

# C3

## Analyse de la parole (1)

1. Amplitude moyenne et maximale
2. Densité d'amplitude
3. L'énergie Moyenne
4. L'énergie TEAGER
5. Nombre de passages par zéro
6. Fréquence fondamentale (F0)
7. Codage TESPAP

# TRAITEMENT DU SIGNAL VOCAL

**1. Analyse** - *consiste à l'acquisition de la parole, le filtrage, l'échantillonnage et le traitement des échantillons pour obtenir certains paramètres et caractéristiques acoustiques*

**2. Synthèse** - fait la voix artificielle des éléments spécifiques du langage (les phonèmes, demi-syllabes, prosodie etc.) obtenu dans la phase d'analyse

**3. Reconnaissance** - repose sur l'extraction des paramètres du signal et les caractéristiques de la voix et en utilisant des algorithmes pour prononcer le discours (phonème, mot, phrase) prononcés par comparaison avec certains modèles ou des références

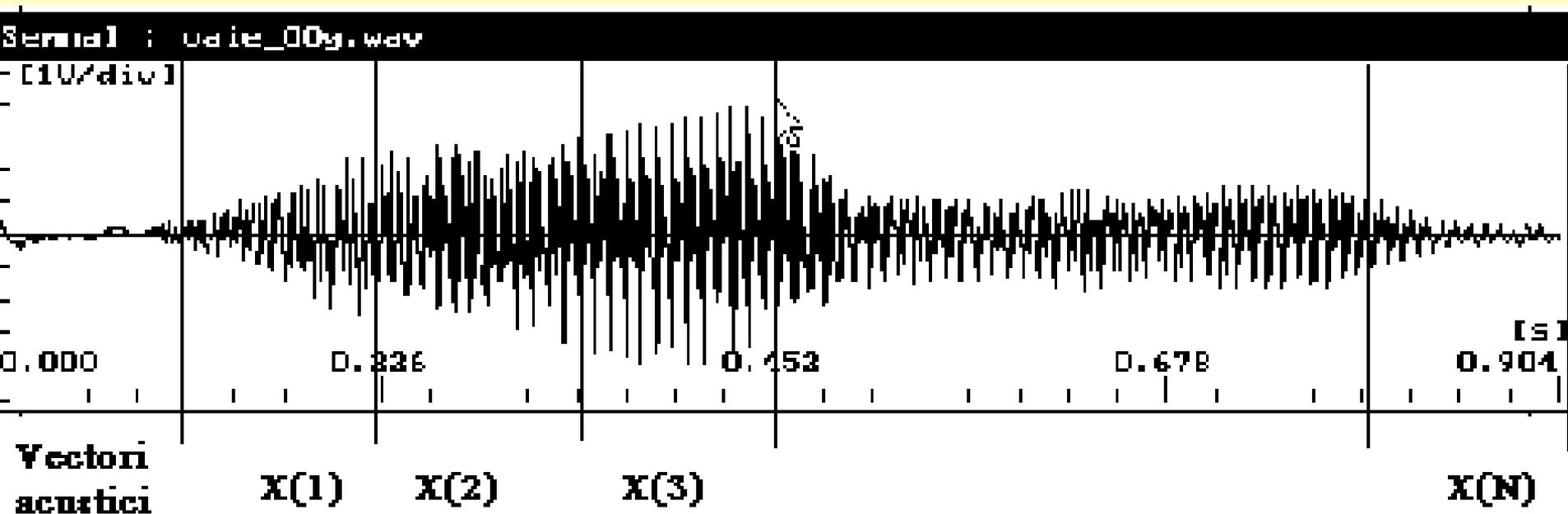
**4. Compression/codage** signal vocal est de réduire la redondance du signal vocal à accroître l'efficacité de la transmission ou le stockage par l'extraction de paramètres et les caractéristiques de la parole

<https://www.clear.rice.edu/elec631/Projects99/mit/index2.htm>

<https://www.mathworks.com/help/signal/ug/practical-introduction-to-time-frequency-analysis.html;jsessionid=b84ab6d49bf2a6cdfd7270d2db82>

**ANALYSE du SV** : consiste en l'acquisition, le filtrage, l'échantillonnage et traitement des échantillons pour obtenir certains paramètres et les caractéristiques acoustiques (vecteurs acoustiques) de la représentation du SV efficace, économique et facile à utiliser dans le traitement ultérieur

**L'analyse a court terme** - le signal vocal est divisé en cadres successives, avec une longueur de  $\sim 10-30$  ms, pendant laquelle le SV est considéré **quasi-stationnaire**



# ANALYSE de la parole

## Domaine Temporel:

1. Amplitude moyenne et maximale
2. Densité d'amplitude
3. L'énergie moyenne
4. Energie TEAGER
5. Nombre de passages par zéro
6. Fréquence fondamentale (FO)
7. Codage TESPAP

## Domaine des fréquences:

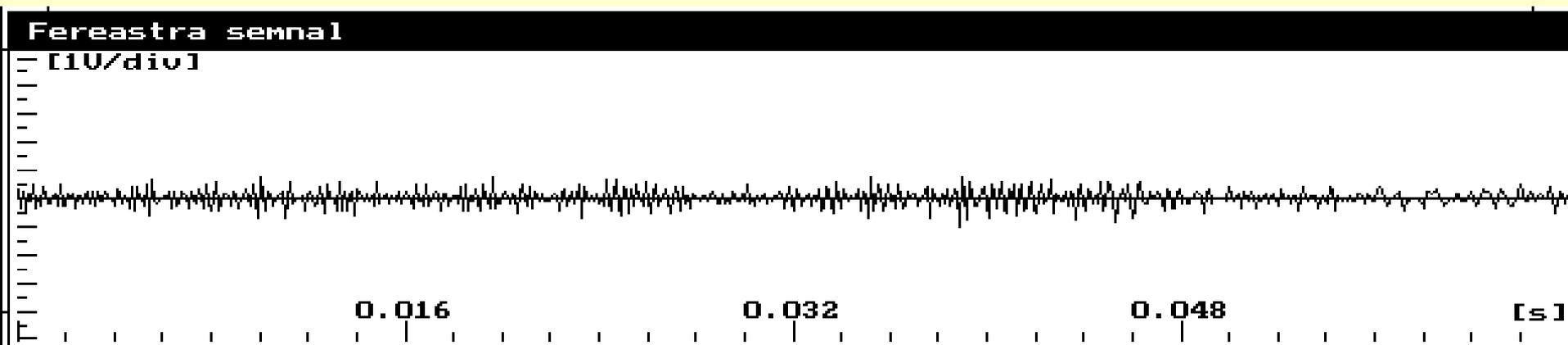
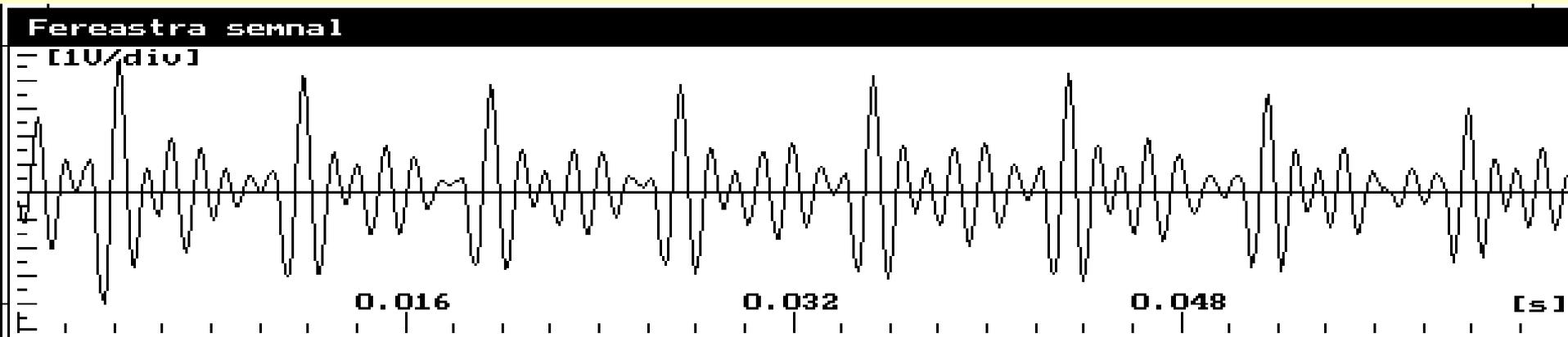
- TFD (TFR)
- Analyse par la PL
- Banc de filtres numériques
- Analyse cepstrale
- Analyse perceptuelle

## Analyse temps-frequence

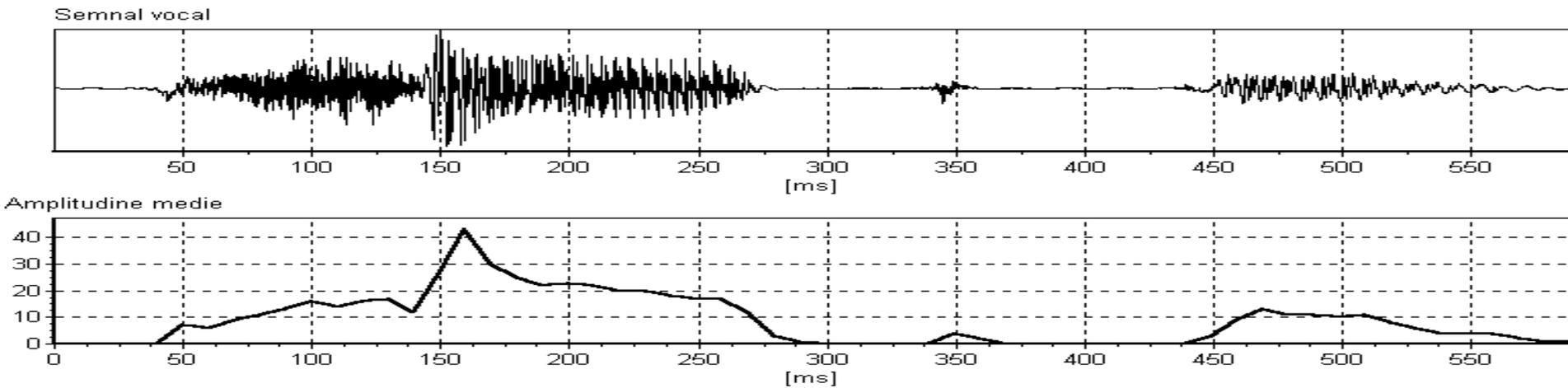
- transformée de Fourier de courte durée (STFT)
- la transformée en ondelettes discrète (Haar) (DWT)
- la transformée en ondelettes continue (Morlet) (CWT)
- la distribution pseudo-Wigner

# 1. Amplitude moyenne et maximale

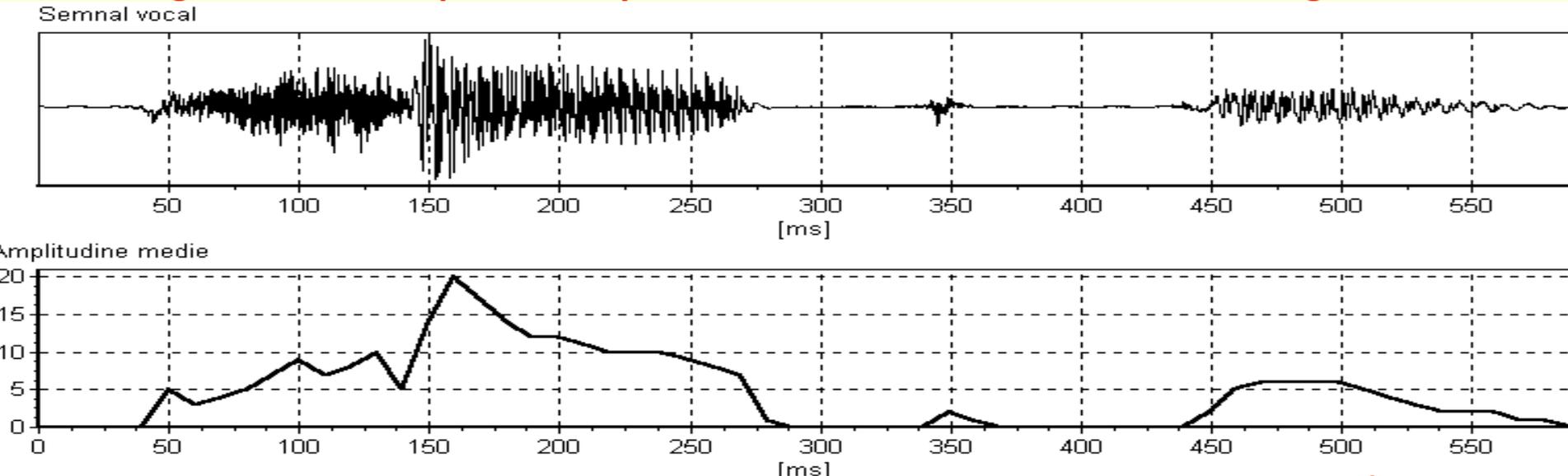
- *L'amplitude maximale et l'amplitude moyenne* d'un segment donne des informations qualitatives sur la présence ou l'absence de la voix ou du bruit ou si le segment analysé est voisée ou non voisée



$$M(n) = \frac{1}{N} \sum_{m=-\infty}^{\infty} |x(m)| \cdot w(n-m) \quad , \text{ou} \quad w(n) = \begin{cases} 1, & 0 \leq n \leq N-1 \\ 0, & \text{în rest} \end{cases}$$

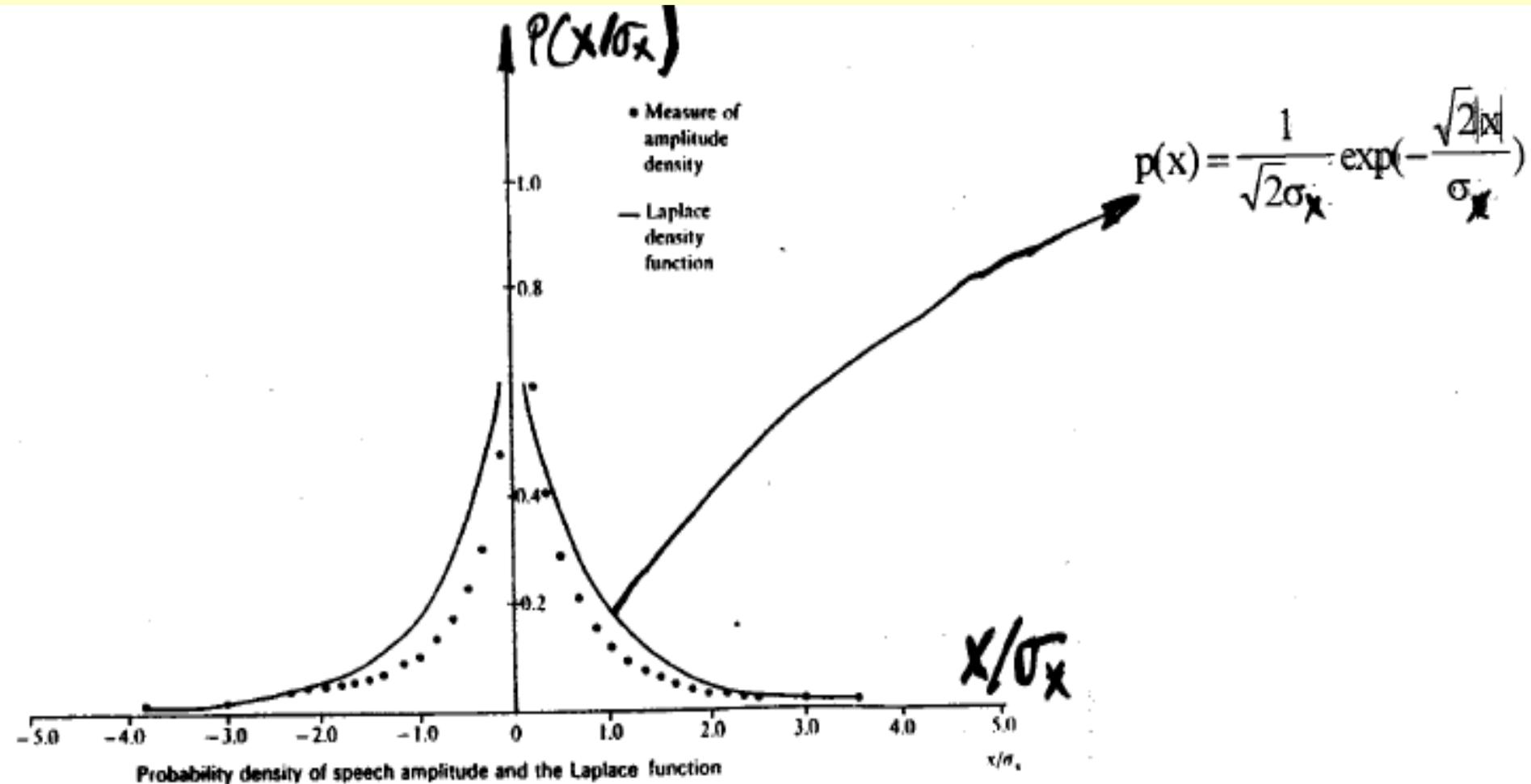


**Signal vocal et amplitude moyenne du SV en utilisant la fenêtre rectangulaire**



**Signal vocal et amplitude Moyenne du SV en utilisant la fenêtre *Hamming***

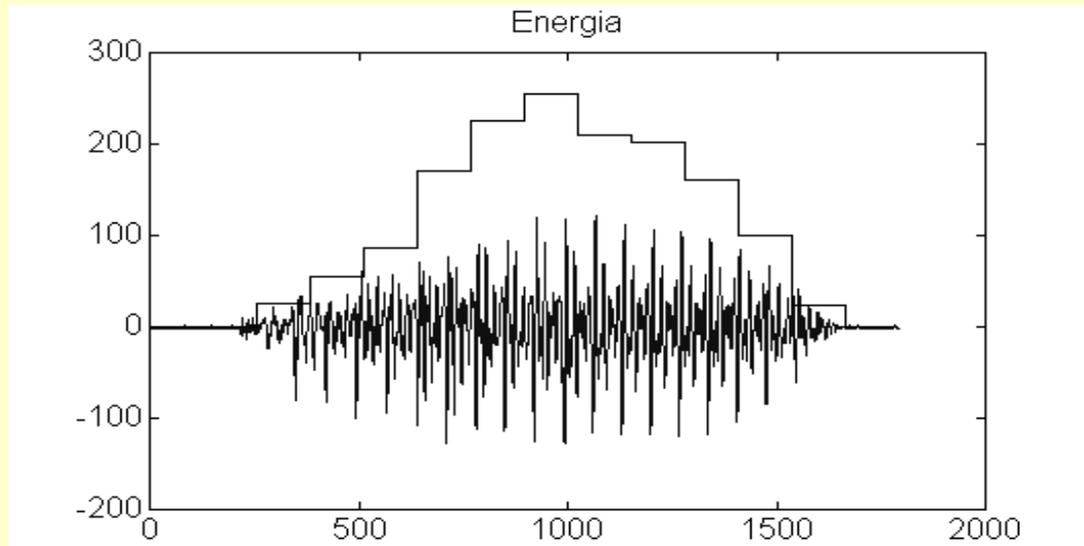
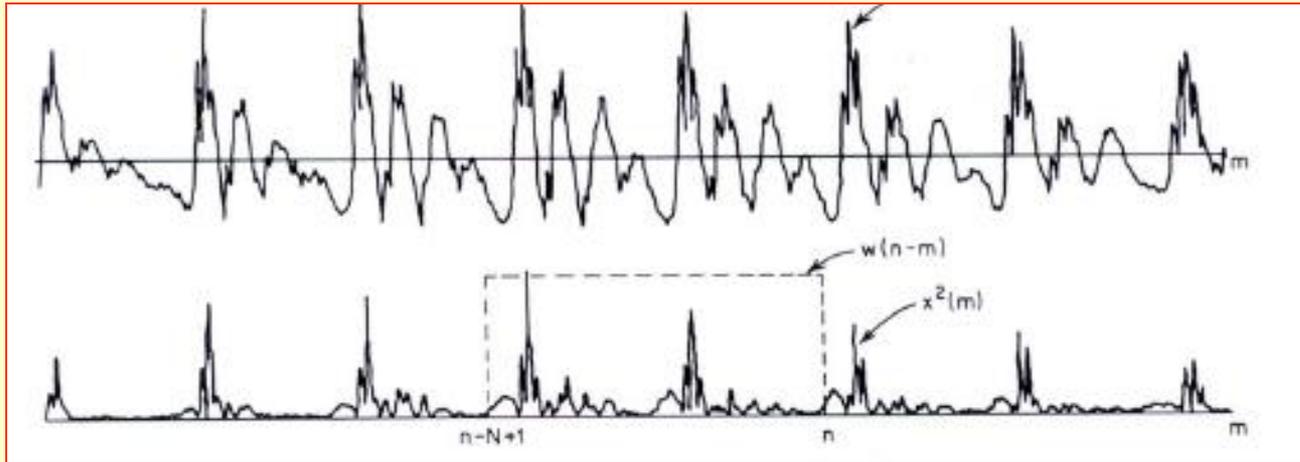
## 2. Densité d'amplitude

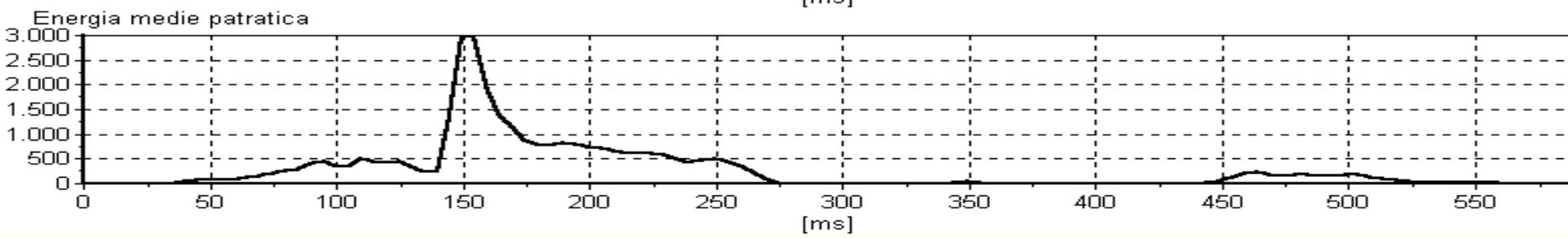
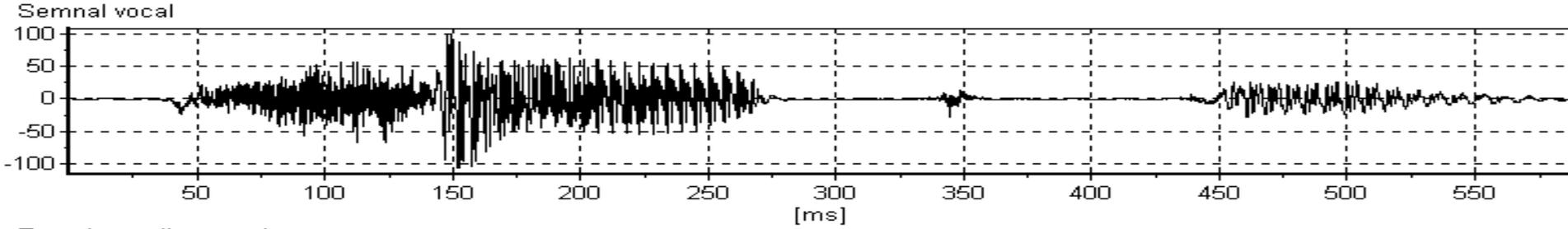


- Distribution statistique des échantillons d'amplitude à long terme
- La fonction densité de l'amplitude (asymétrie - locuteur, microphone, ..)
- Permet d'étalonnage correcte du signal a l'entrée du CAN (AGC)

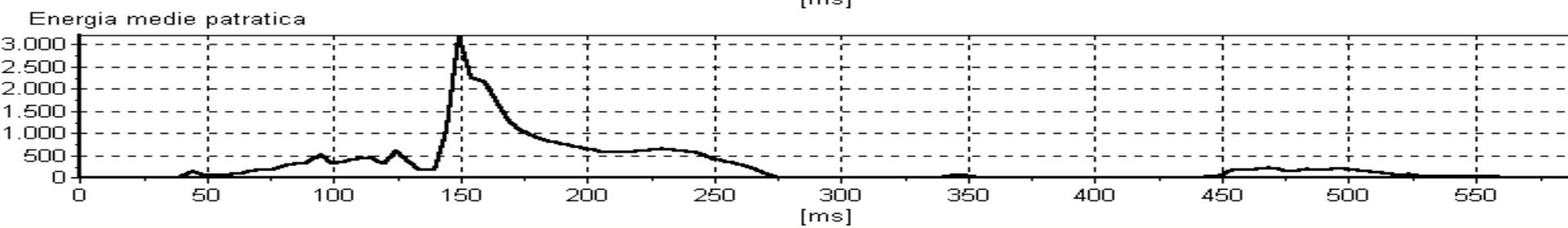
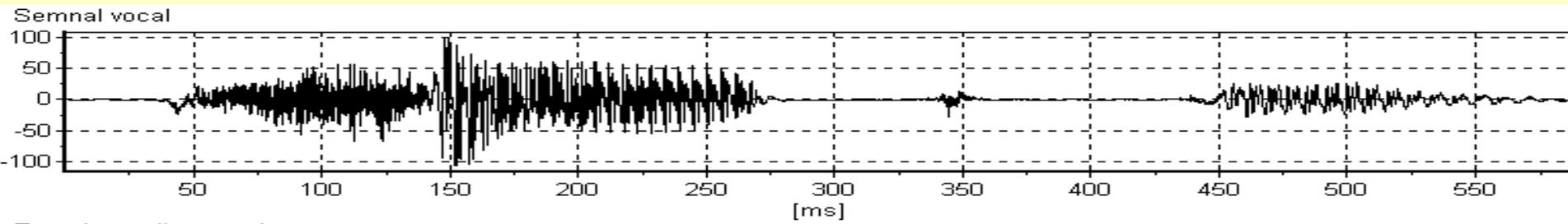
### 3. Énergie moyenne quadratique

$$E(n) = \frac{1}{N} \sum_{m=-\infty}^{\infty} [x(n)w(n-m)]^2$$

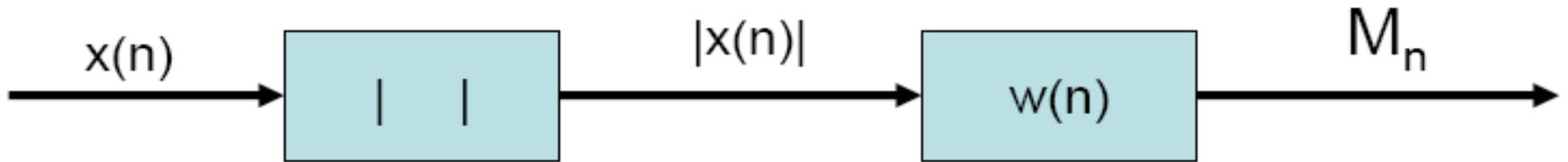




*Signal vocal et l'énergie moyenne quadratique, en utilisant une fenêtre rectangulaire de 10ms et d'un chevauchement de 50%*



*Signal vocal et l'énergie moyenne quadratique, en utilisant une fenêtre Hamming de 10ms et d'un chevauchement de 50%*



$$M_n = \sum_{m=-\infty}^{\infty} |x(m)| w(n-m)$$

-L'énergie moyenne à court terme est très sensible à grandes amplitudes,  $\sim x^2(n)$ , l'amplitude moyenne est utilisée en place

## 4. Opérateur d'énergie Teager

- Teager introduit un nouveau mode de calcul de l'énergie instantanée, comme suit:

$$E_i = x_i^2 - x_{i+1}x_{i-1}$$

En cas de signaux harmoniques :

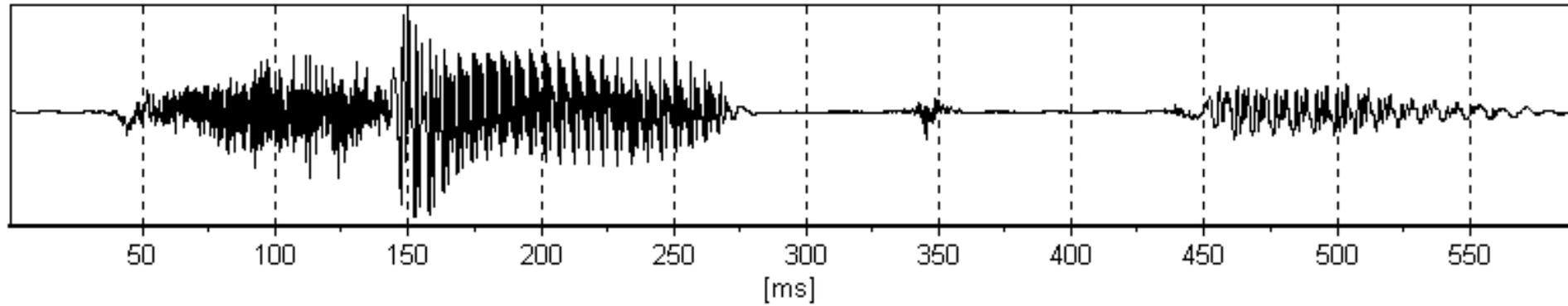
$$x_i = A \cos(\Omega i + \Phi)$$

L'énergie instantanée Teager a la forme :

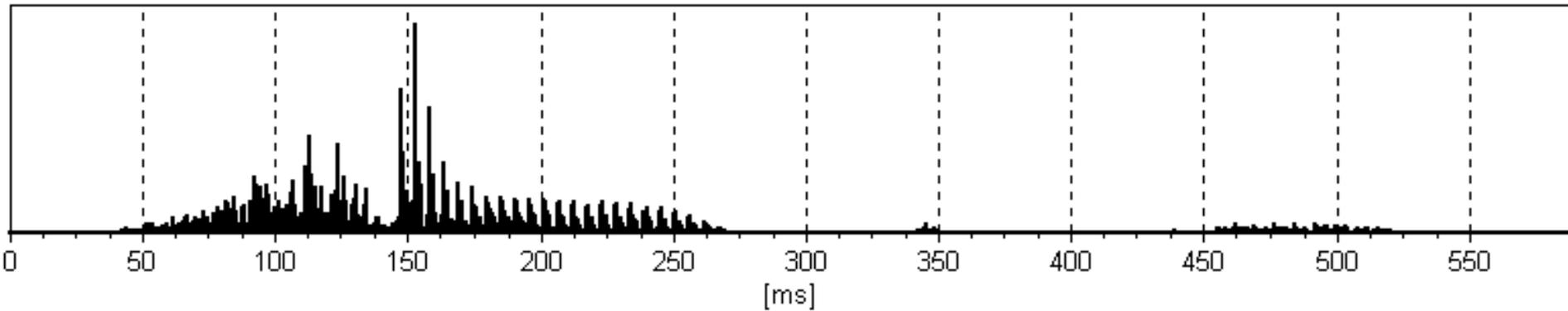
$$E_i = A^2 \sin^2(\Omega) \approx A^2 \Omega^2$$

- cette forme d'énergie instantanée a de variations à la fois l'amplitude et la fréquence du signal vocal
- Cette propriété peut être utilisée pour distinguer les sons fricatives et plosives de bruit
- opérateur Teager peut être utilisé pour séparer les zones de silence et parole dans un signal vocal

Semnal vocal



Energie instantanee Teager



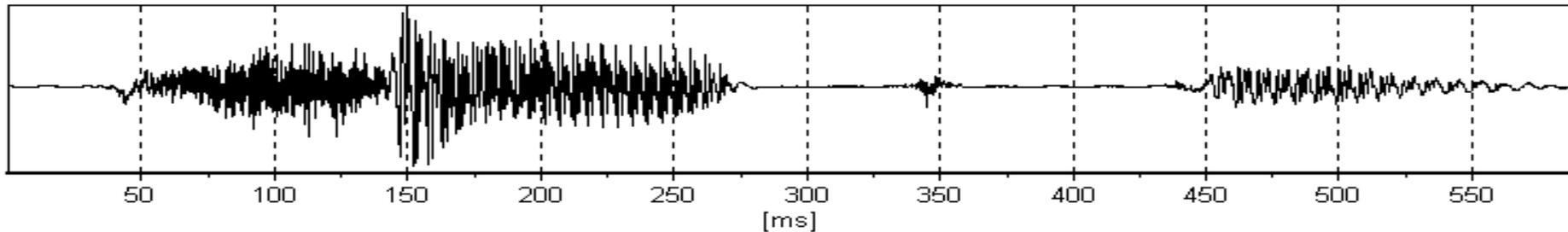
*SV et énergie TEAGER instantanée*

# L'énergie moyenne Teager

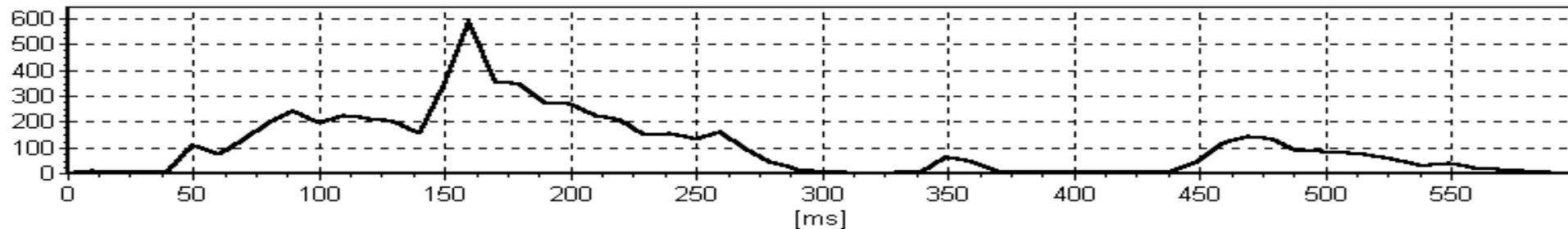
- En cas de l'analyse à court terme, l'énergie moyenne peut être évaluée par sommation et de la médiation (comme dans le cas précédent), soit, tenir compte des propriétés de la transformée de Fourier discrète, TFD, comme suit:

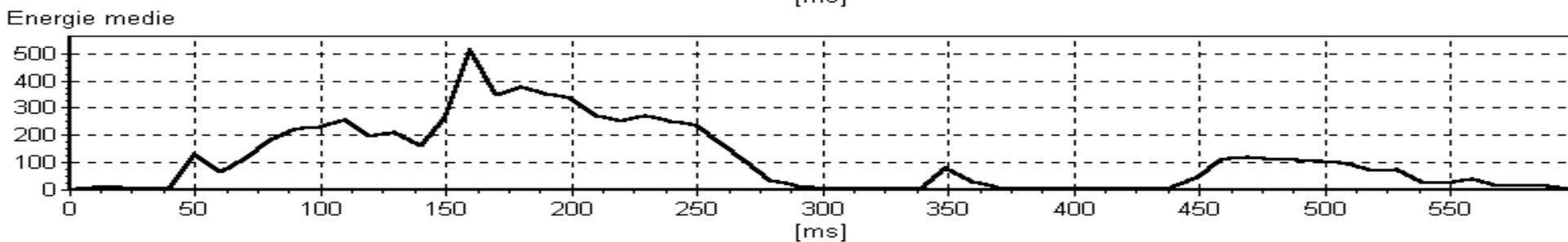
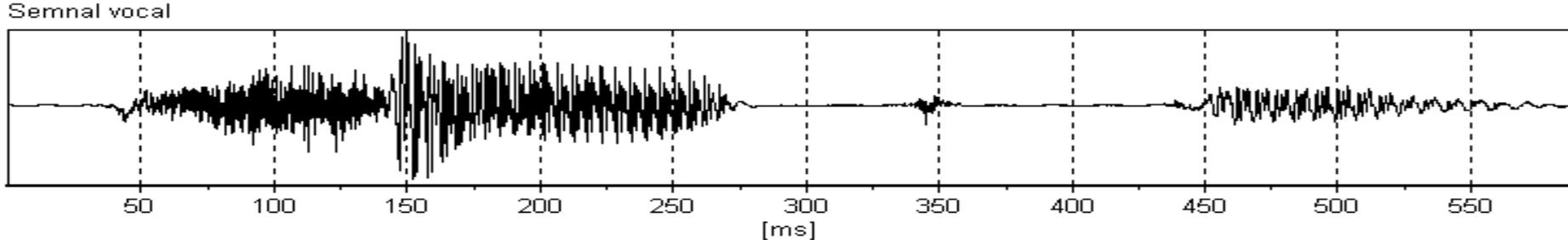
$$E = \sqrt{\sum_{i=1}^N \omega_i^2 |X(\omega_i)|}$$

Signal vocal

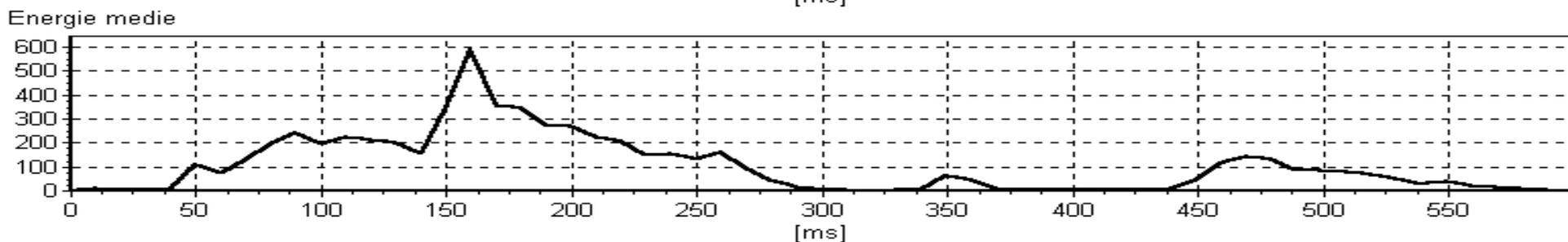
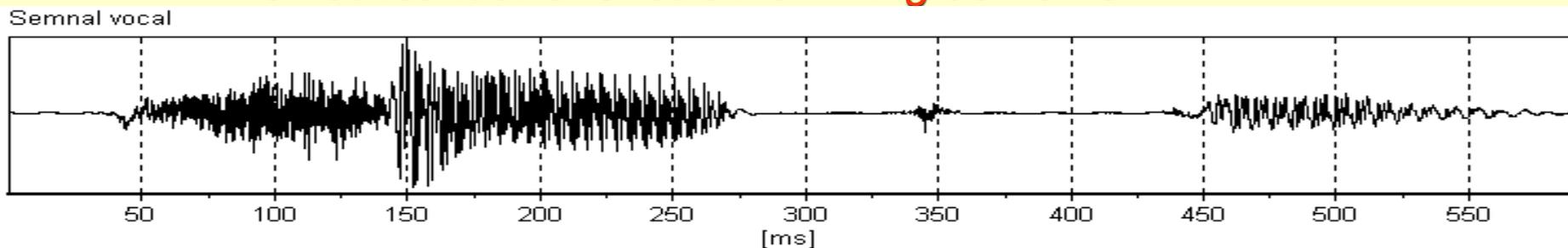


Energie medie





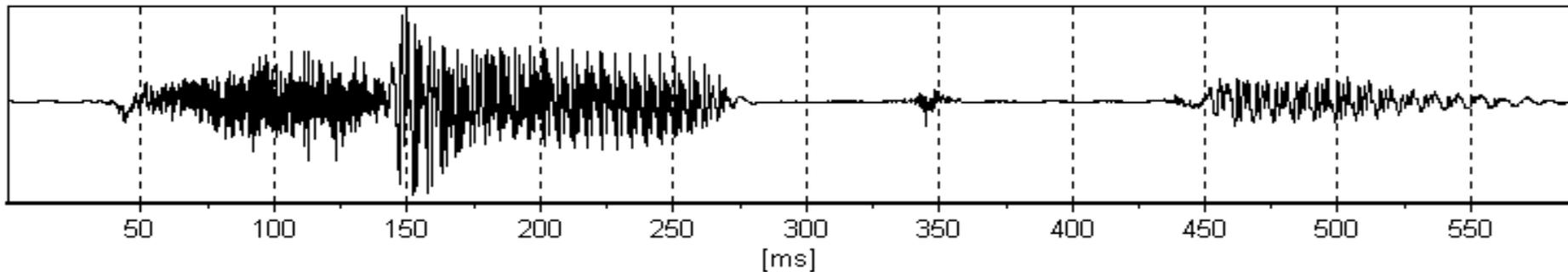
**SV et énergie moyenne TEAGER**  
*en utilisant une fenêtre Hamming de 10ms*



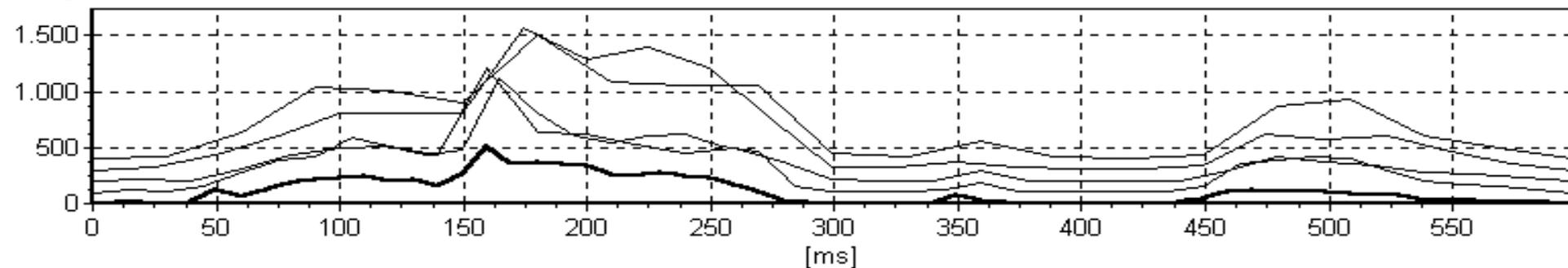
**SV et énergie moyenne TEAGER**  
*en utilisant une fenêtre rectangulaire de 10ms*

- La longueur de la fenêtre influence fortement l'énergie moyenne Teager ainsi, les fenêtres plus longues peuvent cacher les détails de la dépendance en fréquence d'énergie moyenne Teager

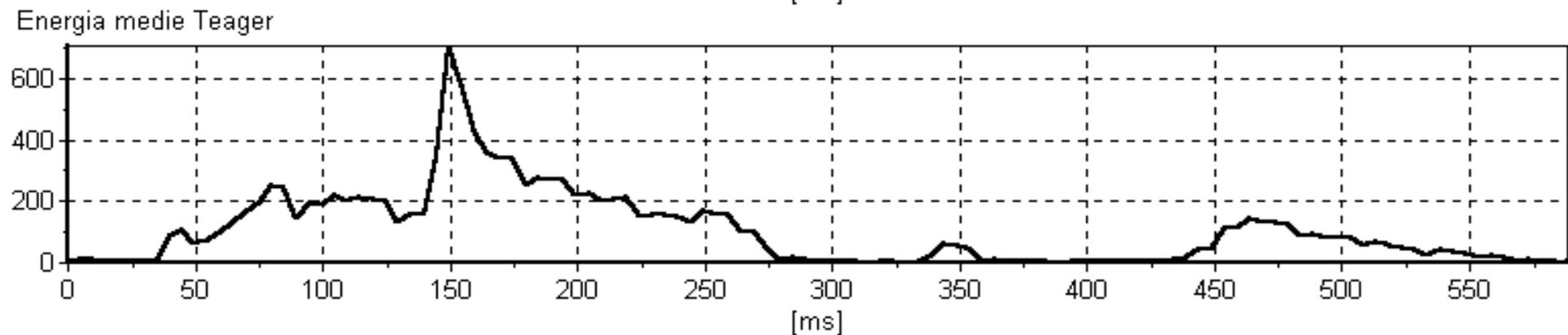
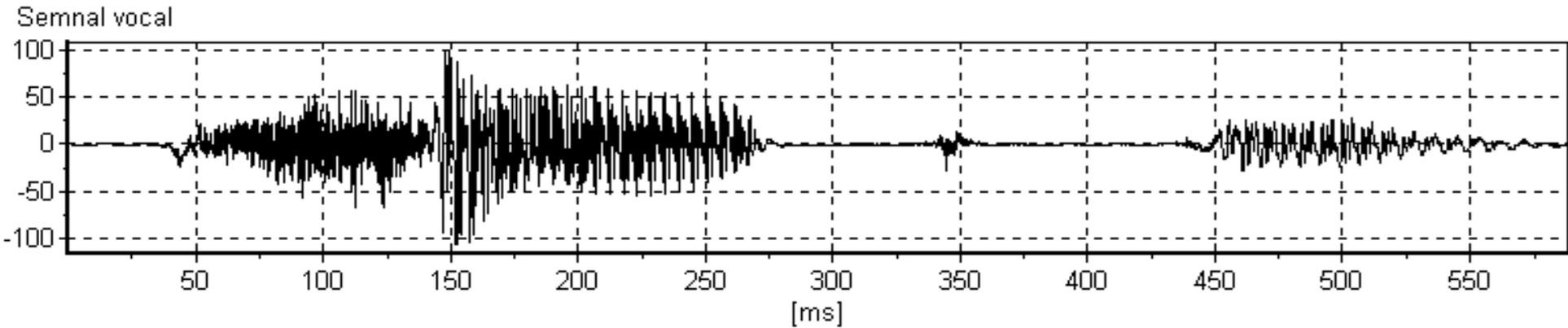
Signal vocal



Energie medie



*SV et énergie moyenne TEAGER en utilisant une fenêtre Hamming normalisée, avec des longueurs de 10ms, 15ms, 20ms, 25ms, 30ms*



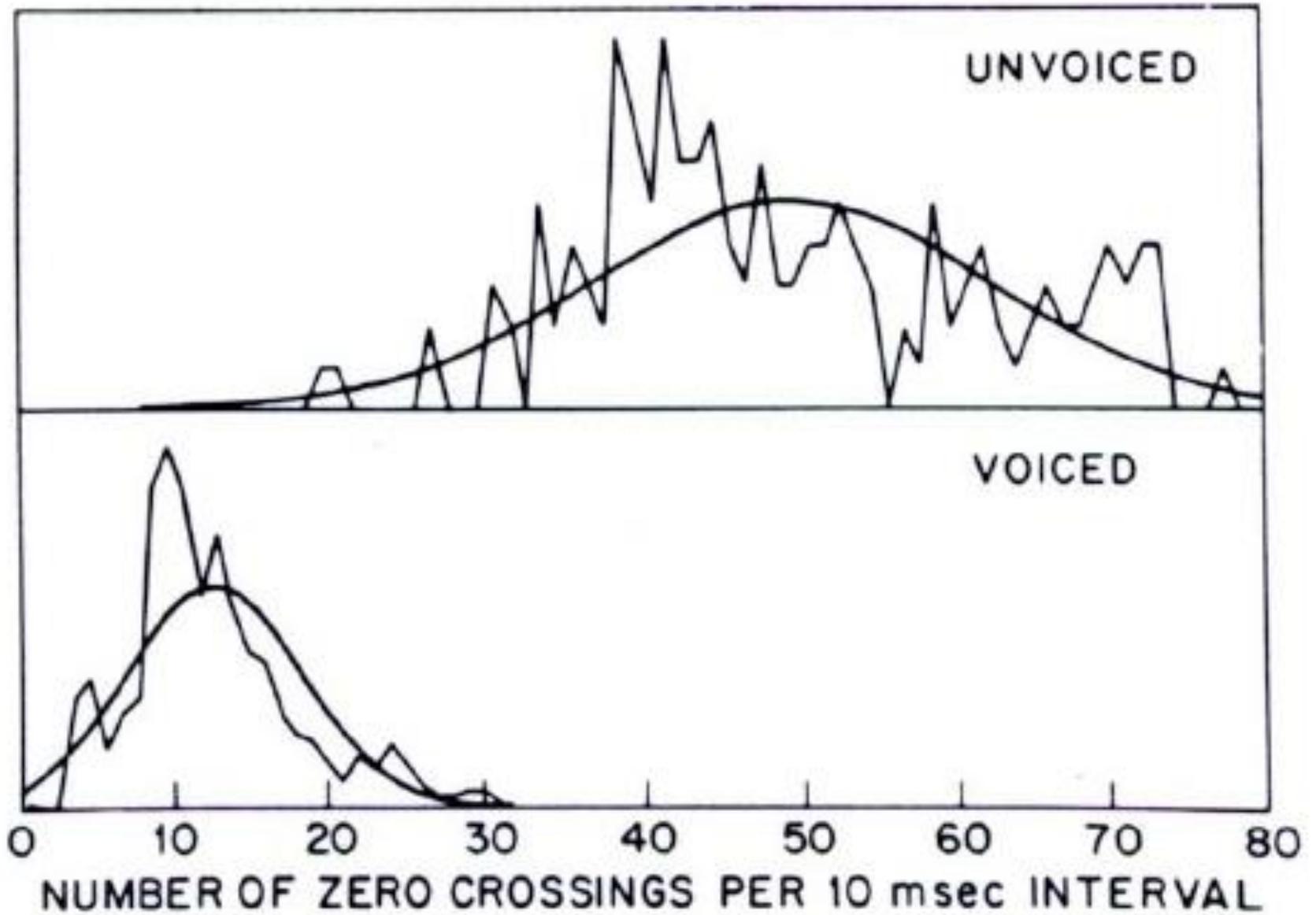
***SV et énergie moyenne TEAGER en utilisant une fenêtre Hamming de 10ms et d'un chevauchement de 50%***

## 5. Nombre de passages par zéro

Nombre de passages par zéro (NPZ=ZCR) est une approximation du contenu spectral du signal vocal

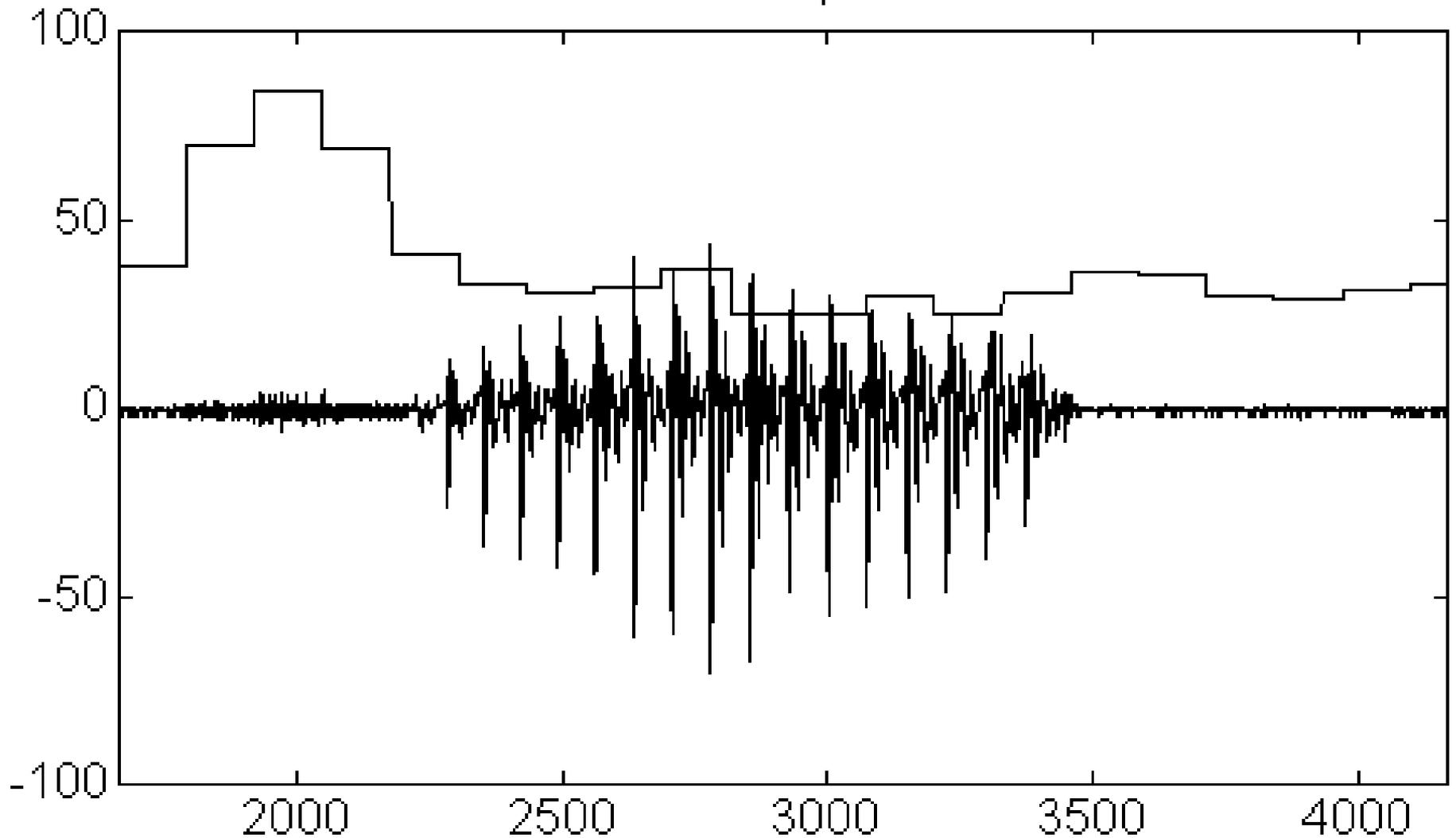
$$\text{NTZ} = \frac{\sum_{n=0}^{N-1} [1 - \text{sgn}(s(n+1)T) \cdot \text{sgn}(s(nT))]}{2}$$
$$\text{sgn}(s(nT)) = \begin{cases} +1, & s(nT) \geq 0 \\ -1, & s(nT) < 0 \end{cases}$$

- NPZ  $\sim F_{es}$  et est affectée par le bruit et l'offset (FPBass) et on fait une normalisation aux NPZ<sub>max</sub>



Distribution du NPZ dans des segments voisée ou nonvoisée

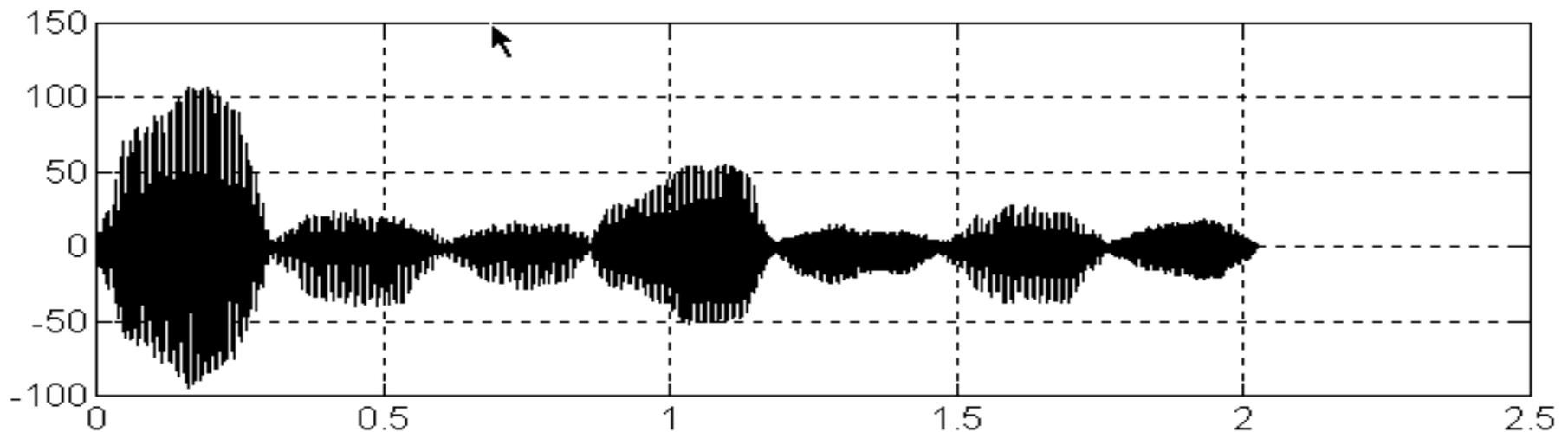
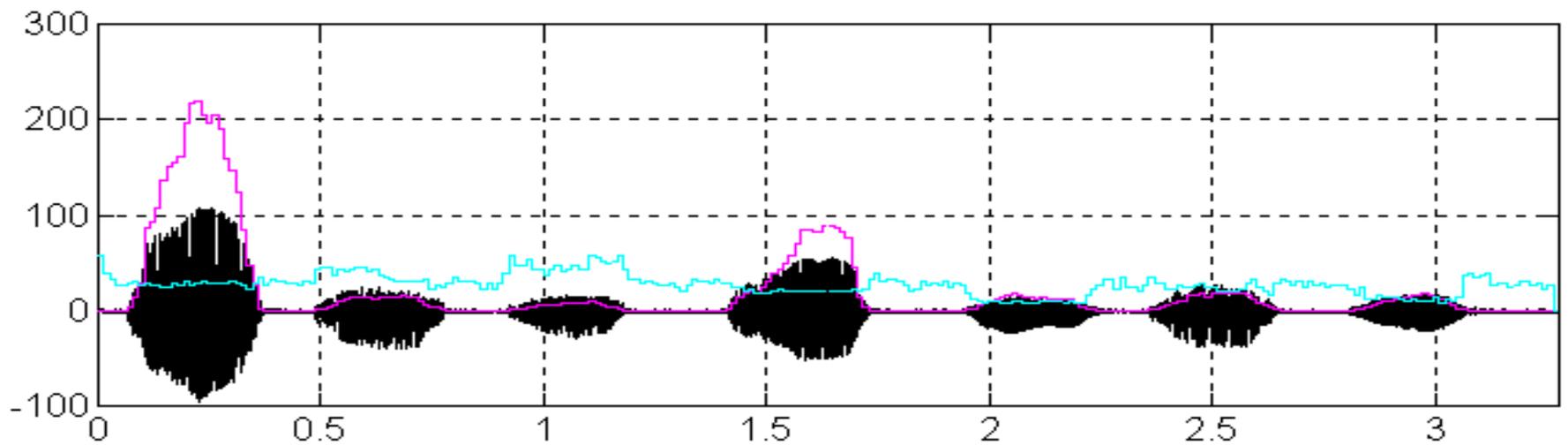
## Numărul trecerilor prin zero - NTZ



**L'Energie + NPZ peuvent nous donner des informations sur la nature du signal: voisée / non voisée, parole / silence**

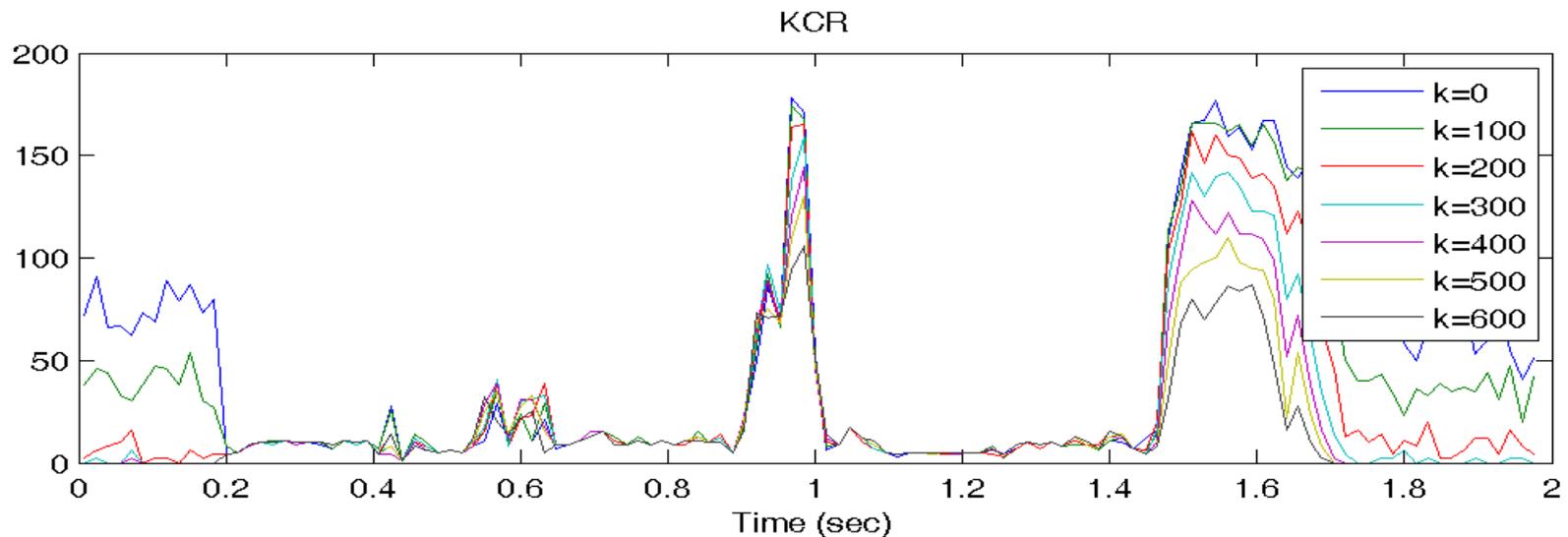
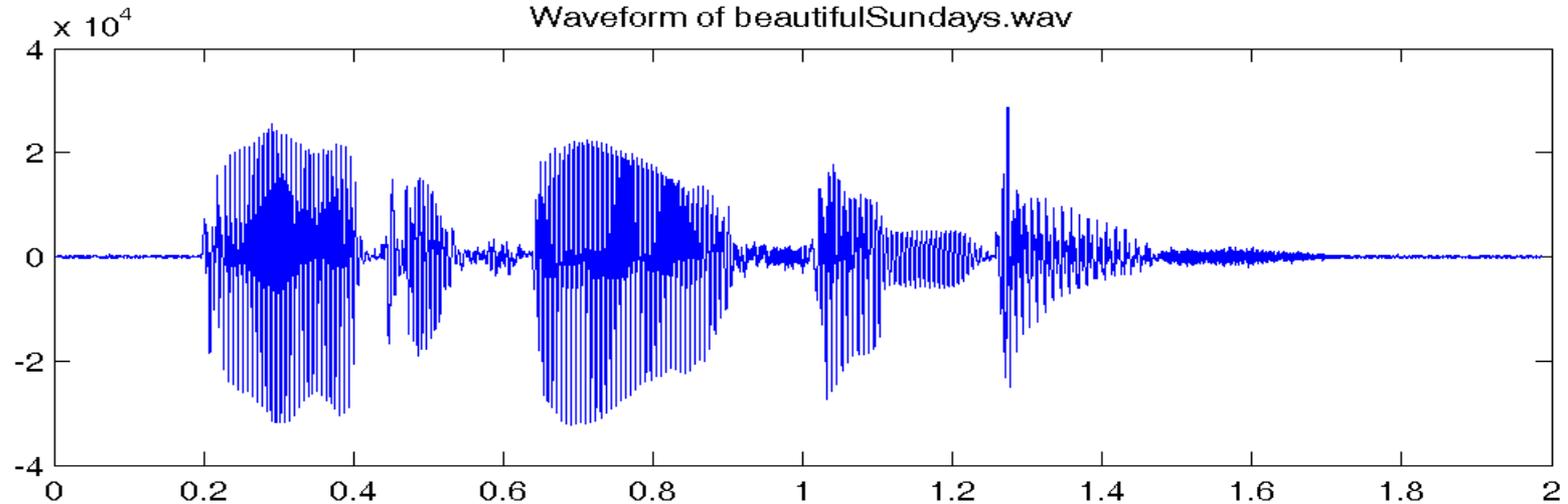
<b>SEGMENT</b>	<b>ENERGIE</b>	<b>NPZ</b>
<b>voisée</b>	<b>grosse</b>	<b>petite</b>
<b>non voisée</b>	<b>moyenne /petite</b>	<b>moyenne/petite</b>
<b>bruit</b>	<b>~0</b>	<b>gros</b>
<b>silence</b>	<b>~0</b>	<b>~0</b>

*La classification des segments selon l'énergie et NPZ*



**Séparation du signal vocal de la silence en cas de l'énonciation des voyelles isolées**

- Nous pouvons étendre la définition de ZCR à KCR (k-pass rate), qui est le nombre de passes sur  $y = k$ , dans une trame de SV



## 6. Méthodes de détermination de la fréquence fondamentale F0 du signal vocal

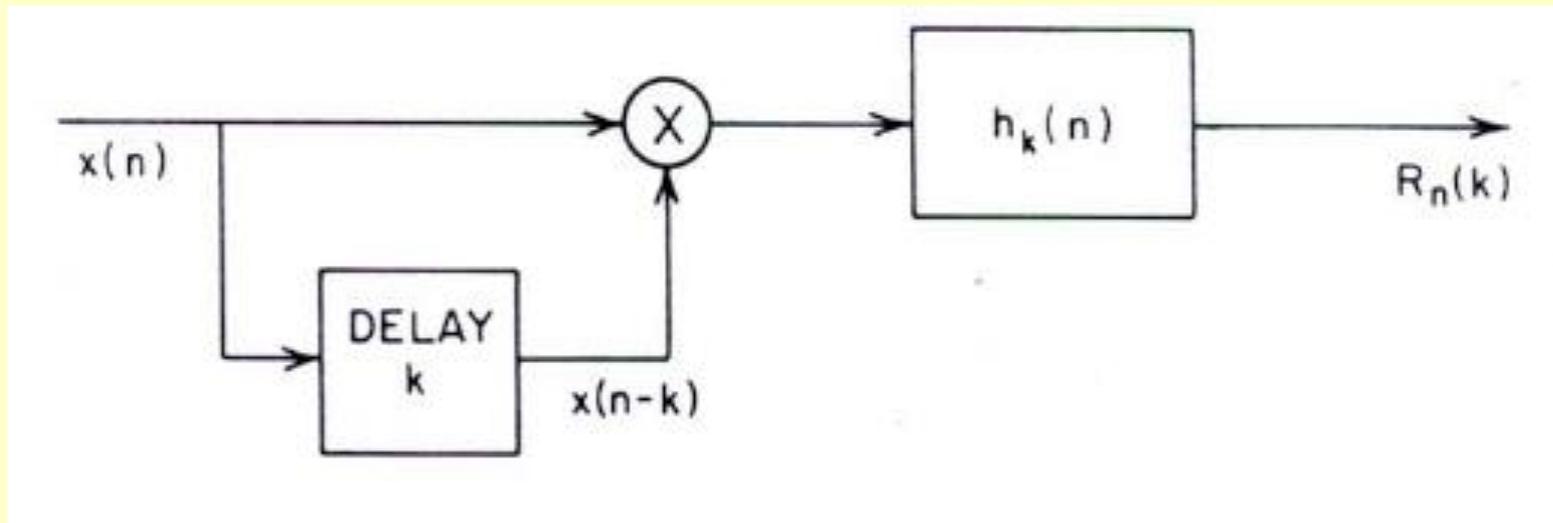
- Détermination de la valeur précise de la fréquence fondamentale est difficile pour les raisons suivantes:
  - les vibrations des cordes vocales ne sont pas tenues à une fréquence constante, en particulier aux extrémités de portions voisées;
  - l'excitation glottale n'est pas un train d'impulsions parce que les *changements de la forme glottale* se font peu à peu, à chaque impulsion (amplitude, forme, période);
  - *C'est difficile de séparer le signal vocal*: l'effet de la source et les effets du conduit vocal;
  - la *gamme de la fréquence F0* a des variations très larges, étant affectée par le sexe, l'âge, l'état de santé des locuteurs;
  - n'est pas clair lorsque on mesure la fréquence fondamentale;

Classification	Methodes	Caracteristiques importantes
<b>I. Traitement de la forme d'onde</b>	Methodes du traitement parallele (Gold-Rabiner)	Utilise des regles de la majorite pour l'extraction de la periode FF en detectant simplement les sommets de la forme d'onde
	Methodes de la reduction des donnees	Elimine les donnees redondantes de la forme d'onde en fonction de divers traitements numeriques et ne laisse que les impulsions FF
	Methodes de comptage des ZCR (NPZ)	Utilise des modeles iteratifs pour determiner le taux de passage par zero du signal
<b>II. Traitement par la correlation</b>	Methodes d'autocorrelation	On utilise la fonction d'autocorrelation de la forme d'onde. On utilise la limitation centrale pour le lissage et le pic de spectre et simplifie les calculs
	Methodes de la correlation modifiee	Utilisez la fonction d'autocorrelation du signal residuelle de l'analyse LPC
<b>III. Traitement du spectre</b>	Methodes cepstrales	Separe l'enveloppe spectrale de la structure fine, par IFFT du spectre logarithmique.
	Methodes d'histogramme des periodes	On utilise l'histogramme des composantes harmoniques du spectre. FF est defini comme diviseur commun des composantes harmoniques

## Méthode d'autocorrélation

$$R_n(k) = \sum_{m=0}^{N-1-k} [x(m+n)w(m)][x(n+m+k)w(m+k)]$$

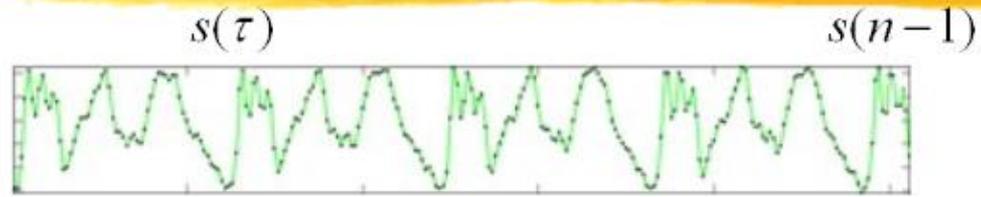
- La valeur de  $R_n(k)$  au moment  $n$ , au delay  $k$  est obtenu par le filtrage de la séquence  $x(n)..x(n-k)$  par un filtre à réponse impulsionnelle  $h_k(n)$



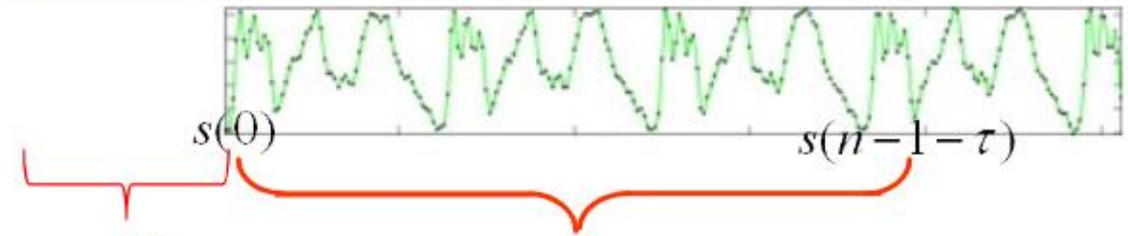
**Le traitement du signal a courte terme par l'autocorrélation**

0-index based,  
[s(0), s(1), ..., s(n-1)]

Original frame s(t):

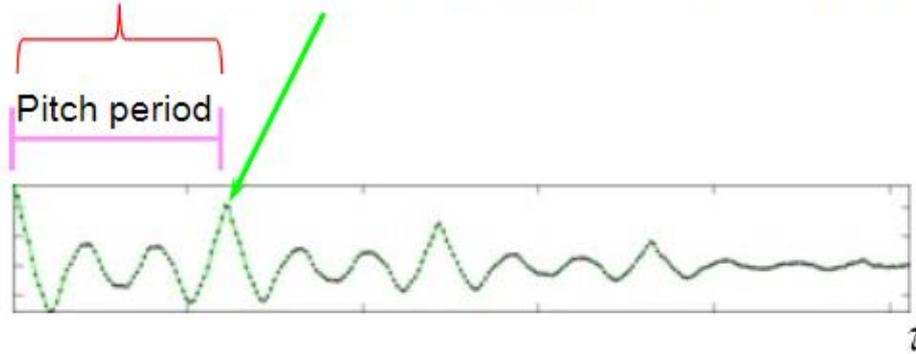


Shifted frame s(t-τ):



τ=30

acf(30) = inner product of the overlap part



Quiz candidate!

$$acf(\tau) = \sum_{t=\tau}^{n-1} s(t)s(t-\tau)$$

Quiz candidate!

To play safe, the frame size needs to cover at least two fundamental periods!

- *pour calculer en toute sécurité, la taille du cadre doit couvrir au moins deux périodes fondamentales*

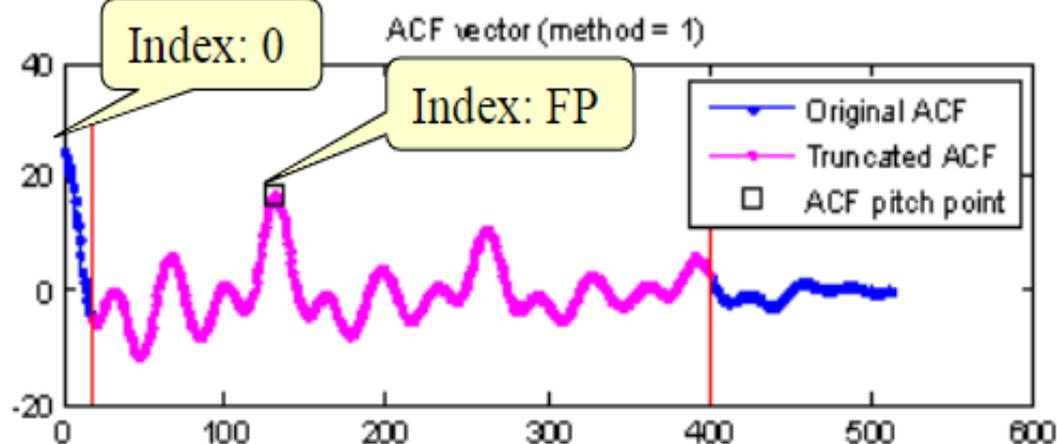
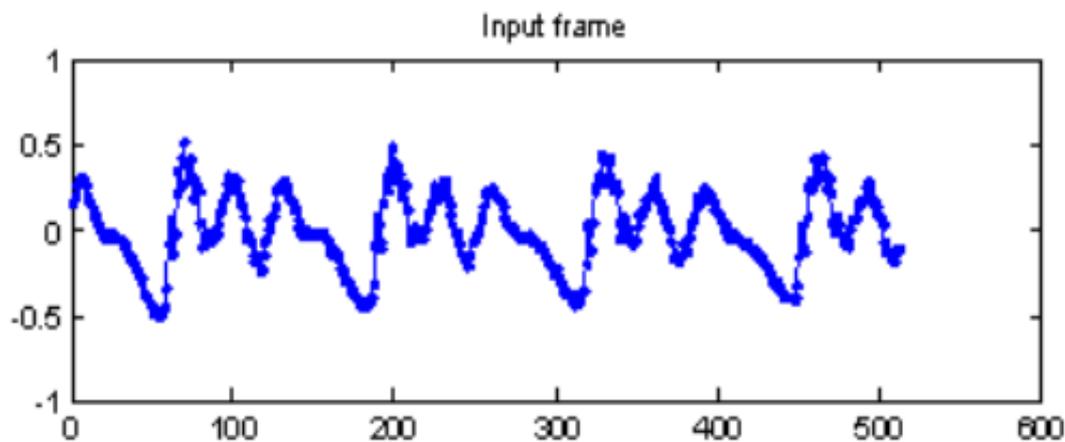
If the range of human's FF is  $[40, 1000]$ , then we have the interval for locating fundamental period (FP):

$$40 \leq \frac{fs}{FP} \leq 1000$$

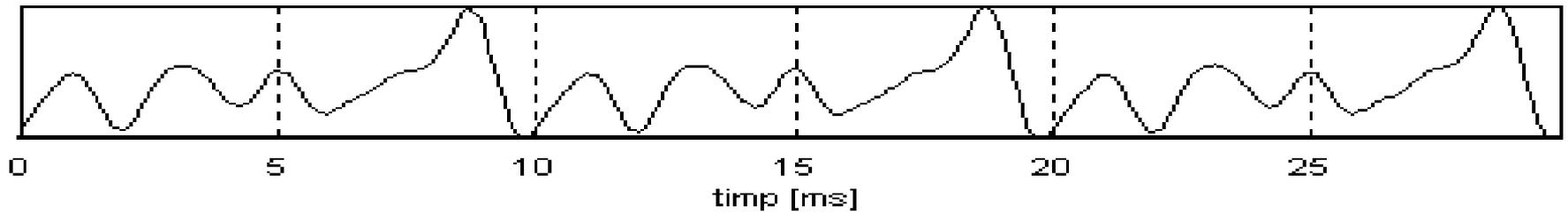
$$\Rightarrow \frac{fs}{1000} \leq FP \leq \frac{fs}{40}$$

Sample rate

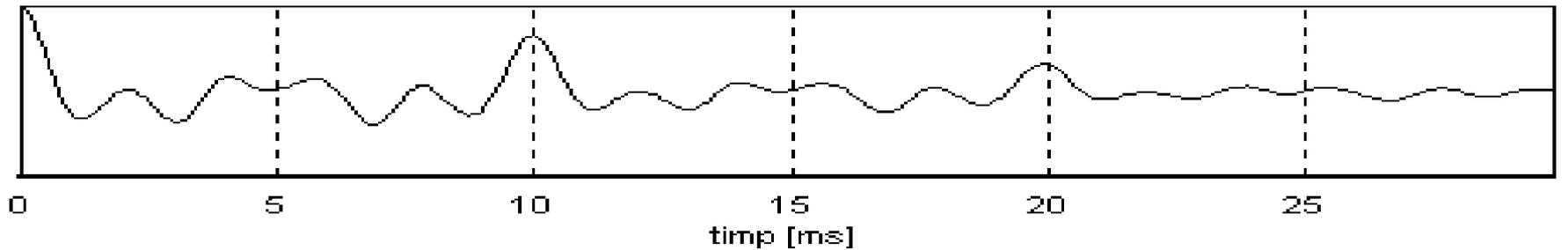
Quiz candidate!



Cadru semnal



Funcția de autocorelație

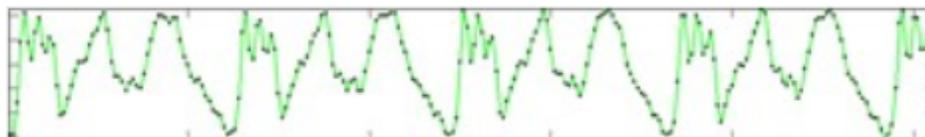


***Segment voisée et sa fonction d'autocorrélation  
associe***

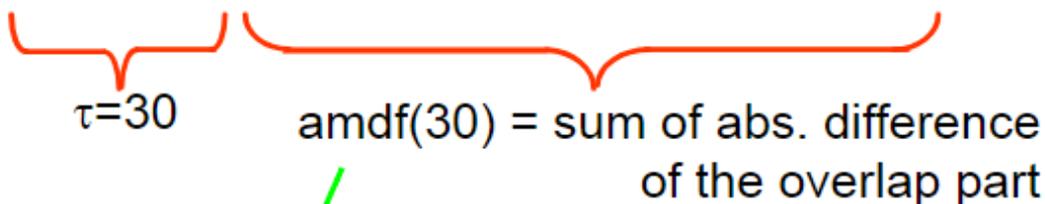
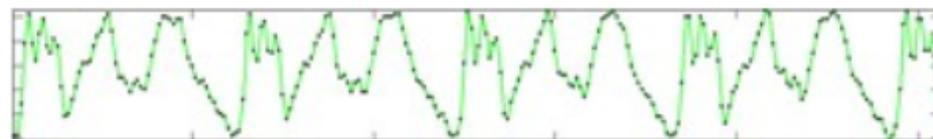
# La méthode de la fonction différence moyenne d'amplitude (FDMA)

$$D(n) = \frac{1}{N} \sum_0^{N-1} |s_k - s_{k-n}|, 0 \leq n \leq N-1.$$

Original frame  $s(i)$ :

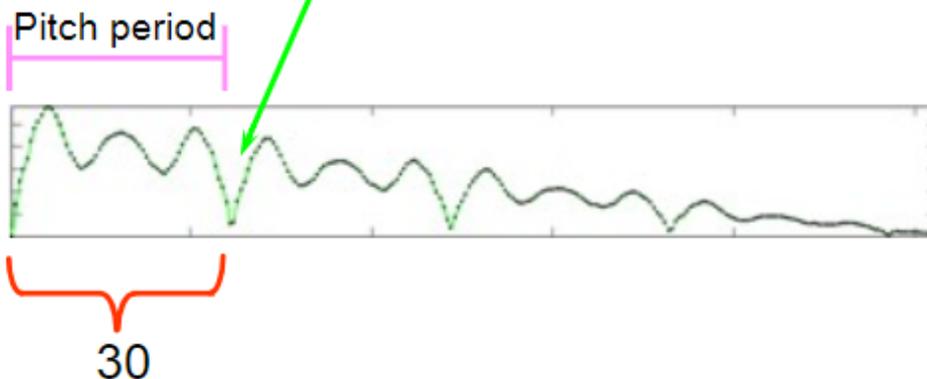


Shifted frame  $s(i-\tau)$ :



Quiz candidate!

$$amdf(\tau) = \sum_{t=\tau}^{n-1} |s(t) - s(t-\tau)|$$



La fonction différence Moyenne d'amplitude (AMDF)

# Formulas

ACF:

$$acf(\tau) = \sum_{t=\tau}^{n-1} s(t) s(t-\tau)$$

AMDF:

$$amdf(\tau) = \sum_{t=\tau}^{n-1} |s(t) - s(t-\tau)|$$

## Two major advantages of AMDF over ACF

AMDF requires less computing power

AMDF is less likely to have the risk of overflow

Les avantages de la FDMA sur la FAC:

- FDMA nécessite moins des calculs
- FDMA est moins susceptible de présenter un risque de débordement

# La fonction différence Moyenne d'amplitude (FDMA)

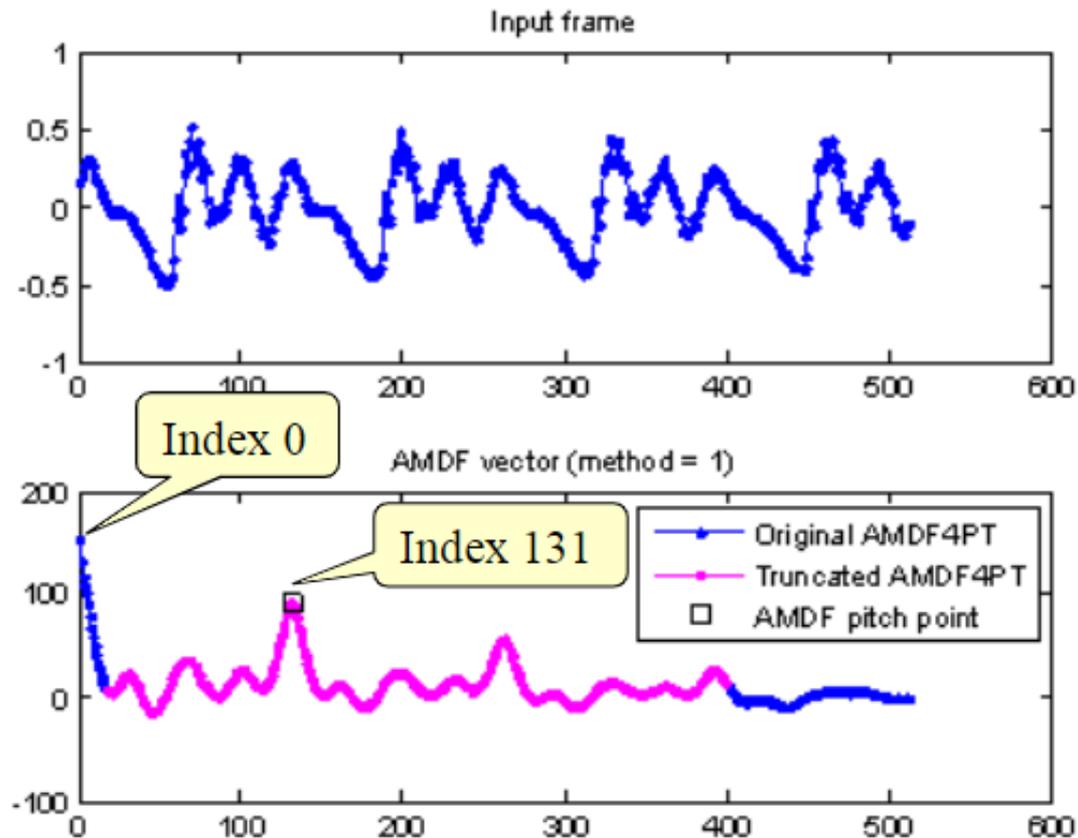
Sample rate = 16kHz

Frame size = 512  
(starting from point  
9000)

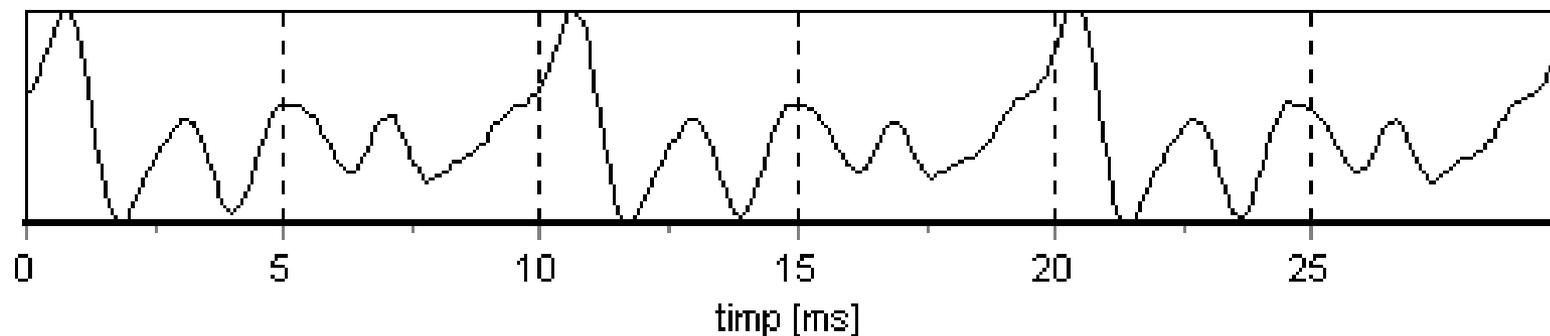
Fundamental frequency

Pitch point occurs at  
index 131, which is  
determined correctly

$$FF = 16000/131 = 123.077 \text{ Hz}$$

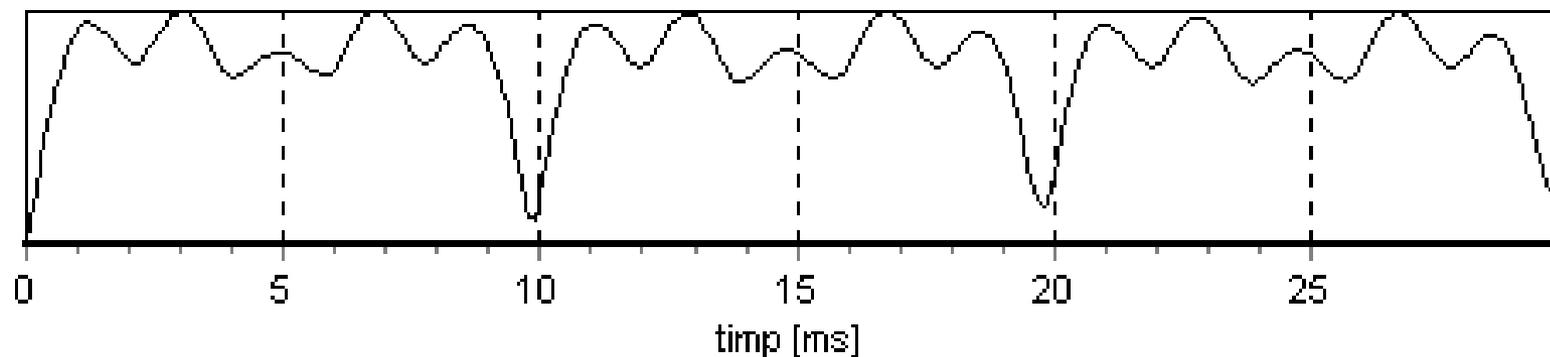


Cadru semnal



Funcția diferentă de amplitudine medie

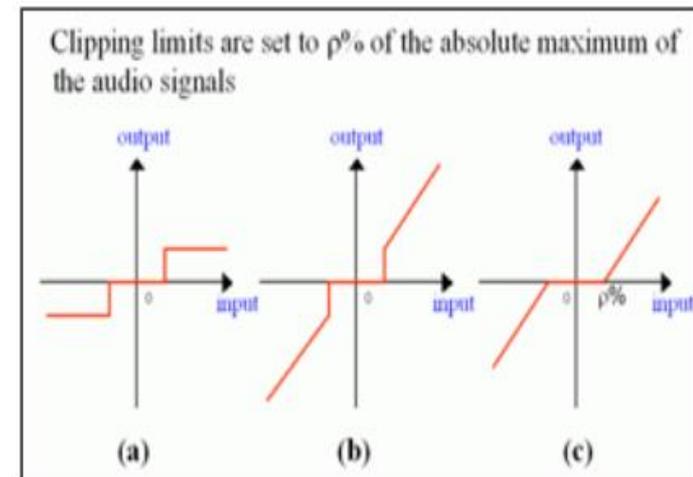
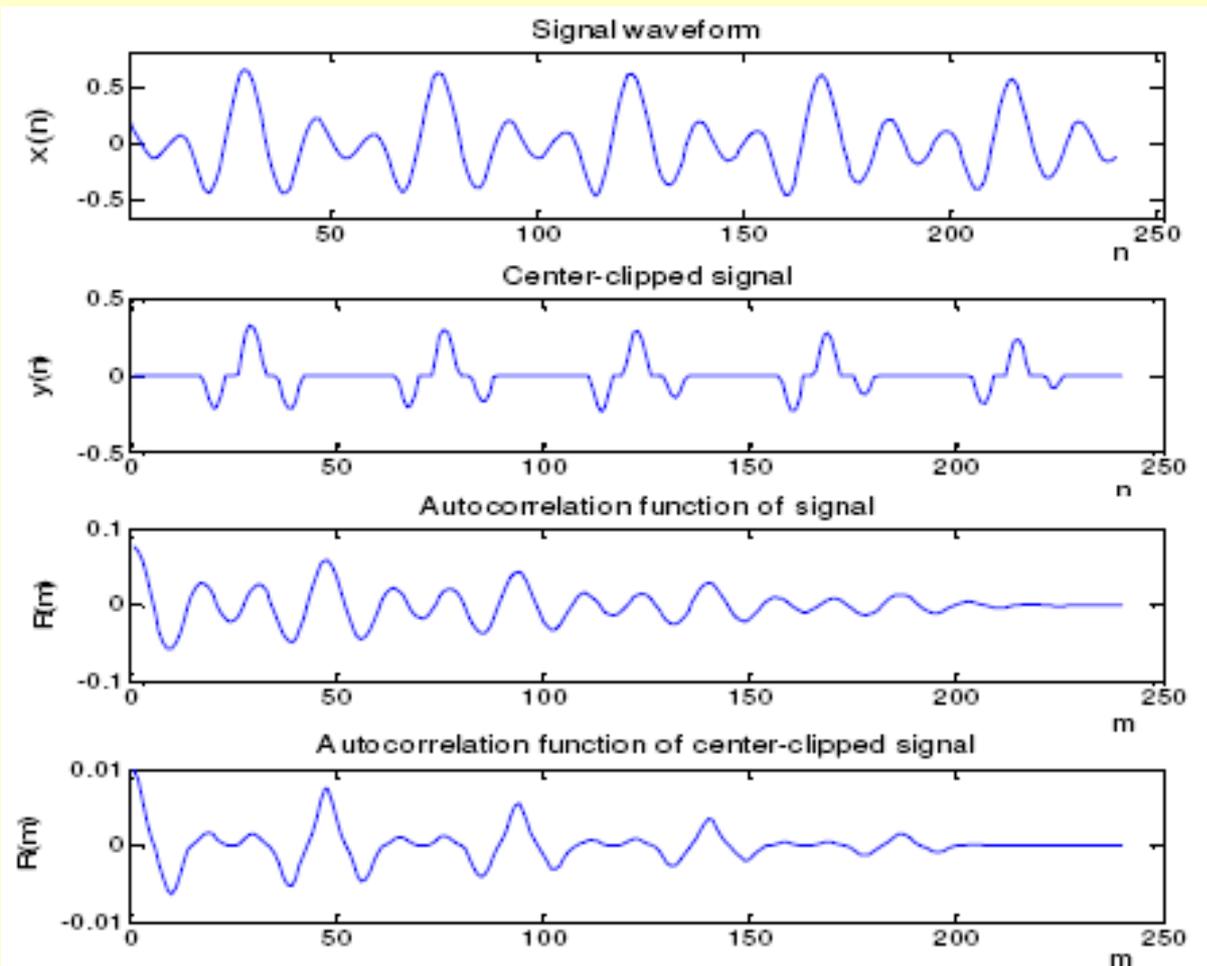
$T_0$ : 9,89 ms  
 $F_0$ : 101,15 Hz

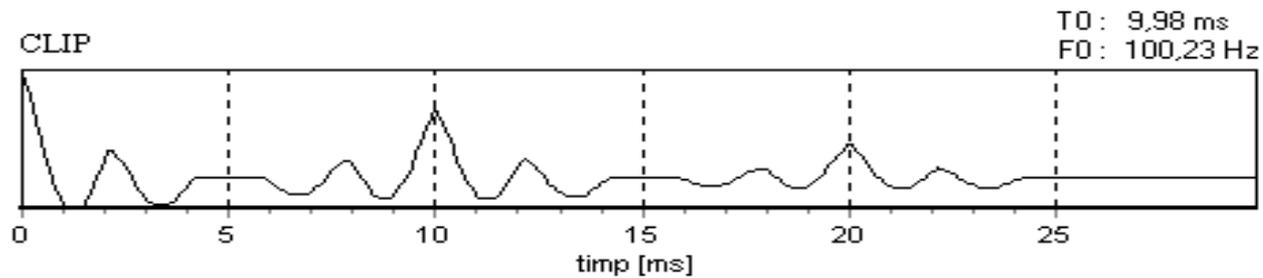
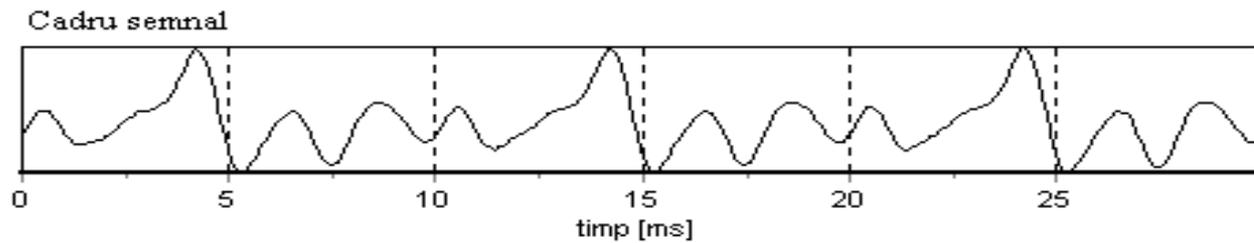


**La fonction différence Moyenne  
d'amplitude (*FDMA*)**

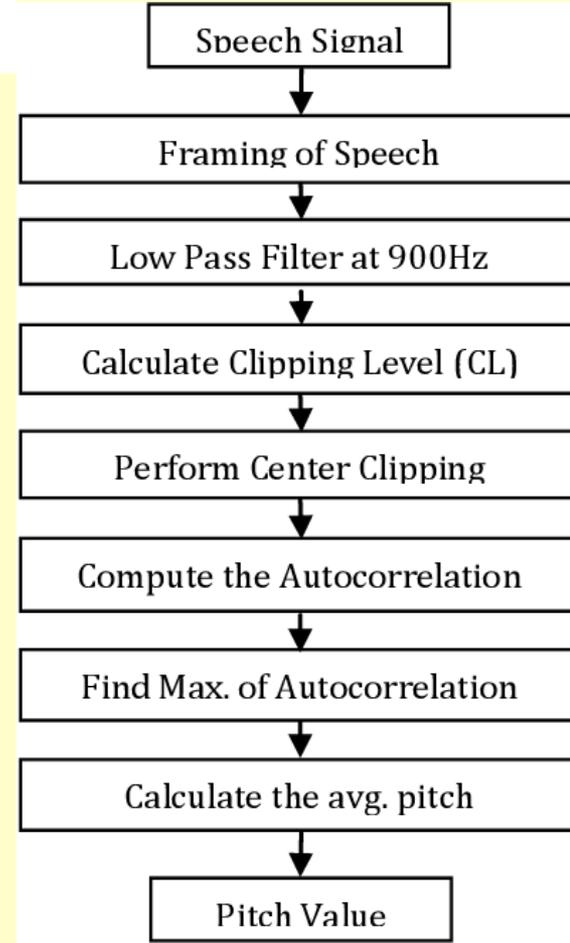
# La méthode a limitation centrale (CLIP: Center Clipping Pitch Detector)

-On utilise un seuil CL et conserve seulement les éléments du signal de la parole dont la valeur absolue dépasse le seuil  $|CL|$ . Pour les valeurs restées: on ajoute la valeur du CL dans le positif, et  $-CL$  de valeurs négatives ou est attribué le maximum positif ou respectivement négatif (limite infinie). Suivant calculer la fonction d'autocorrélation

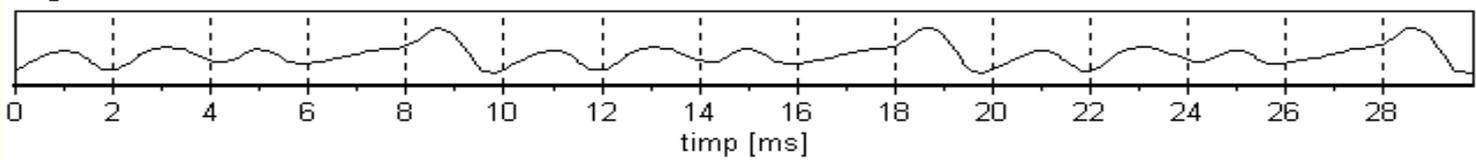




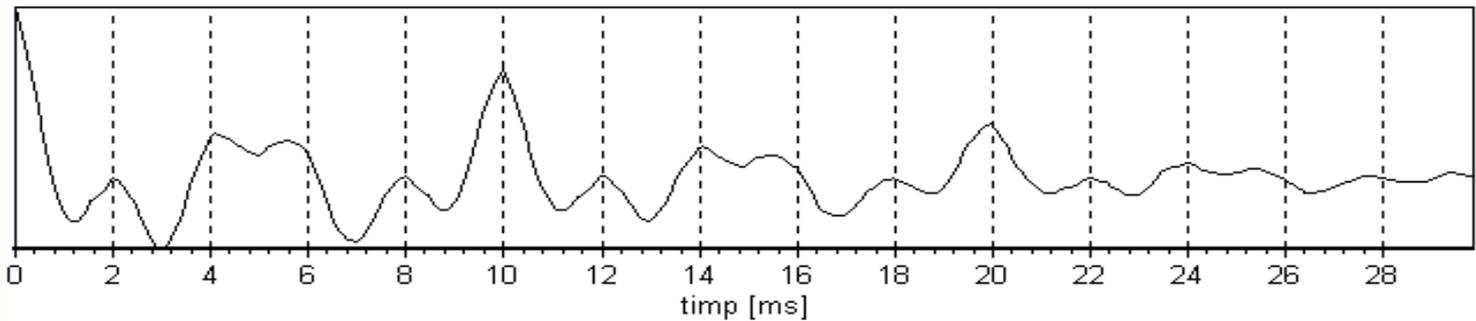
*Determination de la FF en utilisant la méthode CLIP*



Segment de semnal

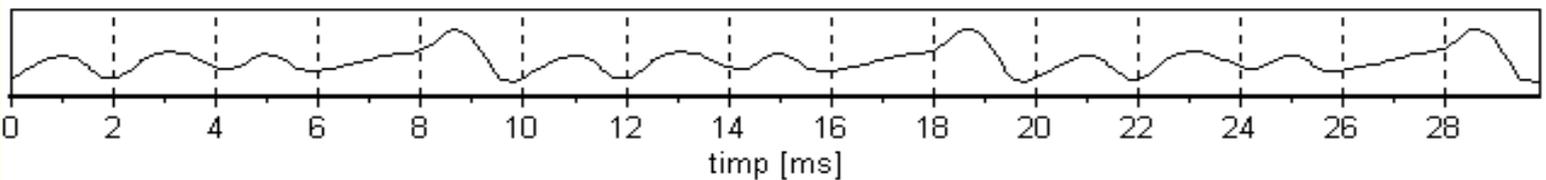


CLIP

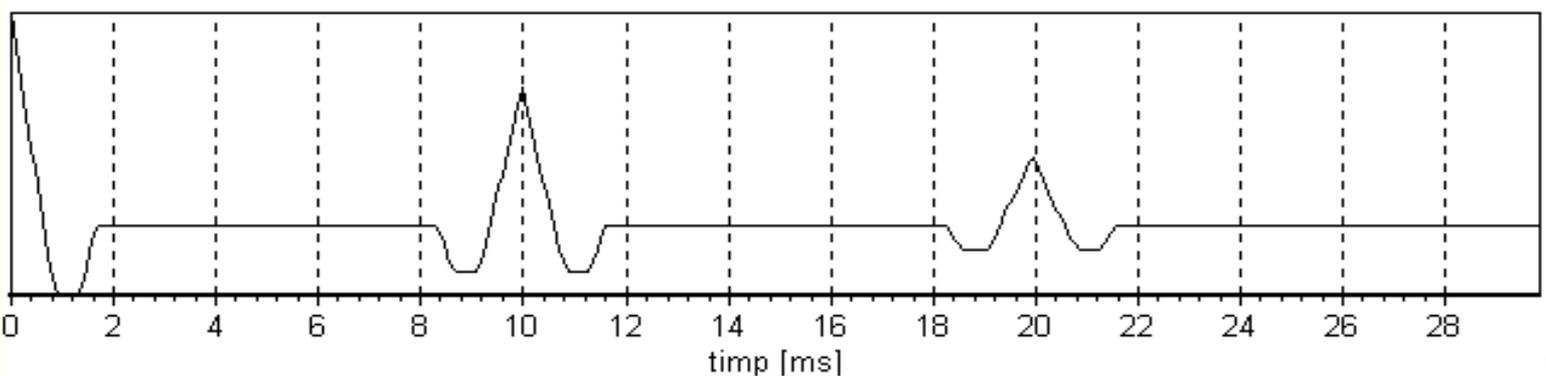


*Determination de la FF en utilisant la méthode CLIP, seuil 20%.*

Segment de semnal

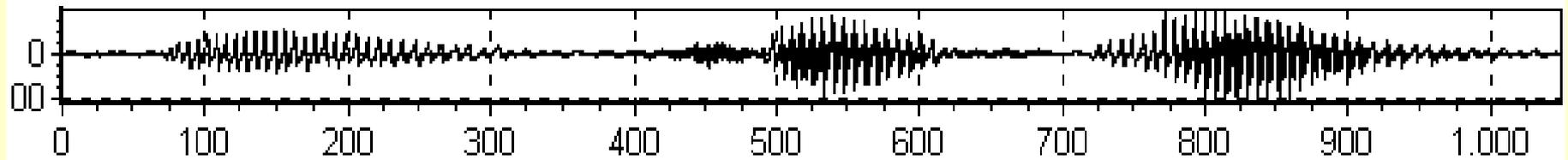


CLIP

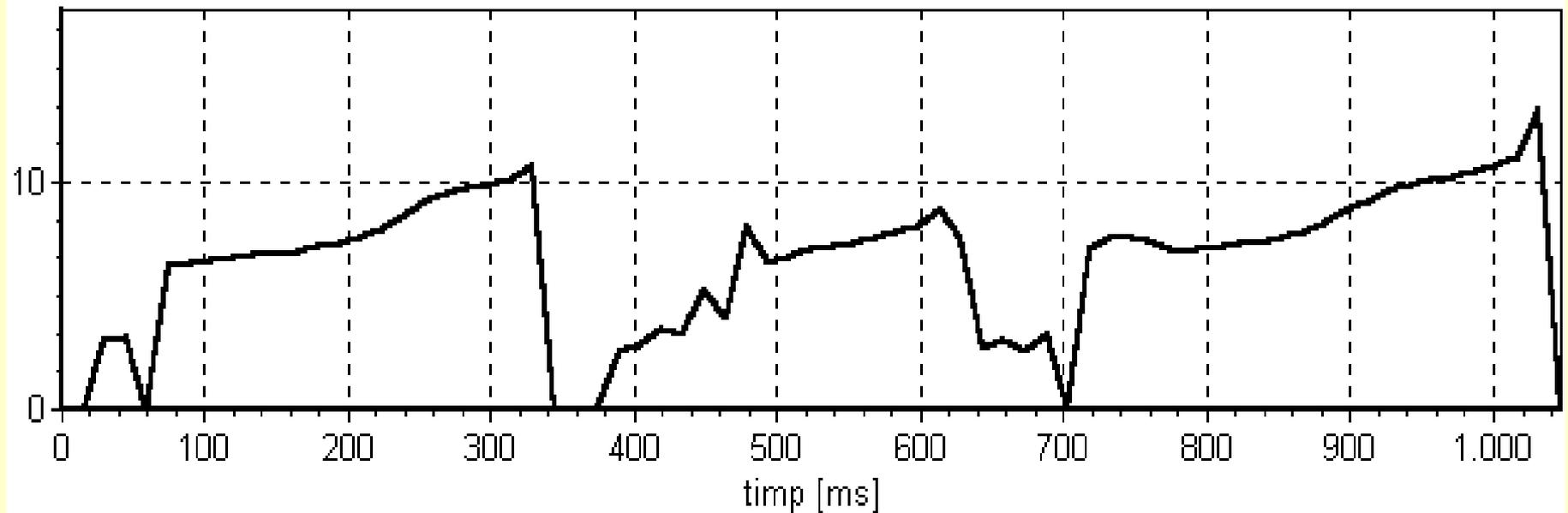


*Determination de la FF en utilisant la methode CLIP, seuil 60%.*

Signal vocal

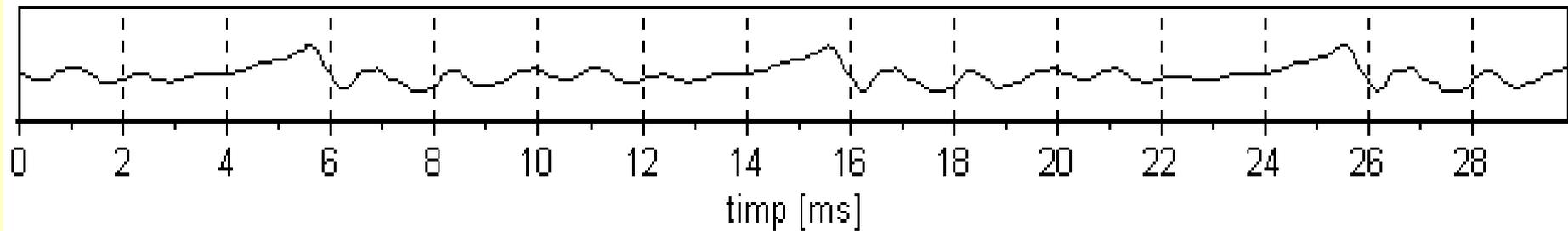


Frecventa fundamentala



*Le contour de la FF en utilisant la metode CLIP, seuil 30%, pour une phrase*

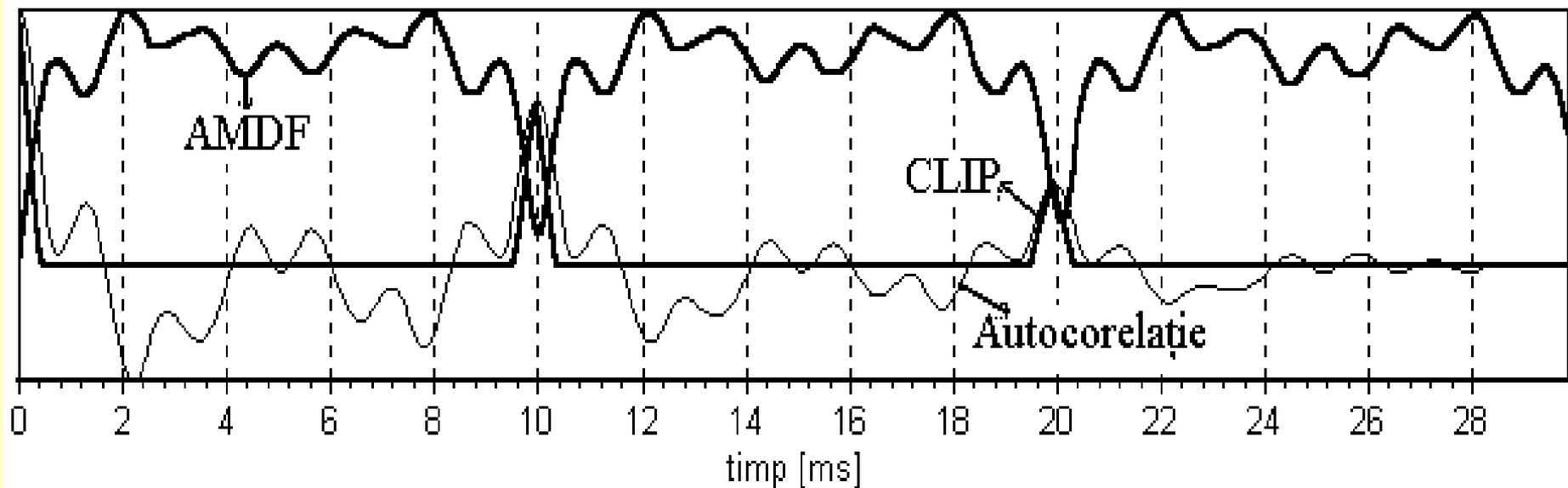
Segment de signal



Autocorrelation, AMDF, CLIP

T0 : 9.98 ms

F0 : 100.23 Hz

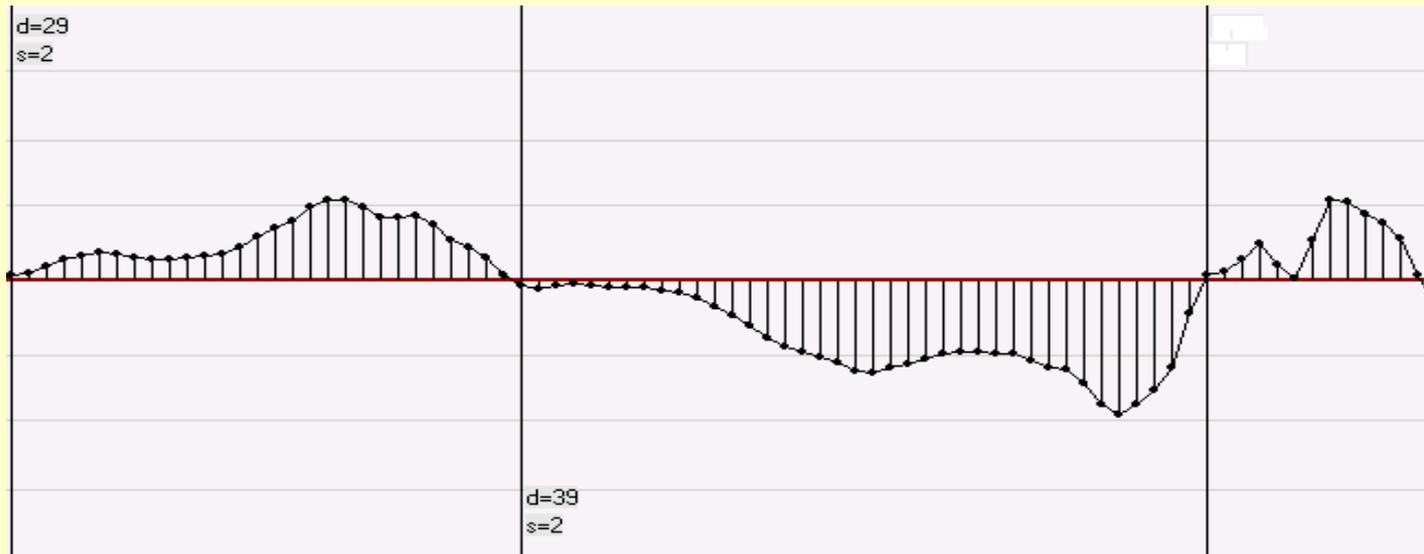


*Convergence des méthodes de détermination de la fréquence fondamentale fondée sur l'autocorrelation*

# 7. Methode de codage TESPAP

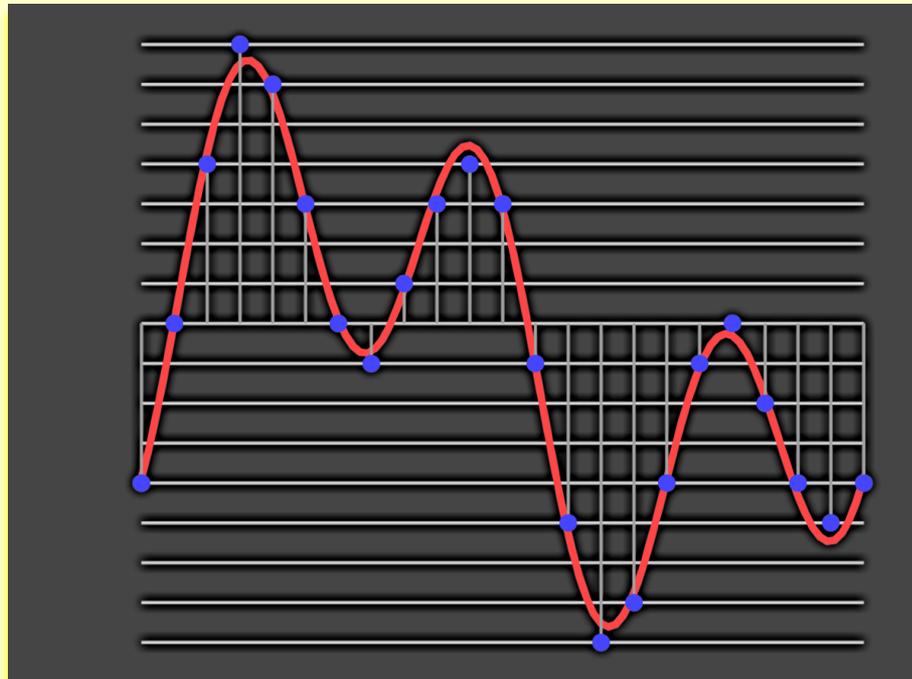
## TESPAR – Time Encoded Signal Processing And Recognition

- C. E. Shannon „A Mathematical Theory Of Communication”, 1948  
„Communication In The Presence Of Noise” 1949
- *Si une fonction  $f(t)$  ne contient pas des fréquences supérieures à  $B(\text{Hz})$ , alors elle est complètement déterminée par ses valeurs à des intervalles de  $\frac{1}{2B}$  sec. les uns des autres*

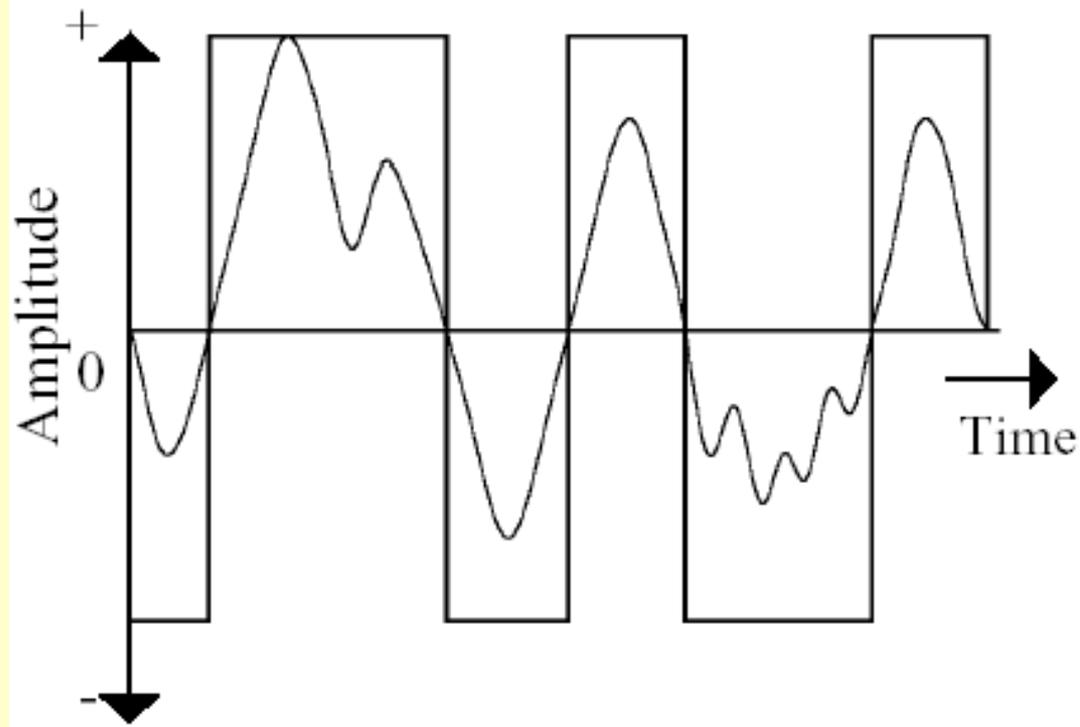


L'échantillonnage régulier des signaux

1. - L'échantillonnage des données ont été ensuite utilisées comme base pour les transformations diverses (généralement linéaire) que: T. Fourier, T. Wavellet et le codeur Walsh pour extraire les paramètres clés dans l'analyse du signal.
2. Les méthodes courantes de codage **demande trois exigences:**
  - *échantillonnage régulier*
  - *l'utilisation des descripteurs d'amplitude*
  - *un domain d'approximation pour l'amplitude*



- C. E. Shannon
- Licklider și Pollack
- Preston C. Hammer
- Bond și Cahn



Infinite clipping

- Licklider et Pollack - effets du manque d'informations d'amplitude dans l'intelligibilité de la parole
- l'exclusion des informations d'amplitude ont conservés l'intelligibilité pendant plus de 97% des cas
- Ce qui montre qu'une partie essentielle de l'information dans la transmission de la voix (intelligibilité) est contenue dans les passages par zéro de l'onde sonore\*

# L'analyse de signaux basée sur des zéros

- La théorie de l'analyse du signal sur la base de zéros commencé par Bond et Cahn en 1958
- Introduction de zéros complexes, atteint deux conclusions importantes:
  - Les groupes de zéros (réels et complexes) sont à la limite Niquist (échantillonnage)
- Les fonctions générées par les sources naturelles de signal, a bande passante limitée, contiennent des zéros complexes qui ne peuvent pas tous être physiquement détecté
- Il n'est pas facile de déterminer *l'ensemble des zéros complexes et réels* qui caractérisent une forme d'onde particulière, la seule méthode connue qui garantit de trouver tous les zéros complexes et réels d'un polynôme de degré  $2TB$  implique la factorisation – n'est pas efficace en temps réel

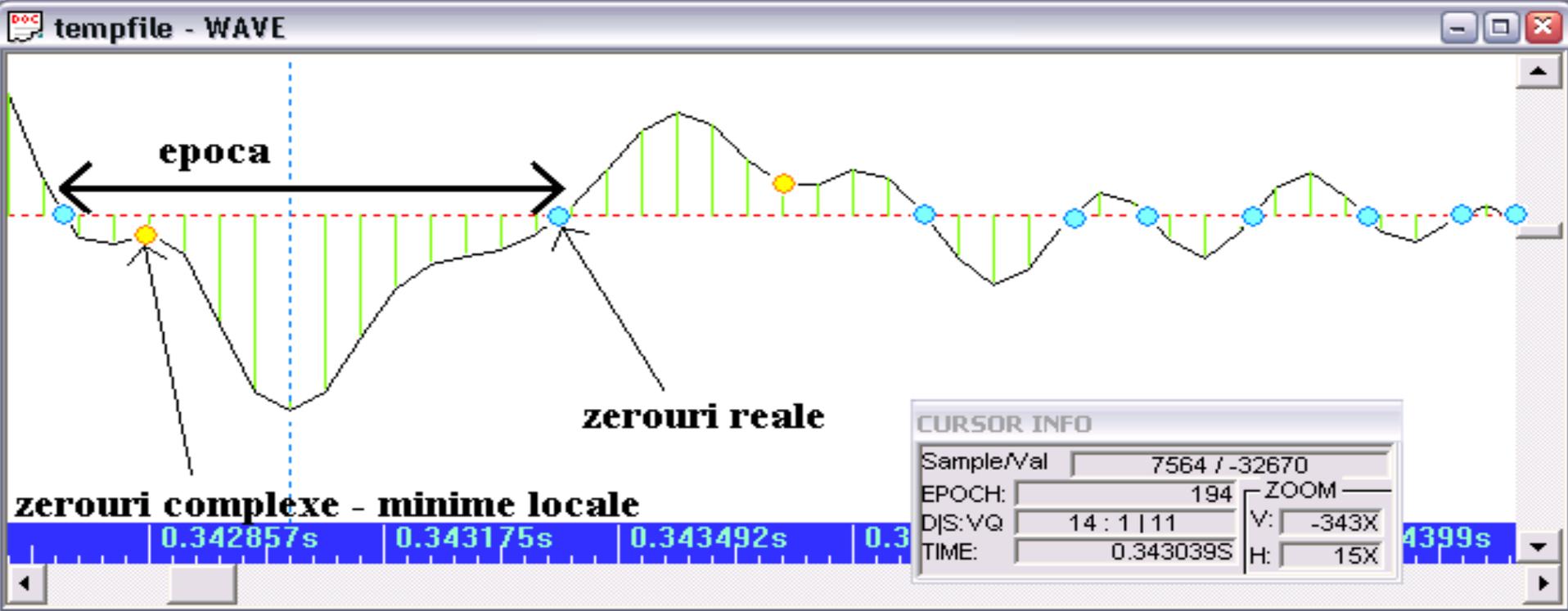
- Nous pouvons considérer cette approche comme une transformée  
**de domain temporel >> dans le domain zéros**

- Alors, que les zéros sont des zéros réels, identique des zéros et des zéros complexes conjugués qui sont associée à des discontinuités: les points minimum, maximum ou d'inflexion qui se produisent dans la forme d'onde entre deux passages par zéro successifs

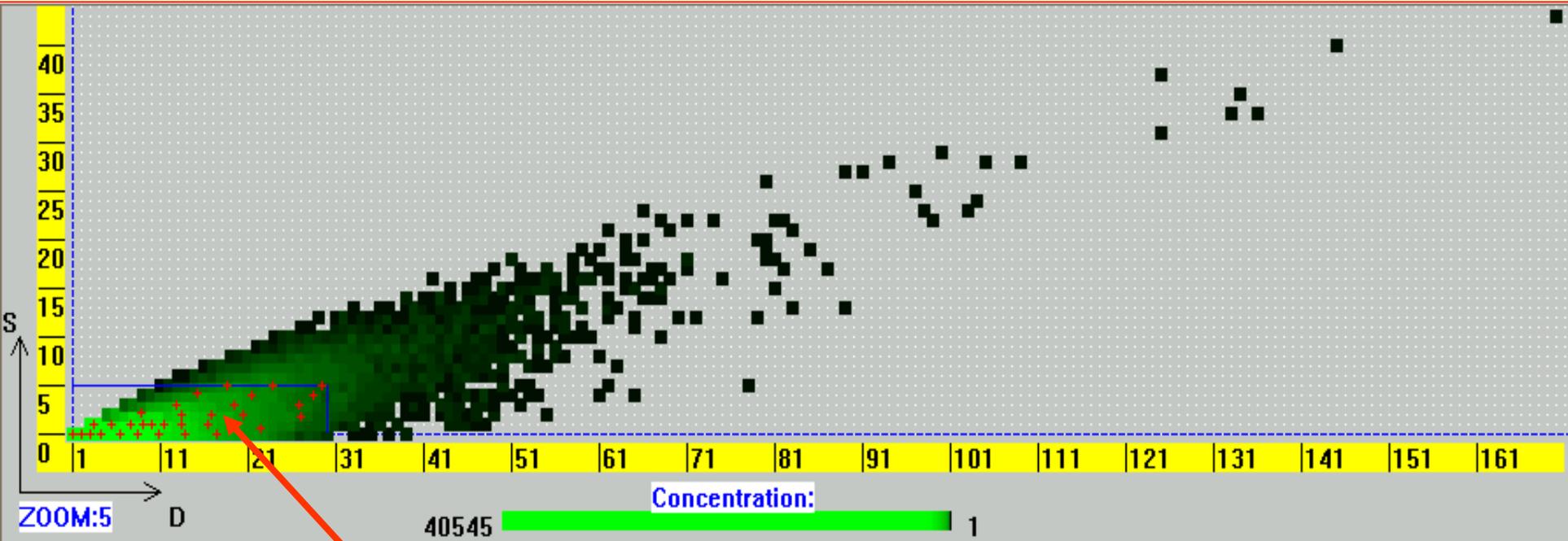
- on fait l'analyse de la forme d'onde entre deux passages par zéro successifs pour obtenir un nombre suffisant de sous-ensembles des zéros pour caractériser la forme d'onde étudiée.

# Methode TESPAP

- une approximation très simple, basé sur la théorie de la forme d'onde est le codage TESPAP
- L'époque (est situé entre deux passages du signal à travers 0)
  - la durée* entre deux passages par zéro successifs- Duration (durée)
  - la forme* entre les deux zéros réels successifs - Shape (forme)
- A chaque époque est associé un symbole sur la base des deux descripteurs S et D
- - on obtienne une tres forte compression  
Tous les symboles forment l'alphabet TESPAP



# La génération de l'alphabet



Distribution des époques dans le plan de SxD pour un fichier contenant un discours qui a la dure environ 2 minutes, échantillonné à une fréquence de 8 kHz

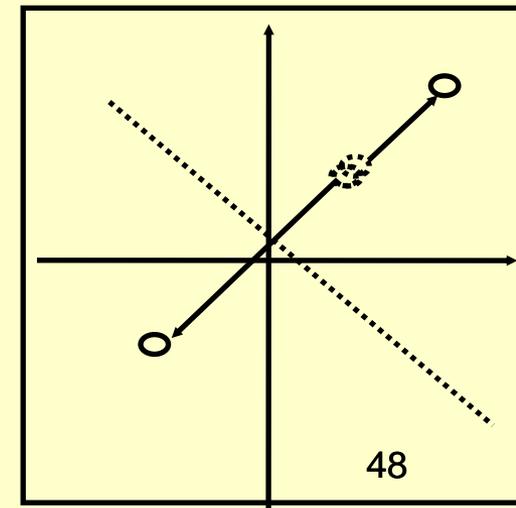
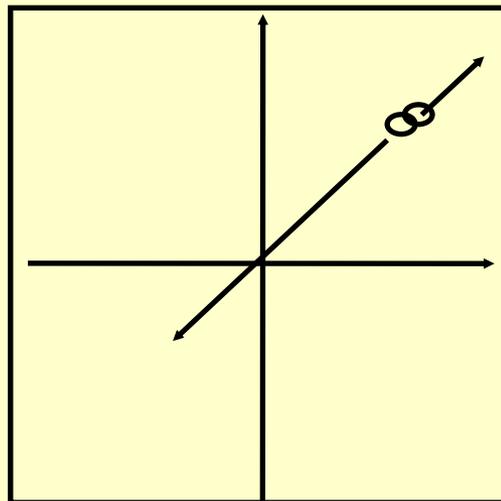
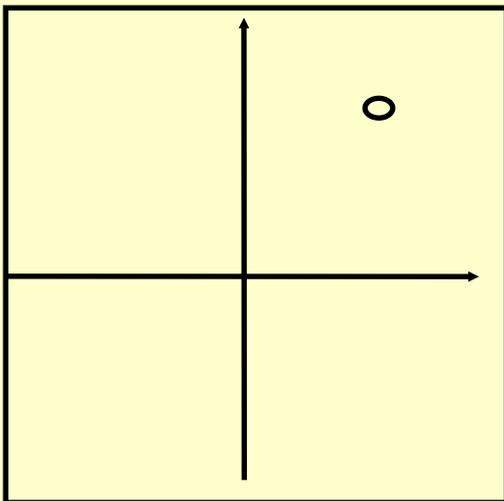
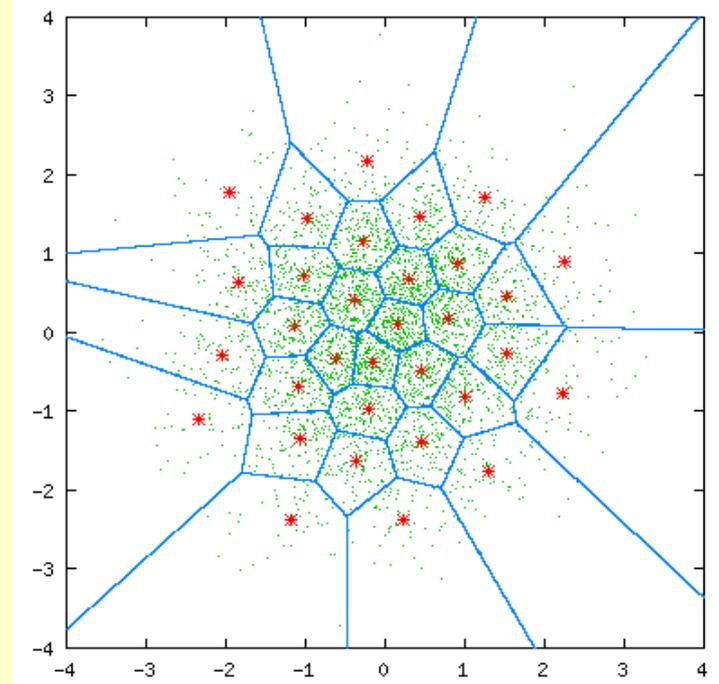
8 kHz	D=30
11 kHz	D=40
22 kHz	D=70

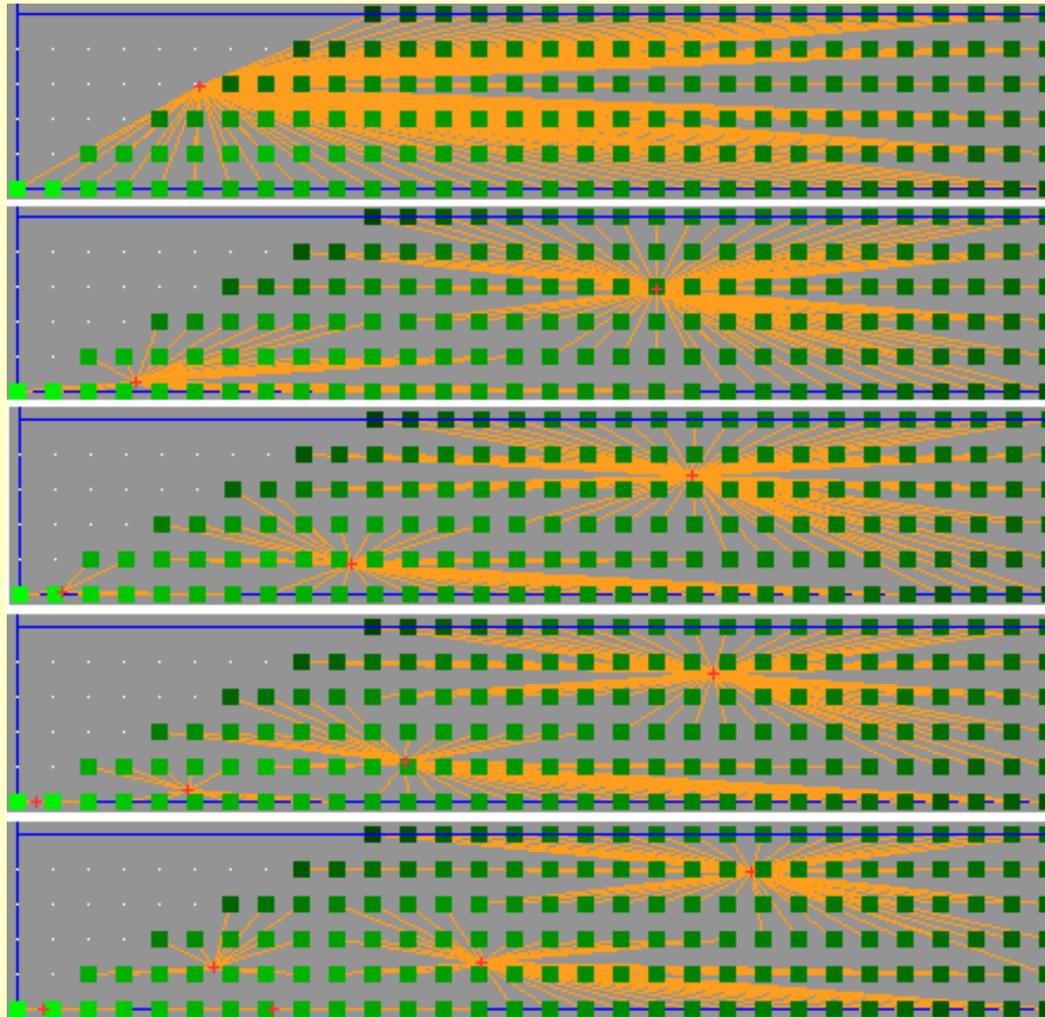
**concentration:  
>96%**

8 kHz	S=5
11 kHz	S=5
22 kHz	S=5

# Alphabet TESPAP – LBG-VQ

- Est déterminé en fonction des propriétés statistiques de la parole humaine
- Permet une plus grande compression
- Élimine la partie ennuyeuse de la variabilité de la voix humaine
- L'algorithme LBG est utilisé et vise à minimiser la distorsion globale





## 5 pas de l'algorithme LBG

40545  1

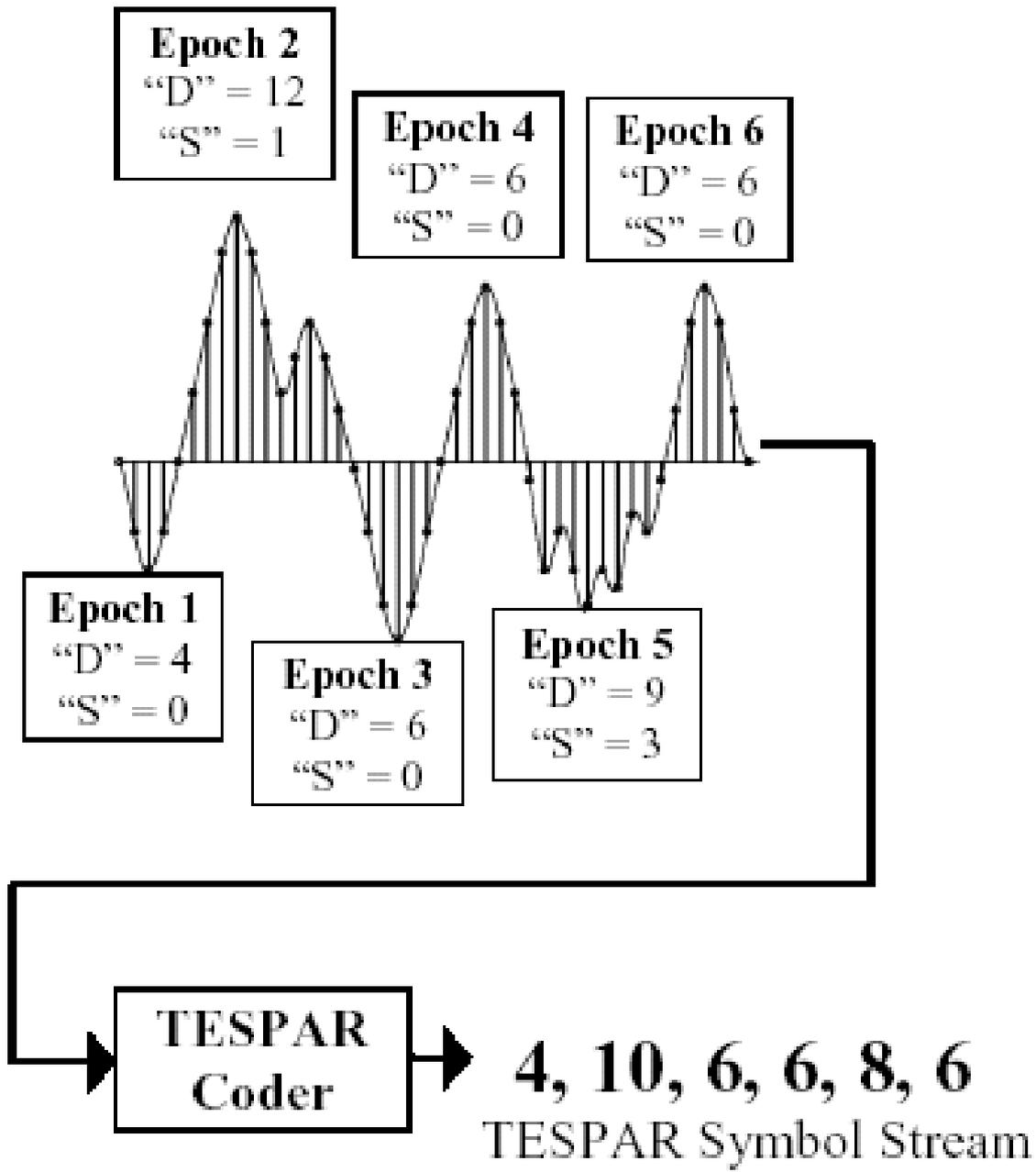
S/D	0	1	2	3	4	5
1	0	-1	-1	-1	-1	-1
2	6	-1	-1	-1	-1	-1
3	14	10	-1	-1	-1	-1
4	4	10	-1	-1	-1	-1
5	30	10	10	-1	-1	-1
6	11	10	25	-1	-1	-1
7	11	9	25	25	-1	-1
8	17	9	25	25	-1	-1
9	1	5	25	25	21	-1
10	1	5	12	21	21	-1
...	.....	.....	.....	.....	.....	...
26	8	20	20	20	3	3
27	8	20	20	20	3	3
28	8	20	20	20	3	3
29	20	20	20	20	3	3
30	20	20	20	20	3	3

Exemple d'alphabet pour la voix échantillonnée à 8 kHz

S/D	0	1	2	3	4	5
1	2	-1	-1	-1	-1	-1
2	17	-1	-1	-1	-1	-1
3	17	7	-1	-1	-1	-1
4	8	7	-1	-1	-1	-1
5	8	7	21	-1	-1	-1
6	25	7	21	-1	-1	-1
7	25	7	21	21	-1	-1
8	25	7	21	21	-1	-1
9	25	7	21	21	9	-1
10	12	11	21	21	9	-1
11	12	11	21	21	9	-1
12	12	11	21	21	9	-1
13	12	11	21	9	9	27
14	12	11	21	9	9	27
15	12	11	21	9	22	27
16	18	11	5	9	22	27
17	18	11	5	9	22	27
18	18	11	5	9	22	27
19	18	23	5	9	22	27
20	18	23	5	9	22	27
21	18	23	5	9	22	27
22	18	23	5	9	22	27
23	18	23	5	9	22	27
24	3	23	29	9	22	27
25	3	23	29	9	22	27
26	3	23	29	9	22	27
27	3	15	29	9	22	27
28	3	15	29	24	31	27
29	3	15	14	24	31	27
30	3	15	14	24	31	27

S/D	0	1	2	3	4	5
31	3	15	14	24	31	4
32	3	15	14	24	31	4
33	3	15	14	24	31	4
34	3	15	30	24	31	4
35	3	15	30	24	13	4
36	6	15	30	24	13	4
37	6	15	30	24	13	4
38	6	15	30	24	13	4
39	6	6	30	24	13	4
40	6	6	30	13	13	4
41	6	6	30	13	13	20
42	6	6	30	13	13	20
43	6	6	30	13	13	20
44	6	6	30	13	13	20
45	6	6	16	13	13	20
46	6	16	16	13	13	20
47	6	16	16	13	10	20
48	6	16	16	13	10	20
49	6	16	16	10	10	20
50	6	16	16	10	10	20
51	6	16	16	10	10	20
52	6	16	16	10	10	28
53	6	16	16	10	10	28
54	6	16	16	10	10	28
55	6	16	16	10	10	28
56	6	16	19	19	10	28
57	6	16	19	19	10	28
58	6	16	19	19	10	28
59	6	16	19	19	10	28
60	6	16	19	19	10	28
61	6	16	19	19	10	28
62	16	16	19	19	10	28
63	16	16	19	19	10	28
64	16	19	19	19	26	28
65	16	19	19	19	26	28
66	16	19	19	19	26	28
67	16	19	19	19	26	28
68	16	19	19	19	26	28
69	16	19	19	19	26	28
70	16	19	19	19	26	28

Tabel 16: Alfabetul implicit pentru frecvența de 22kHz

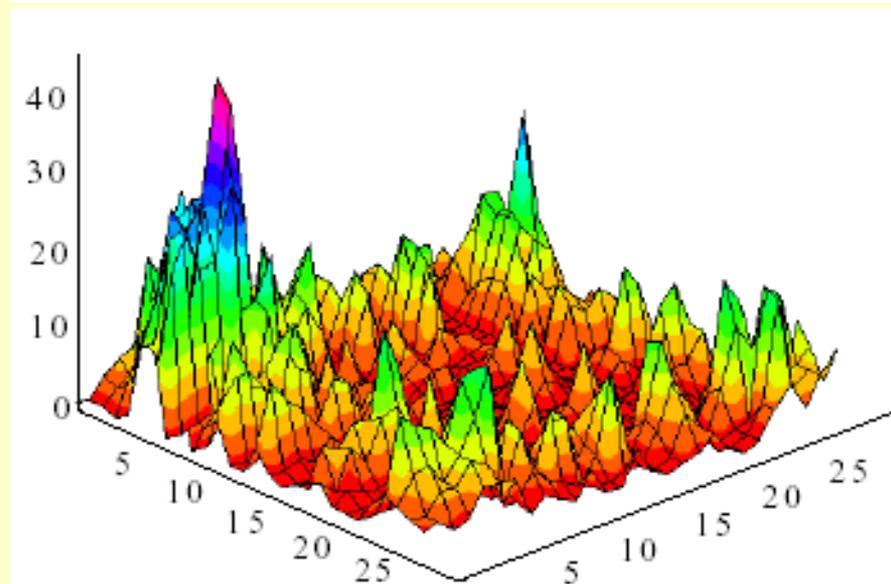
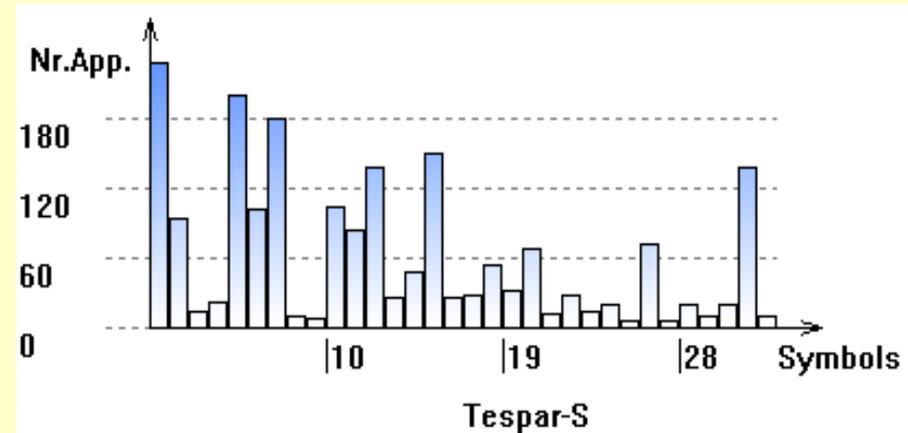


# Les Matrices S et A

- Du débit de symboles obtenus a la sortie du codeur TESPAP on peut créer des tableaux:

Matrice S - (1 dimensionnelle) ne retenir que le nombre d'occurrences de symboles

Matrice A- (2 dimensions) contiennent des informations temporelles, impliquant un autre paramètre  $L = \text{LAG}$

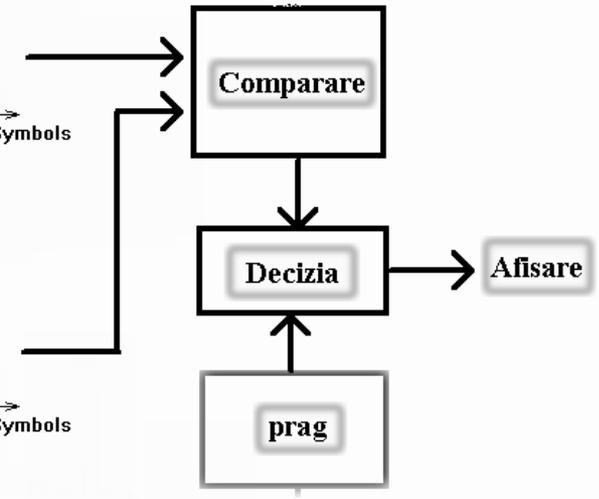
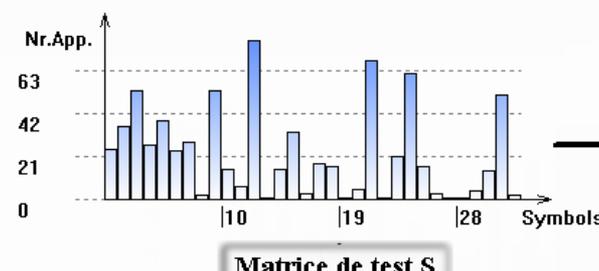
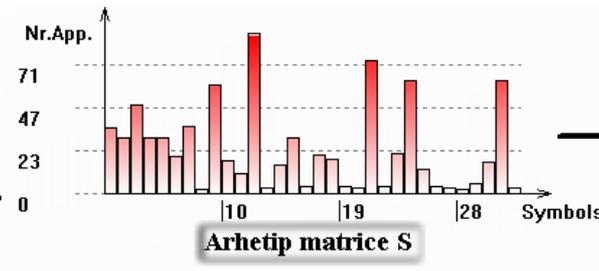
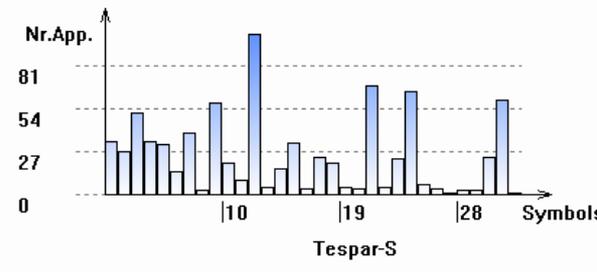
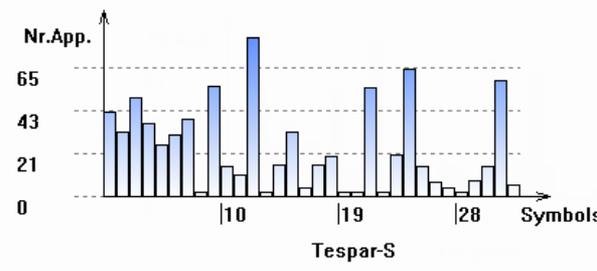
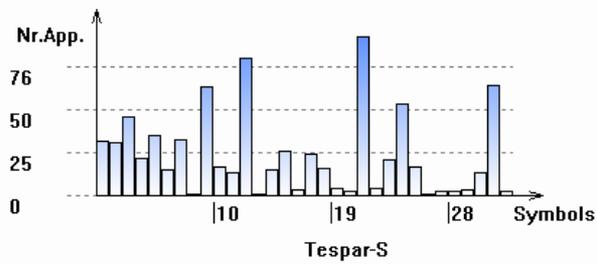


- Il existe deux méthodes principales de classification en utilisant cette méthode:

- *classification en utilisant des archétypes*

- *classification en utilisant les réseaux de neurones*

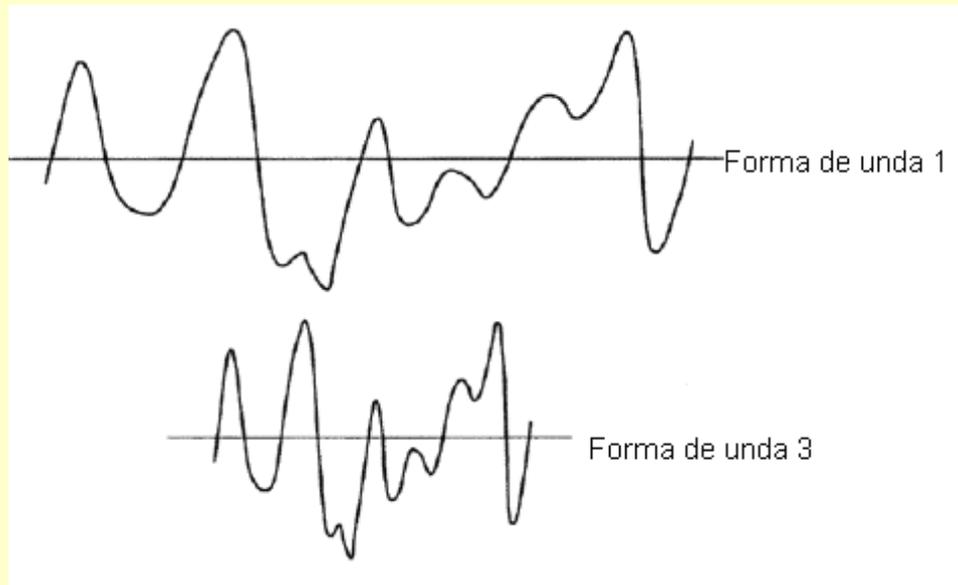
Le principal avantage des variables de traitement du signal dans le temps en utilisant TESPAP - aux méthodes classiques qui utilisent des descripteurs dans le domaine fréquentiel est que *les tableaux TESPAP sont de taille fixés*



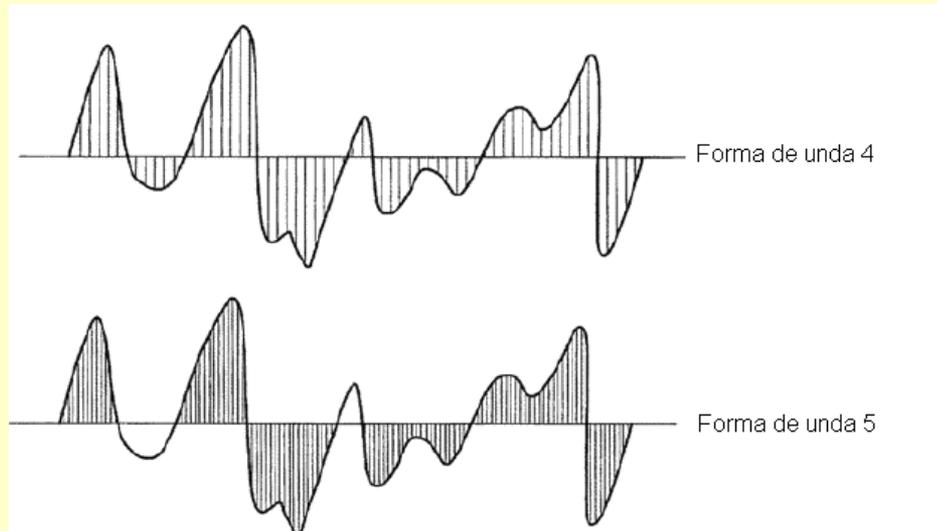
génération d'archétype et son utilisation pour la classification

- Archétype = nombre moyen de matrices d'entrée, élimine la variabilité
- Vous pouvez suivre les changements des valeurs des données sur chaque dimension et de leur donner plus d'importance que ceux qui varient

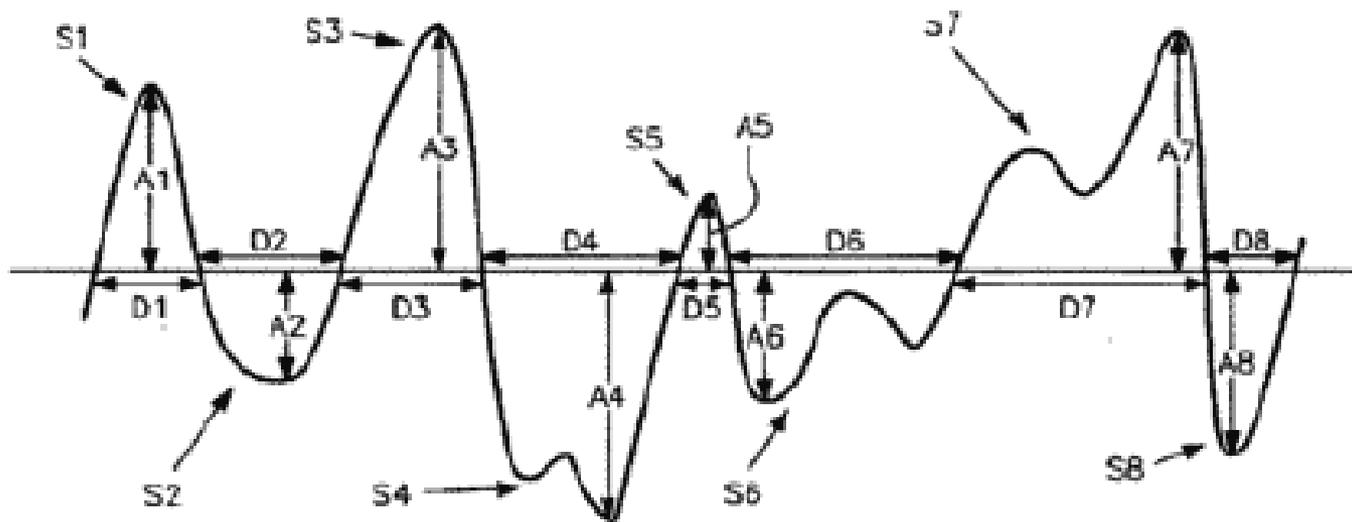
# Méthode TESPAP DZ



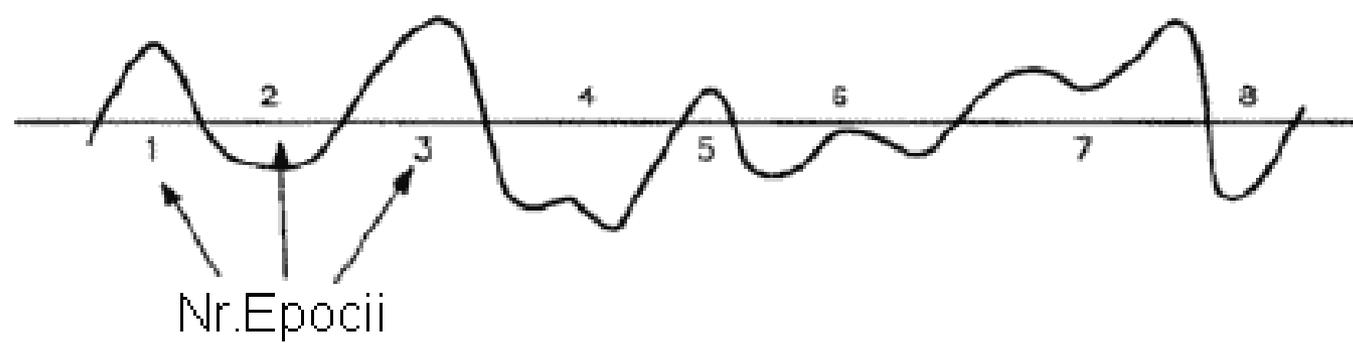
Forme d'onde originale et échantillonné à une fréquence double



Formes d'ondes échantillonnées à différentes fréquences

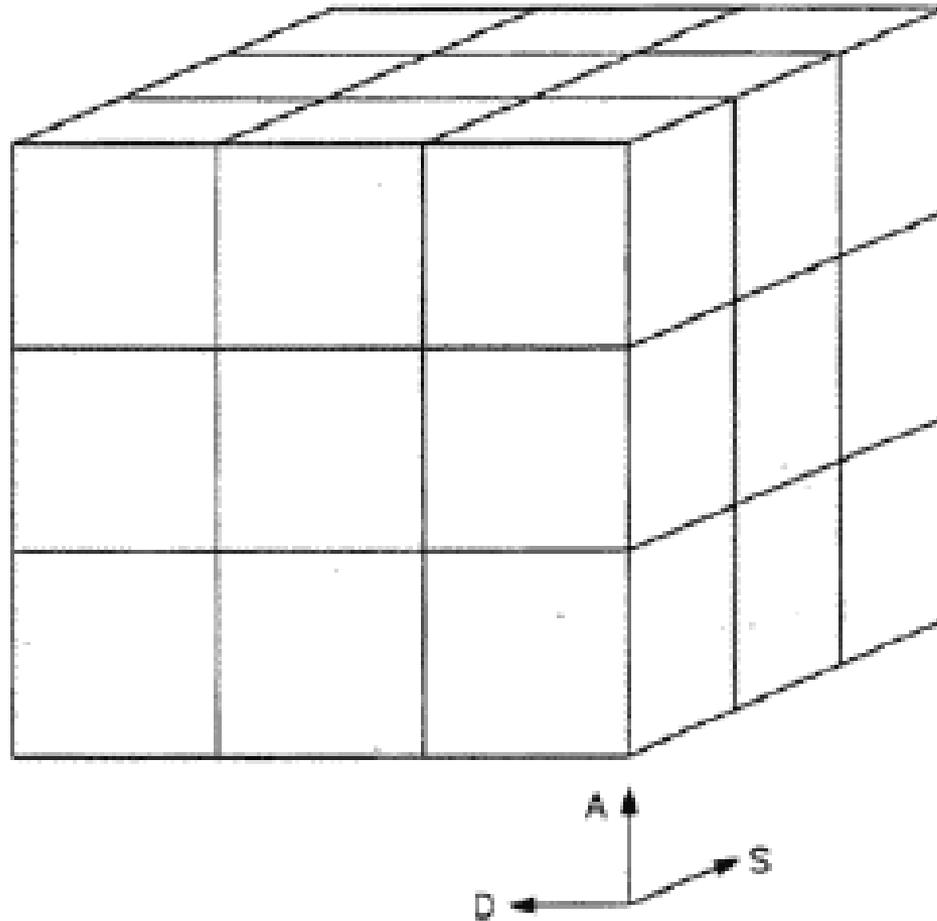


Forma de unda 1



Forma de unda 2

- Des comparaisons sont faites entre les paires d'époque, ou sont comparés les caractéristiques individuelles D, S et A de chaque paire d'époques

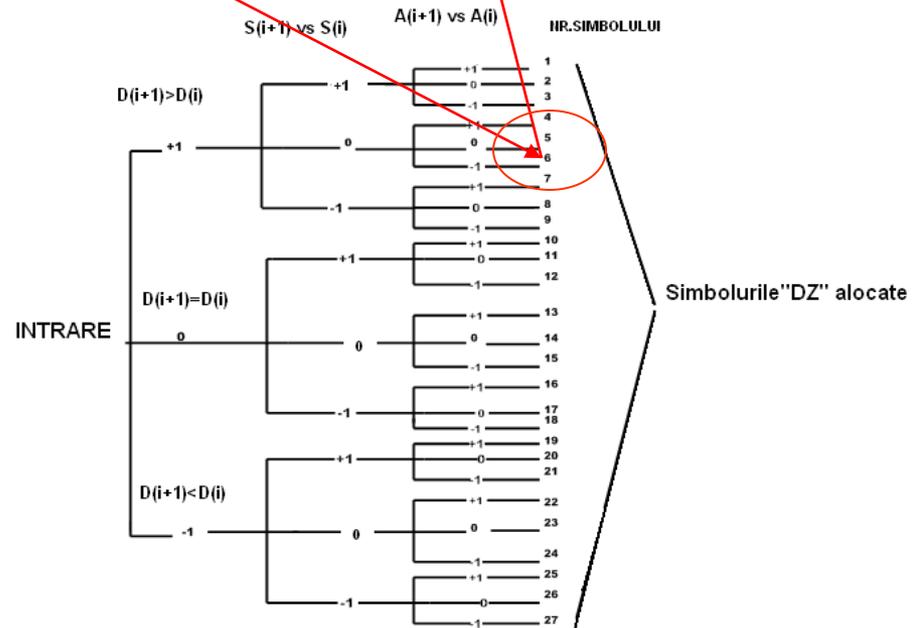
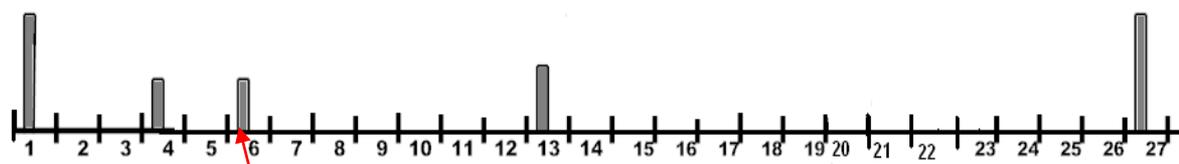


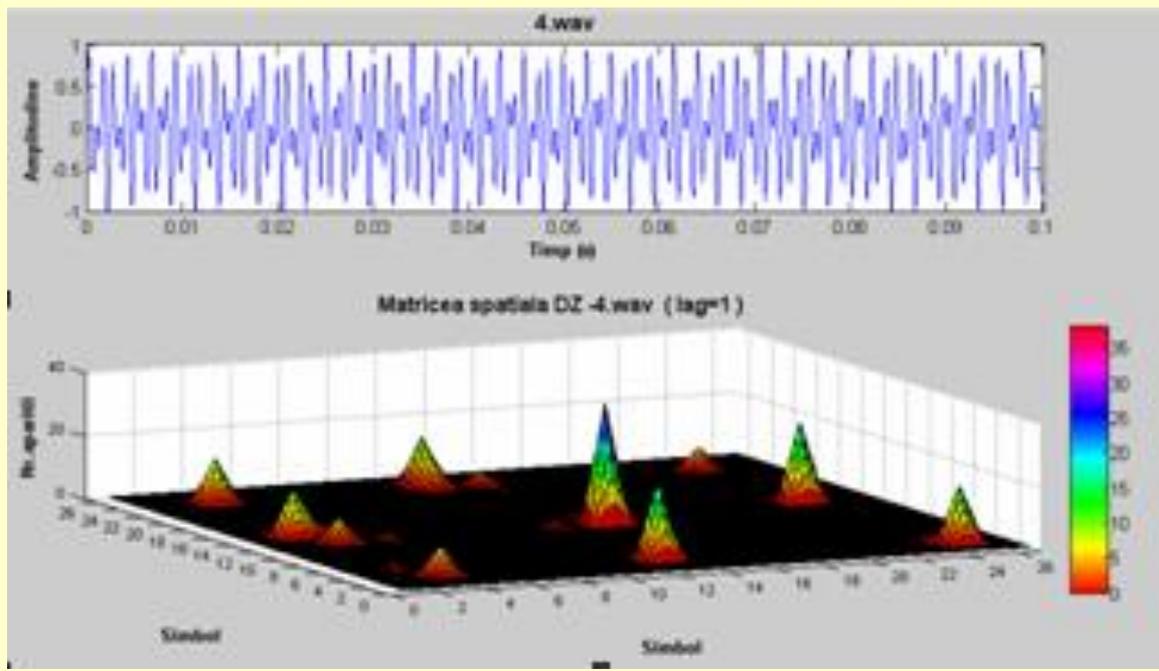
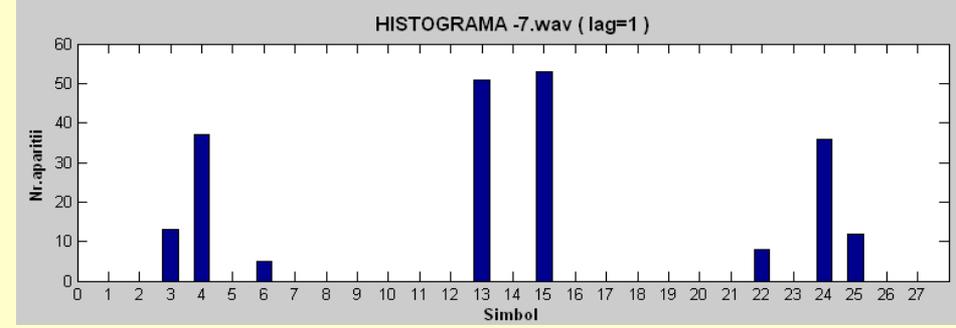
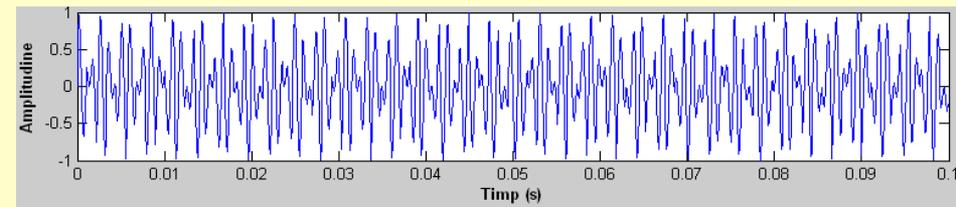
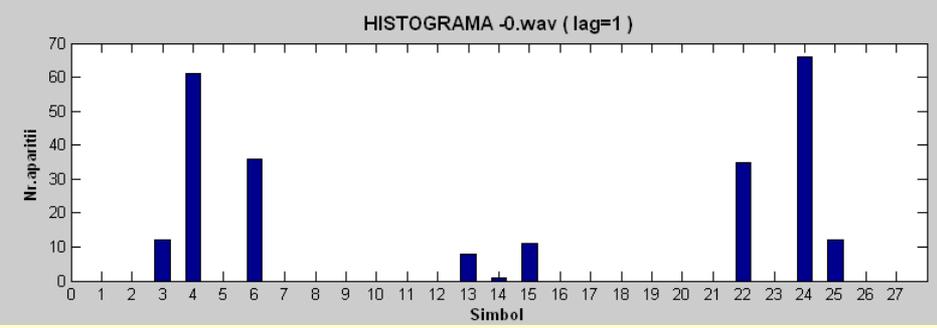
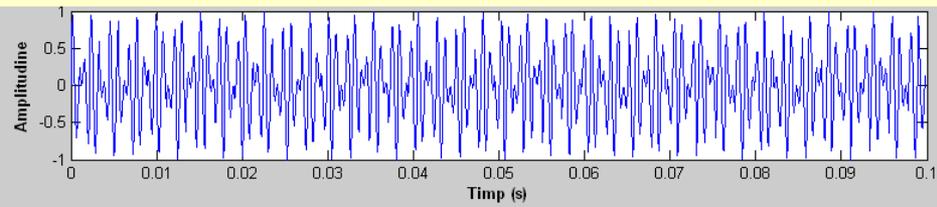
L'espace tridimensionnel  
les 3 descripteurs



• On voit attribuer un symbole à partir de 1 ..... 27 sur un

E2/E1	E3/E2	E4/E3	E5/E4	E6/E5	E7/E6	E8/E7
D2>D1=>+	D3=D2=>0	D4>D3=>+	D5<D4=>-	D6>D5=>+	D7>D6=>+	D8<D7=>-
S2=S1=>0	S3=S2=>0	S4>S3=>+	S5<S4=>-	S6>S5=>+	S7=S6=>0	S8<S7=>-
A2<A1=>-	A3>A2=+	A4=A3=>+	A5<A4=>-	A6>A5=>+	A7>A6=>+	A8<A7=>-
6	13	1	27	1	4	7





EX. – Decodeur DTMF

- Tableaux des distances du module entre les 12 tons DTMF:

	0	1	2	3	4	5	6	7	8	9	*	#
0	0	125	133	125	215	115	93	195	80	84	133	97
1	125	0	36	112	90	26	82	98	57	83	92	60
2	133	36	0	94	88	34	82	110	67	83	112	52
3	125	112	94	0	180	104	52	190	109	61	184	118
4	215	90	88	180	0	102	164	38	139	165	98	126
5	115	26	34	104	102	0	70	102	45	71	88	42
6	93	82	82	52	164	70	0	168	65	33	140	74
7	195	98	110	190	38	102	168	0	127	169	76	118
8	80	57	67	109	139	45	65	127	0	50	79	29
9	84	83	83	61	165	71	33	169	50	0	125	61
*	133	92	112	184	98	88	140	76	79	125	0	72
#	97	60	52	118	126	42	74	118	29	61	72	0

## AVANTAGES

- la puissance de calcul nécessaire est au moins d'un ordre de grandeur plus faible que les autres approches (fréquence)
- Convient pour les applications - en temps réel
- Simplicité de mise en œuvre - à faible coût
- Robustesse
- Peut différencier les signaux presque identiques dans le domaine fréquentiel

## INCONVENIENTS

- La formation MLP nécessite beaucoup de temps
- Les distances ne sont pas suffisantes pour les applications biométriques

## DOMAINES D'APPLICATION

- La vérification biométrique, la reconnaissance vocale
- Le diagnostic repose sur les vibrations des ponts, des turbines, des moteurs ou d'autres sources de vibrations naturelles, même le diagnostic médical