de 5ms);
6. Excitation adaptives de type CELP algébrique; ;
7. Retard algorithmique 7.5 ms;
8. Qualité RTC pour les hauts débits;;
9. Il supporte DTX grâce à un détecteur d' activité vocale (VAD);
10. Testé et validé sur pratiquement toutes les plateformes commerciales.

Le codeur G728

Caractéristiques

1. Débit binaire : 16kbps;
2. Il utilise une excitation de type CELP à très faible retard (Low Delay Code Excited Linear Prediction - LD-CELP).
3. Il utilise adaptation « backward », pour atteindre un retard algorithmique record de 0.625 ms.
4. En conditions de transparence du canal il a la qualité perceptuelle d'un PCM à 32 kb/s;
5. Excellente robustesse au bruit de fond et à la musique;
6. Il supporte très bien les tonalités DTMF;
7. Il tolère très bien les modems numériques à faible débit;
8. Particulièrement bien adapté pour des applications VoIP;
9. Utilisé dans les vidéotéléphones suivant la norme H.320

Le codeur G723.1

Caractéristiques

1. C' est un codeur du standard ITU;
2. Double débit : 5.3 et 6.3 kbps.
3. Filtre LPC d' ordre 10 pour les deux débits;
4. Excitation Multi-pulse pour 6.3, CELP algébrique pour le 5.3;
5. Il code des trames de 30 ms;
6. Retard algorithmique 7.5 ms;
7. Faible distorsion pour les tonalités DTMF;
8. Entrée/sortie codées sur du PCM 16 bits uniformes;
9. Le meilleur taux de codage du marché avec la plus faible distorsion;
10. Applications : compression full duplex pour le multimédia, vidéo-téléphonie et vidéo- conférence;
11. Il supporte DTX grâce à un détecteur d' activité vocale (VAD);
12. Testé et validé sur pratiquement toutes les plateformes commerciales

Le codeur MELP du FS

1. Il fait partie du « Federal Standard » des USA;
2. Débit binaire : 2.4 kb/s
3. Il remplace les codeurs FS1015 (2.4 kb/s) et le FS1016 (4.8 kb/s);
4. Il utilise un mélange d'excitations (« Mixed Excited Linear Prediction »);
5. Il est particulièrement robuste au bruit de fond;
6. Il est très bien adapté aux erreurs de transmission;
7. Il est inter opérable avec d'autres codeurs, d'autres systèmes;
8. La qualité perceptive supérieure à celle du FS1016;
9. Complexité et retard algorithmique importants;
10. Nécessite des architectures adaptées.

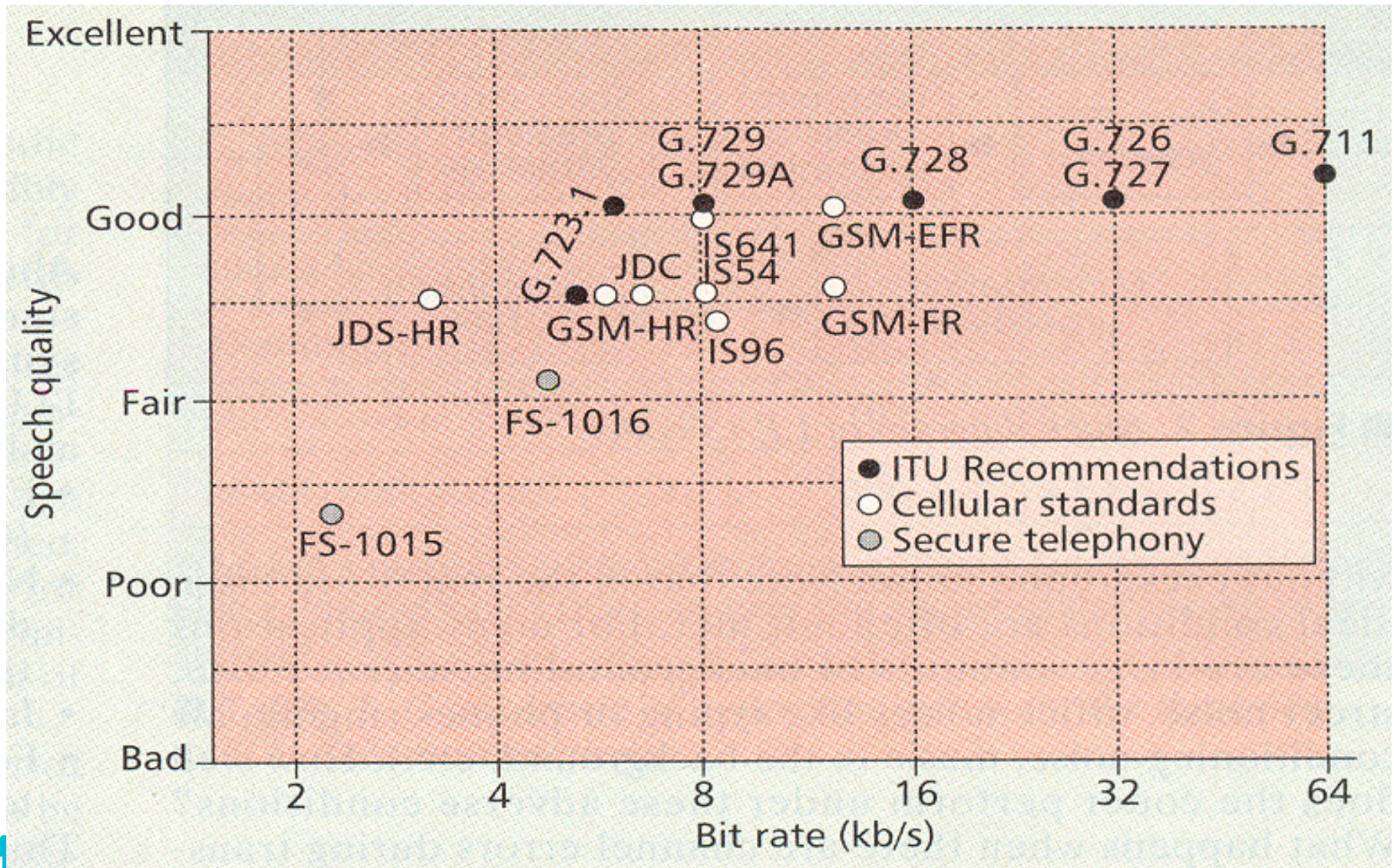
Plan (VI)

- I. Conversion analogique/numérique
- II. Codage PCM
- III. Les paramètres fondamentaux d'un CP
- IV. Le codage par analyse-synthèse
- V. Exemples
- VI. Etude comparative**

Tableau comparatif

Standard	G711	G728	G723.1	RPE-LTP	FS1015	FS1016	MELP
Débit (kb/s)	64	16	5.3, 6.3	13	2.4	4.8	2.4
Qualité	RTC	RTC	<RTC	~RTC	<<RTC	<RPELTP	>FS1016
Complexité MIPS RAM(kO)	<<1 1 Octet	30 2	20 2	4.5 1	20 2	19 2	<80 ?
Retard (ms)	0.125	0.625	40	20	110	40	?
Type	PCM	CELP	CELP	RPELTP	LPC	CELP	MELP

Performances



Fin