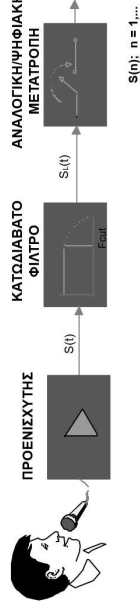


### Αναλογική Επεξεργασία

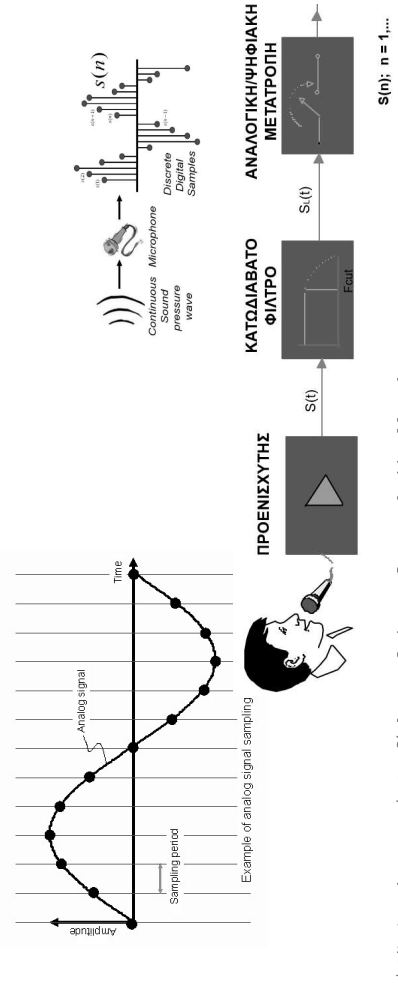
- Μικρόφωνο,
- Μετατροπή ακουστικό σε ηλεκτρικό,
- Πηγή μεγάλης παραμόρφωσης (θορύβου).
- Προενίσχυση,
- Ενίσχυση του αναλογικού σήματος,
- Πηγή μεγάλης παραμόρφωσης (θορύβου).



# Προεπεξεργασία Σήματος Ομιλίας Αναλογική/Ψηφιακή Μετατροπή (A/D) Κατωδιαβατό φίλτρο (Περιορισμός Φάσματος): - Περιορισμός φάσματος: $0-f_{cut}$ Hz, - Φάσμα ομιλίας Ανδρών: 0,2–3,2 kHz (4 Formants ). - Φάσμα ομιλίας Γυναίκων, Παιδιών, : 0,2–3,6 kHz I (4 Formants ). ► Α. Ιδανικό κατωδιαβατό φίλτρο. B. Διάγραμμα στοιχειώδους Low Pass φίλτρου ΠΡΟΕΙΣΧΥΤΗΣ ΚΑΤΩΔΙΑΒΑΤΟ ΦΙΛΤΡΟ ΑΝΑΛΟΓΙΚΗ/ΨΗΦΙΑΚΗ ΜΕΤΑΤΡΟΠΗ $s(n); n = 1, \dots$

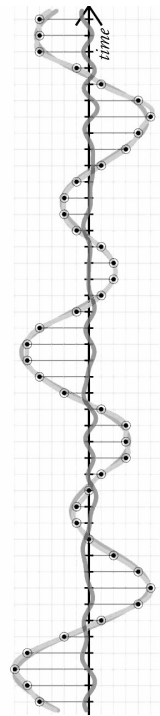
# Προεπεξεργασία Σήματος Ομιλίας Αναλογική/Ψηφιακή Μετατροπή (A/D) Δειγματοληψία (Sampling): - $s(t)$ , δειγματοληψία $T_s (=1/fs)$ , διακριτό σήμα: $s(nT)$ - Με κανονικοποιημένα διαστήματα ( $T=1$ ): $s(n) = s(nT)$ - Διακριτό σήμα μετασχηματισμένο κατά $z (=exp(j\omega T))$ $$S(Z) = \sum_{n=-\infty}^{\infty} s(n)z^{-n}$$

# Προεπεξεργασία Σήματος Ομιλίας Αναλογική/Ψηφιακή Μετατροπή (A/D) Δειγματοληψία (Sampling): - Τυλάχιστον δύο δείγματα ανά κύκλο (Hz)\*, - $fs (1/T_s) \geq 2f_{cut}$ Hz



\* Aliasing: Φασματικής Αναδίπλωσης, θεωρημα δειγματοληψίας Nyquist

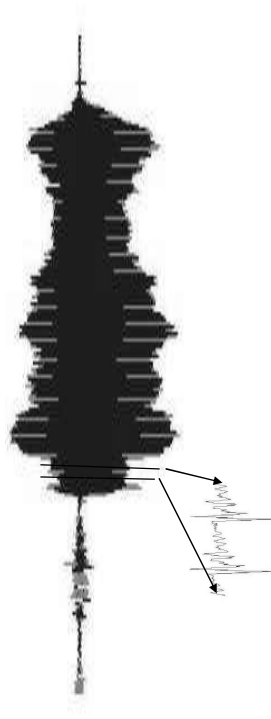
# Προεπεξεργασία Σήματος Ομιλίας Αναλογική/Ψηφιακή Μετατροπή (A/D) Κβαντισμός (Quantization) - Αντιστοίχιση Αναλογικών δειγμάτων σε διακριτές Ψηφιακές τιμές - 8-bit (-128 to 127), - 16-bit (-3,2768 to 3,2767) (16-bit PCM) - 8-bit mu-law; - Σφάλμα Κβαντισμού.



## Βραχύχρονη Ανάλυση – ΠΛΑΙΣΙΟΠΟΙΗΣΗ

### Πλαίσια (Frames)

- ▶ Χαρακτηριστικά ομιλίας (Άρθρωση) μεταβάλλονται αργά.
- ▶ Ρυθμός Άρθρωσης  $\ll$  Ρυθμός Μεταβολής Κυματομορφής,
- ▶ Τμήμα ομιλίας  $N$  δειγμάτων απομονώνεται σαν ήχος με σταθερές ιδιότητες. Πλαίσιο.

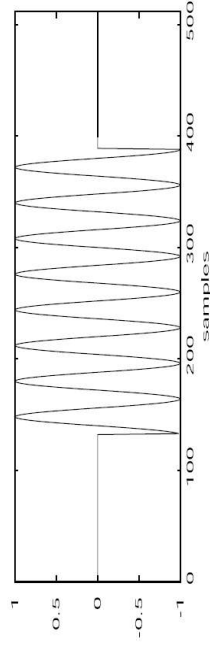


## Βραχύχρονη Ανάλυση – Πλαίσια (Frames)

Παράθυρο:

- ▶ Τετραγωνικό Παράθυρο:

$$w[n] = \begin{cases} 1 & 0 \leq n \leq L-1 \\ 0 & \text{otherwise} \end{cases}$$

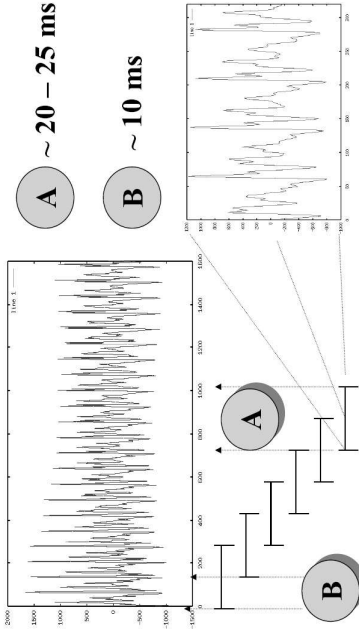


(a) Rectangular window

## Βραχύχρονη Ανάλυση – Πλαισιοποίηση

### Πλαίσια (Frames)

- ▶ Μήκος Πλαισίου (**A**): τυπικά, 20-25ms
  - ▶ Για  $f_s = 10,000 \text{ Hz} \rightarrow N$  (200-250) δείγματα,
- ▶ Μετακίνηση Πλαισίου (**B**) (επικαλυπτόμενα): τυπικά, 5-10ms,
  - ▶ Για  $f_s = 10,000 \text{ Hz} \rightarrow B$  (50-100) δείγματα (βήμα μετατόπισης πλαισίων)

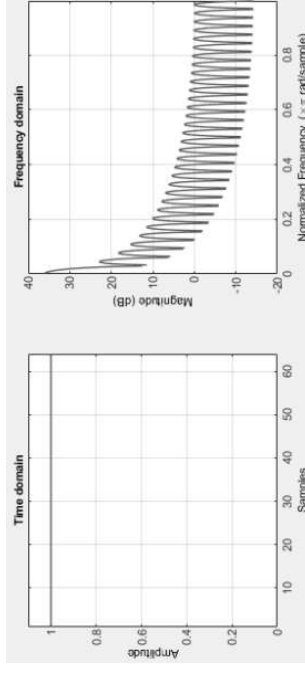


## Βραχύχρονη Ανάλυση – Πλαίσια (Frames)

Παράθυρο:

- ▶ Τετραγωνικό Παράθυρο:

$$w[n] = \begin{cases} 1 & 0 \leq n \leq L-1 \\ 0 & \text{otherwise} \end{cases}$$

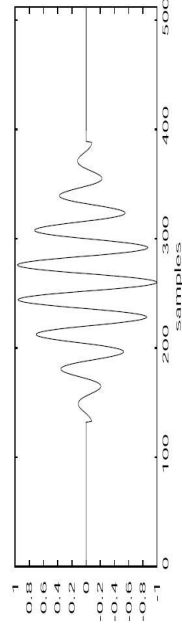
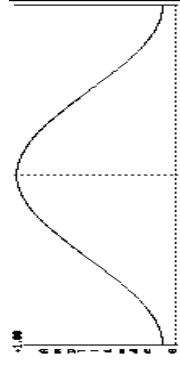


Προεπεξεργασία Σήματος Ομιλίας  
Βραχύχρονη Ανάλυση – Πλαίσια (Frames)

Παράθυρο:

- Hamming Παράθυρο:

$$w[n] = \begin{cases} 0.54 - 0.46 \cos\left(\frac{2\pi n}{L-1}\right) & 0 \leq n \leq L-1 \\ 0 & \text{otherwise} \end{cases}$$

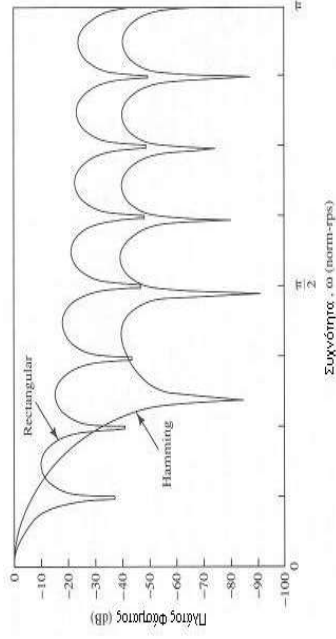


(c) Hamming window

Προεπεξεργασία Σήματος Ομιλίας  
Βραχύχρονη Ανάλυση – Πλαίσια (Frames)

Παράθυρο:

- Τετραγωνικό – Hamming στη συχνότητα:

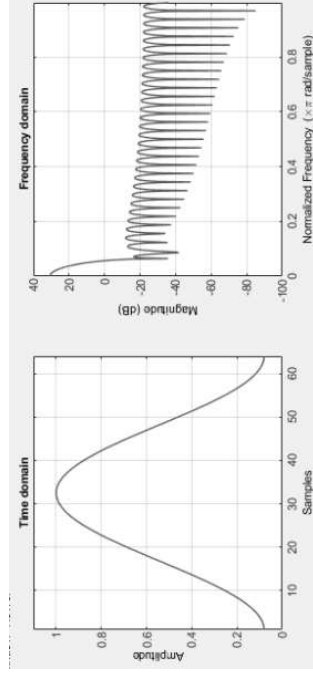


Προεπεξεργασία Σήματος Ομιλίας  
Βραχύχρονη Ανάλυση – Πλαίσια (Frames)

Παράθυρο:

- Hamming Παράθυρο:

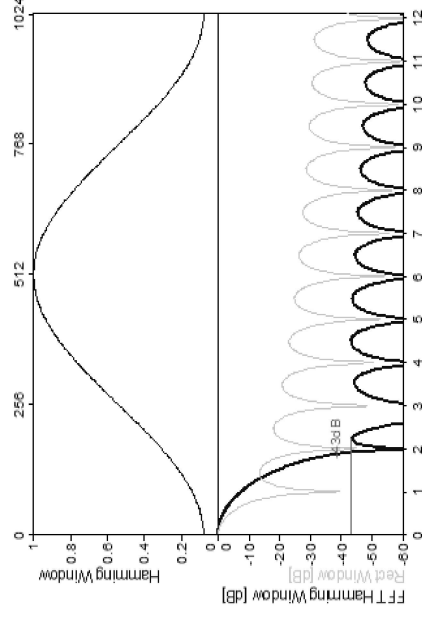
$$w[n] = \begin{cases} 0.54 - 0.46 \cos\left(\frac{2\pi n}{L-1}\right) & 0 \leq n \leq L-1 \\ 0 & \text{otherwise} \end{cases}$$



Προεπεξεργασία Σήματος Ομιλίας  
Βραχύχρονη Ανάλυση – Πλαίσια (Frames)

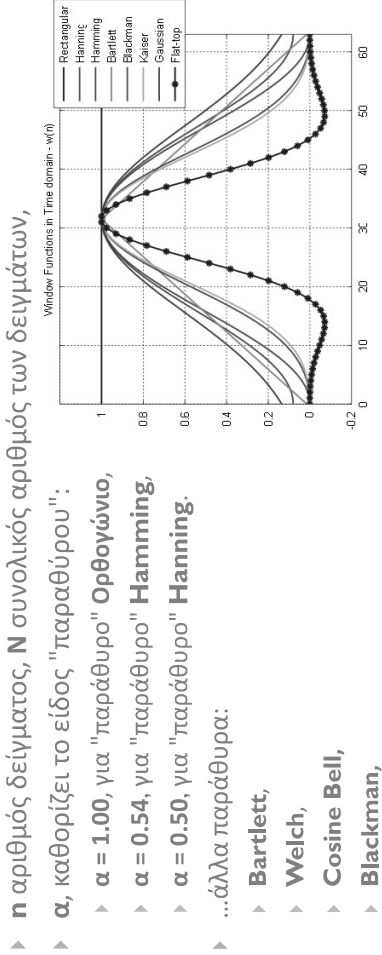
Παράθυρο:

- Τετραγωνικό – Hamming στη συχνότητα:



....διαφορετικά Φίλτρα “Παραθύρων”

$$w(n) = \begin{cases} a + (1-a) \cos(2\pi n/N), & 0 \leq n \leq N-1 \\ 0, & \text{π α ν τ ά α λ λ ο υ} \end{cases}$$



- ▶ ...άλλα παράθυρα:

- ▶ **Bartlett**,
- ▶ **Welch**,
- ▶ **Cosine Bell**,
- ▶ **Blackman**,

Φίλτρο μόνο πόλων

$$H(z) = S(z) / E(z) = L(z) \cdot V(z) \cdot R(z)$$

- ▶ οι 2 πόλοι της γλωττίδας (glottal pulse),

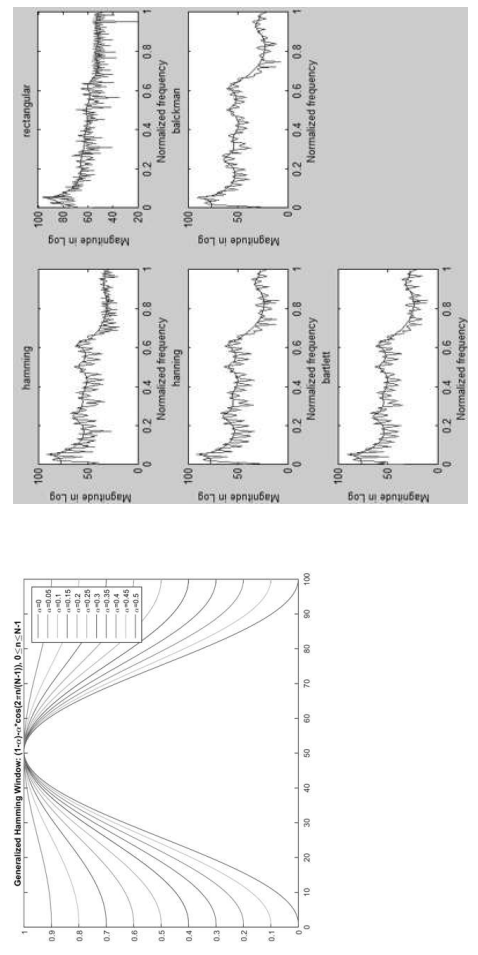
$$L(z) = \frac{1}{(1 - e^{-cT} z^{-1})^2}, \quad \dots cT < 1$$

- ▶ ο 1<sup>ος</sup> απλοποιείται με το Μηδενικό της **Ακτινοβολίας των χειλιών**,

$$R(z) = 1 - z^{-1}$$

- ▶ ο 2<sup>ος</sup>, δημιουργεί την **φασματική κλίση** (spectral tilt),  
...και εξασθενεί τη συνάρτηση μεταφοράς (ΦΚ), με κλίση 6dB/οκτάβα.

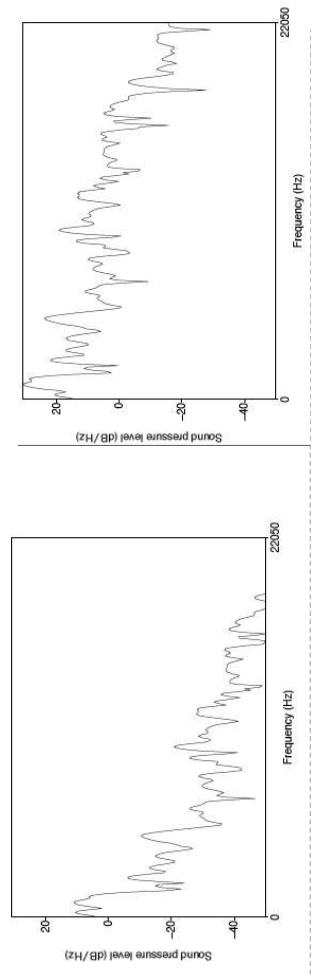
....διαφορετικά Φίλτρα “Παραθύρων”



- ▶ 22

Προ-Έμφασης (Pre-Emphasis)

- ▶ ο 2<sup>ος</sup>, δημιουργεί την **φασματική κλίση** (spectral tilt),  
...και εξασθενεί τη συνάρτηση μεταφοράς (ΦΚ), με κλίση 6dB/οκτάβα.
- ▶ Ζητούμενο: η εξάλειψη του ανεπιθύμητου δεύτερου πόλου
  - ▶ **Ενίσχυση** της ενέργειας στις **υψηλές συχνότητες**
  - ▶ **Πριν** και **μετά** την προ-Έμφαση, για το Φάσμα του φωνήματος [as]



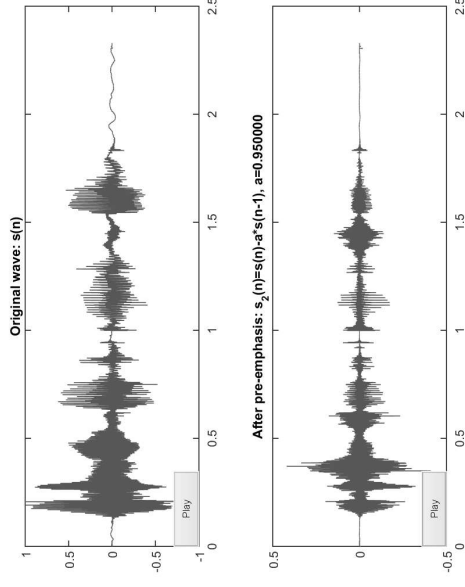
### Προ-Έμφασης (Pre-Emphasis)

- ▶ Προέμφαση: Φίλτρο ενός μηδενικού.
$$P(z) = 1 - \mu z^{-1}$$
  - ▶ ...δηλαδή Πολλαπλασιασμός της ΣΜ του ΦΚ με Φίλτρο ενός μηδενικού
- $$H(z) = S(z) / E(z) = L(z) \cdot V(z) \cdot R(z)$$
  - ▶ Στο χρόνο, η Προέμφαση:
$$s'(n) = s(n) - \mu \cdot s(n-1)$$
    - ▶  $\mu$ , κλίση του φάσματος: (άηχοι,  $\mu < 1$ , ηχηροί  $\mu = 1$ )
    - ▶  $\mu = R(1)/R(0)$ 
      - ▶  $R(0)$ , ενέργεια του πλαισίου,
      - ▶  $R(1)$ , 1<sup>ος</sup> συντελεστής αυτοσυσχέτισης.

▶ 25

Νίκος Φακωτάκης    Τεχνολογία Ομιλίας

### Προ-Έμφασης (Pre-Emphasis)



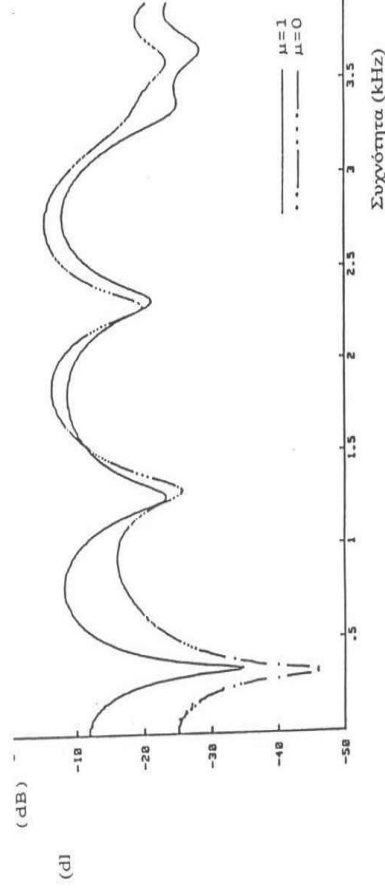
▶ 26

Νίκος Φακωτάκης    Τεχνολογία Ομιλίας

## ΠΡΟΕΠΕΞΕΡΓΑΣΙΑ ΣΗΜΑΤΟΣ ΟΜΙΛΙΑΣ... Ψηφιακή Προεπεξεργασία

### ΠΡΟΕΠΕΞΕΡΓΑΣΙΑ ΣΗΜΑΤΟΣ ΟΜΙΛΙΑΣ Ψηφιακή Προεπεξεργασία

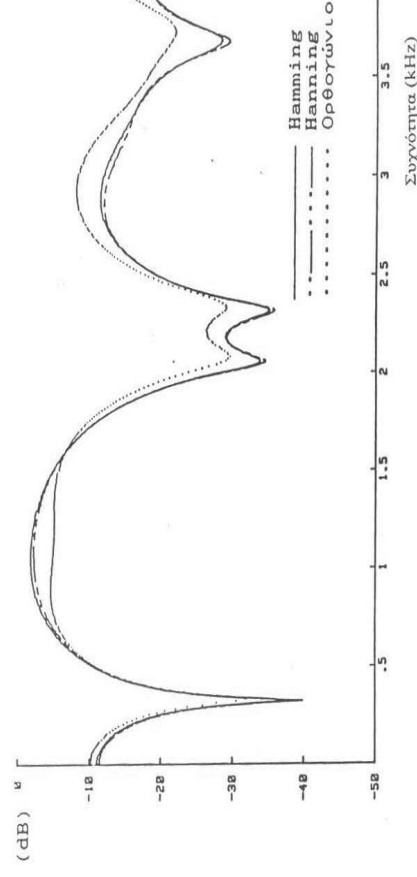
.....Προέμφαση εξομαλυσμένου ηχηρού πλαισίου / $\mu$ /



▶ 27

Νίκος Φακωτάκης    Τεχνολογία Ομιλίας

Ηχηρό πλαίσιο ομιλίας με παράθυρο: α) Ορθογώνιο, β) **Hamming**, γ) **Hanning**.



▶ 28

Νίκος Φακωτάκης    Τεχνολογία Ομιλίας