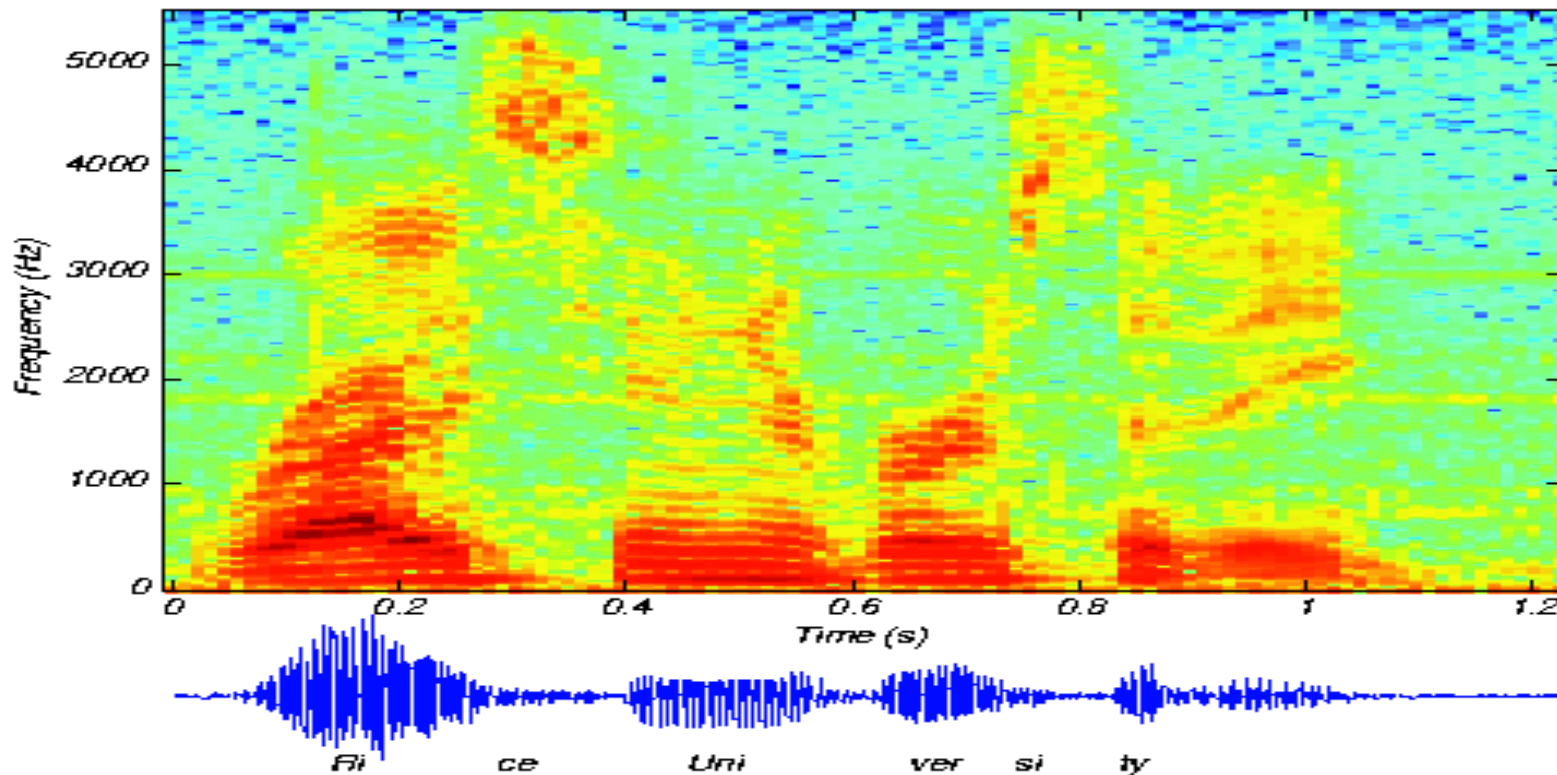


Η ΦΩΝΗ ΩΣ ΣΗΜΑ



Αναστασία Γεωργάκη, Αν. Καθηγήτρια,
Τμήμα Μουσικών Σπουδών, Εργαστήριο Μουσικής Ακουστικής Και τεχνολογίας,
Labmat.music.uoa.gr
georgaki@music.uoa.gr
ΕΚΠΑ

Η φωνή ως σήμα

- Περι φωνής (είδη, σημειολογία, ποιότητα)
 - Η φυσιολογία/λειτουργία της φωνής
 - Ανατομικά χαρακτηριστικά
 - Φυσιολογικά χαρακτηριστικά
 - Μοντέλο λειτουργίας
 - Μέθοδοι και τεχνολογίες ανάκτησης και διερεύνησης του φωνητικού σήματος
 - Το φωνητικό σήμα στην ομιλούσα φωνή
 - Αντικειμενικά χαρακτηριστικά
 - Πηγή- φίλτρο (σήμα-σύστημα)
 - Φωνήεντα –σύμφωνα
- Διαφοροποίηση μεταξύ τραγουδιστής και ομιλούσας φωνής

ΠΕΡΙ ΦΩΝΗΣ

- ΔΙΕΠΙΣΤΗΜΟΝΙΚΗ ΠΡΟΣΕΓΓΙΣΗ
- ΙΑΤΡΙΚΗ (φωνιατρική, ωτορινολαρυγγολογία και ψυχιατρική)
- ΠΛΗΡΟΦΟΡΙΚΗ (Ακουστική ανάλυση, επεξεργασία σήματος, Μοντελοποίηση, συνθετική φωνή)
- ΓΛΩΣΣΟΛΟΓΙΑ (φωνολογία)
- ΜΟΥΣΙΚΟΛΟΓΙΑ (ιστορική/ανθρωπολογική/μουσική ακουστική)
- ΦΩΝΗΤΙΚΗ (Ερμηνεία)
- ΜΟΥΣΙΚΗ ΕΚΠΑΙΔΕΥΣΗ

Είδη φωνής

-
- «Ο άνθρωπος έχει τρία είδη φωνών, ήτοι την ομιλούσα ή αρθρωμένη φωνή, την άδουσα ή μελωδική φωνή και την παθητική ή τονισμένη φωνή, η οποία χρησιμεύει ως γλώσσα των παθών». (jean Luc Nancy)
-
- Jean-Luc Nancy , «Vox clamans in deserto”. *Demande: Philosophie, littérature*, Paris:Galilée, 2015, σελ. 331-343.

Είδη φωνής σε σχέση με το ύφος του κειμένου

- Η ΑΡΘΡΩΜΕΝΗ ΦΩΝΗ

- Η ΦΩΝΗ ΤΟΥ ΦΙΛΟΣΟΦΟΥ
- Κώστας Αξελός (19:07-19:52)
- <https://www.youtube.com/watch?v=TYI8GRJGnBY>



- Η ΦΩΝΗ ΤΟΥ ΠΟΙΗΤΗ Ι
- Οδ. Ελύτης (10.34-10.57)
- <https://www.youtube.com/watch?v=v851rSIZbBQ>



- Η ΦΩΝΗ ΤΟΥ ΠΟΙΗΤΗ ΙΙ
- Αγγ. Σικελιανός (0:00-0:22)
- <https://www.youtube.com/watch?v=wAI1SnRSxEw>



- ΦΩΝΗ ΤΩΝ ΠΑΘΩΝ

- Η ΦΩΝΗ ΤΗΣ ΤΡΑΓΩΔΟΥ (0:00-0:46)
- Κατίνα Παξινού
- <https://www.youtube.com/watch?v=PlbSc6kSQXo>



- ΑΔΟΥΣΑ ΦΩΝΗ

- ΦΩΝΗ στη Μελοποιημένη ποίηση
- Πλάτωνος –Καρυωτάκης (Σαβίννα) (0:00-0:33)
- <https://www.youtube.com/watch?v=aZ-aWOyGm60>



- ΦΩΝΗ ΣΤΗΝ ΟΠΕΡΑ
- Maria Callas (Casta Diva)
- 0:00-0:33
- <https://www.youtube.com/watch?v=TYI8GRJGnBY>



ΣΗΜΕΙΟΛΟΓΙΑ ΤΗΣ ΦΩΝΗΣ ΣΤΟ ΠΡΟΦΟΡΙΚΟ ΛΟΓΟ

- *Σημειολογία*

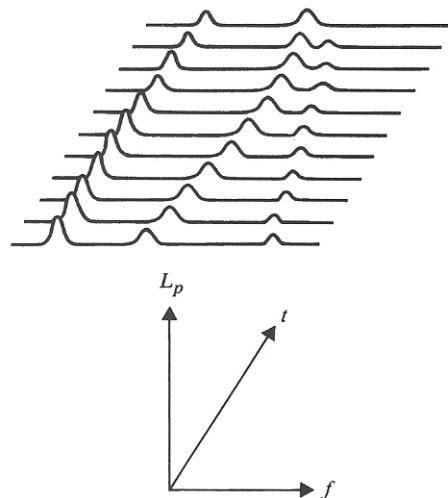
(πολιτισμική καταγωγή, εκπαίδευση, χαρακτήρας, ψυχολογική διάθεση, ηλικία)

- *Προφορική γλώσσα*

(φωνολογικοί, συντακτικοί, λεκτικοί, προσωδιακοί και παραγλωσσικοί υποκώδικες)

Αναγνώριση ομιλίας

- Μπορούμε να αναγνωρίσουμε πάνω απο 30 φωνήματα/sec και με γρήγορο ρυθμό μέχρι 400 λέξεις /min
- Τεστ της άρθρωσης



Ποιότητα της Φωνής (Noth, 1995)

2.2.2 VOICE QUALITY

Voice qualities are speech variables which characterize a speaker's "tone of voice" in adjustment to situational factors. Trager classifies them as follows (1958: 5):

- pitch range and control (spread or narrowed [as in monotone speech])
- vocal lip control (from hoarseness to openness)
- glottis control (sharp or smooth transitions)
- articulatory control (forceful vs. relaxed speech)
- rhythm control (smooth or jerky)
- resonance (from resonant to thin)
- tempo (increased or decreased)

For further paralinguistic studies in **voice** quality, see Crystal (1969; 1975), Ostwald (1973), Laver (1980), Scherer & Giles, eds. (1979), and Scherer, ed. (1982).

Δημιουργία του έναρθρου λόγου.

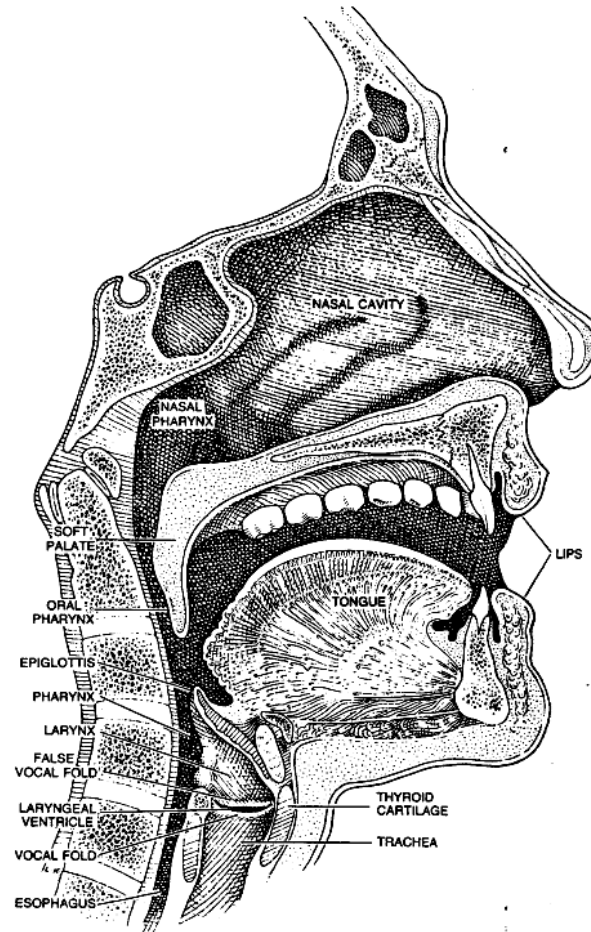
- Από την αναπνοή ...στη φωνή (πηγή ενέργειας)
- Οι χορδές πάλλονται... (ταλαντωτής)
- Το εύπλαστο αντηχείο
- Ο σχηματισμός των φωνηέντων, συλλαβών και φράσεων (αρθρωτές)
- Το συναίσθημα σαν έκφραση στο τραγούδι
- Τα ρετζίστρα
- <http://www.youtube.com/watch?v=iYpDwhpILkQ>

Ακουστική/φυσιολογία/λειτουργία

- ΣΥΜΠΙΕΣΤΗΣ--> ΤΑΛΑΝΤΩΤΗΣ---> ΑΝΤΗΧΕΙΟ
- ΠΝΕΥΜΟΝΕΣ→ ΦΩΝ. ΧΟΡΔΕΣ-> ΦΩΝ. ΣΩΛΗΝ
— (φωνή)
- ΑΝΑΠΝΟΗ--> ΦΩΝΗΣΗ-----> ΑΡΘΡΩΣΗ
- Κοιλιακοί λαρυγγικοί χείλη, γλώσσα
- Διάφραγμα
- <https://www.youtube.com/watch?v=h7zsjD1PLvw>

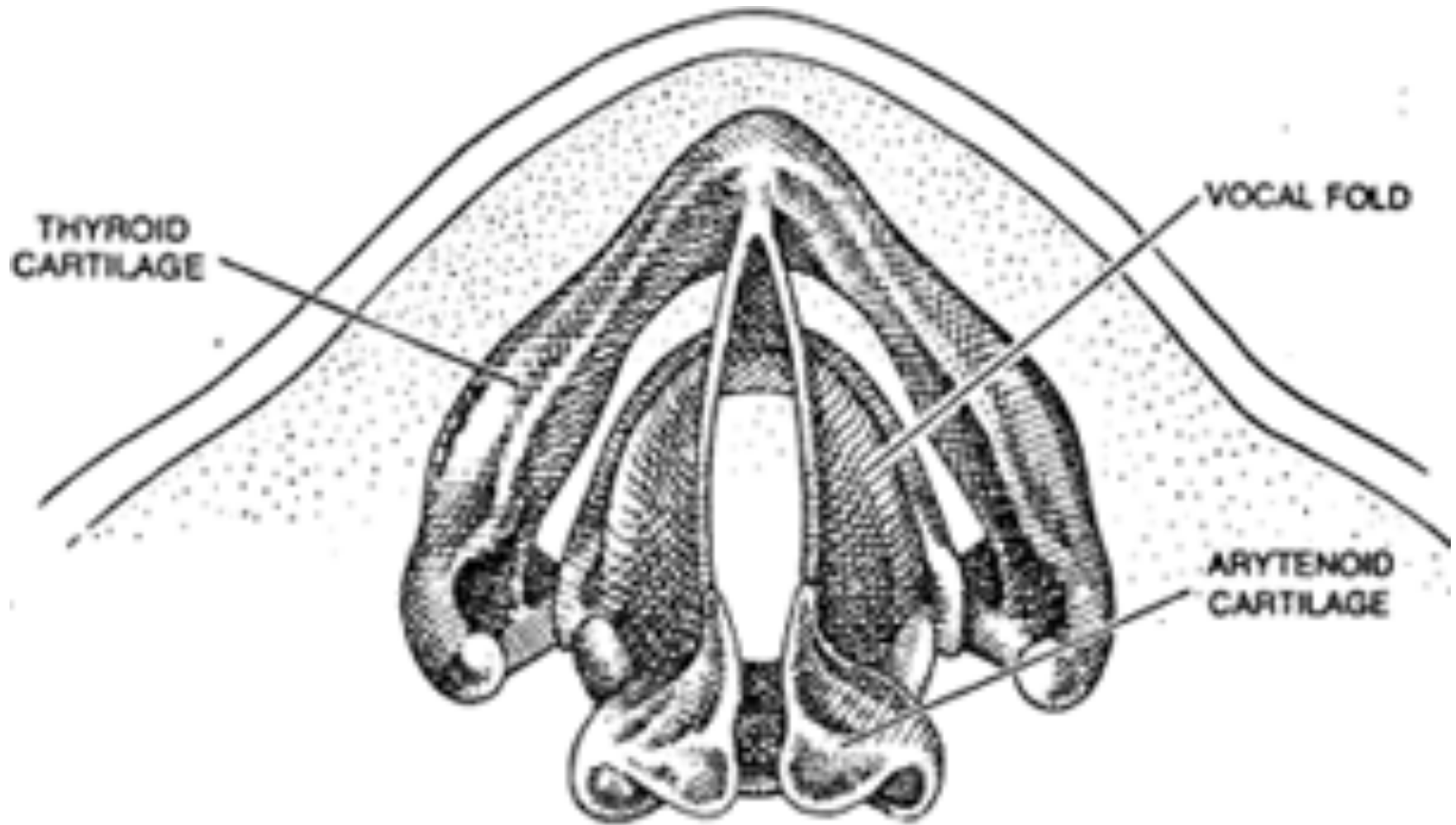
Ανατομία της φωνητικής συσκευής

<https://www.youtube.com/watch?v=uTOhDghCKQs&NR=1>



Οι Φωνητικές χορδές

<http://www.youtube.com/watch%3Fv=iYpDwhpILkQ>



ΤΕΧΝΟΛΟΓΙΕΣ ΑΠΟΚΩΔΙΚΟΠΟΙΗΣΗΣ ΤΗΣ ΛΕΙΤΟΥΡΓΙΑΣ ΤΗΣ ΦΩΝΗΤΙΚΗΣ ΣΥΣΚΕΥΗΣ ΚΑΙ ΤΟΥ ΣΗΜΑΤΟΣ

ΣΥΓΧΡΟΝΕΣ ΜΕΘΟΔΟΙ ΑΞΙΟΛΟΓΗΣΕΩΝ ΚΑΙ ΜΕΤΡΗΣΕΩΝ ΤΩΝ ΠΑΡΑΜΕΤΡΩΝ ΤΗΣ ΦΩΝΗΣ (ΟΜΙΛΟΥΣΑΣ ΚΑΙ ΤΡΑΓΟΥΔΙΣΤΗΣ).

ΑΝΑΠΝΕΥΣΤΙΚΗ ΛΕΙΤΟΥΡΓΙΑ

-ΣΠΙΡΟΜΕΤΡΙΚΕΣ ΑΞΙΟΛΟΓΗΣΕΙΣ
-ΚΙΝΗΜΑΤΙΚΕΣ ΑΞΙΟΛΟΓΗΣΕΙΣ
-ΑΕΡΟΔΥΝΑΜΙΚΕΣ ΑΞΙΟΛΟΓΗΣΕΙΣ

ΛΑΡΥΓΓΙΚΗ ΛΕΙΤΟΥΡΓΙΑ

-ΗΛΕΚΤΡΟΓΛΩΤΤΟΓΡΑΦΙΚΗ ΑΞΙΟΛΟΓΗΣΗ
ELECTROGLOTTOGRAPHY (EGG)
-ΑΕΡΟΔΥΝΑΜΙΚΗ ΑΞΙΟΛΟΓΗΣΗ
AERODYNAMIC ASSESSMENT (AEROPHONE)
-ΗΛΕΚΤΡΟΜΥΟΓΡΑΦΙΚΗ ΑΞΙΟΛΟΓΗΣΗ
ELECTROMYOGRAPHY (EMG)

ΑΡΘΡΩΤΙΚΗ ΛΕΙΤΟΥΡΓΙΑ

-ΧΕΙΛΙΚΗ ΚΑΙ ΓΛΩΣΣΙΚΗ ΠΙΕΣΗ
-ΟΥΡΑΝΙΣΚΟΓΡΑΦΙΚΗ ΑΞΙΟΛΟΓΗΣΗ
ELECTROPALATOGRAPHY (EPG)
-ΗΛΕΚΤΡΟΜΑΓΝΗΤΙΚΗ ΑΡΘΡΩΤΟΓΡΑΦΙΑ
ELECTROMAGNETIC ARTICULOGRAPHY (EMA)

ΦΩΝΗΤΙΚΗ ΟΔΟΣ
(ΜΗΚΟΣ Κ ΣΧΗΜΑ)
ΚΑΙ ΕΚΤΙΜΗΣΗ
ΦΩΝΗΤΙΚΗΣ ΛΕΙΤΟΥΡΓΙΑΣ

-ΤΟΜΟΓΡΑΦΙΚΗ ΑΞΙΟΛΟΓΗΣΗ
MAGNETIC RESONANCE IMAGING (MRI)
-ΥΠΟΛΟΓΙΣΤΙΚΗ ΤΟΜΟΓΡΑΦΙΚΗ ΑΞΙΟΛΟΓΗΣΗ
COMPUTATIONAL TOMOGRAPHY (CT)
-ΑΞΙΟΛΟΓΗΣΗ ΜΕΣΩ ΑΚΤΙΝΩΝ-Χ
X-RAY IMAGING (XRI)
-ΦΑΣΜΑΤΟΓΡΑΦΗΜΑ

ΑΝΑΛΥΣΗ ΤΟΥ ΣΗΜΑΤΟΣ ΤΗΣ ΦΩΝΗΣ

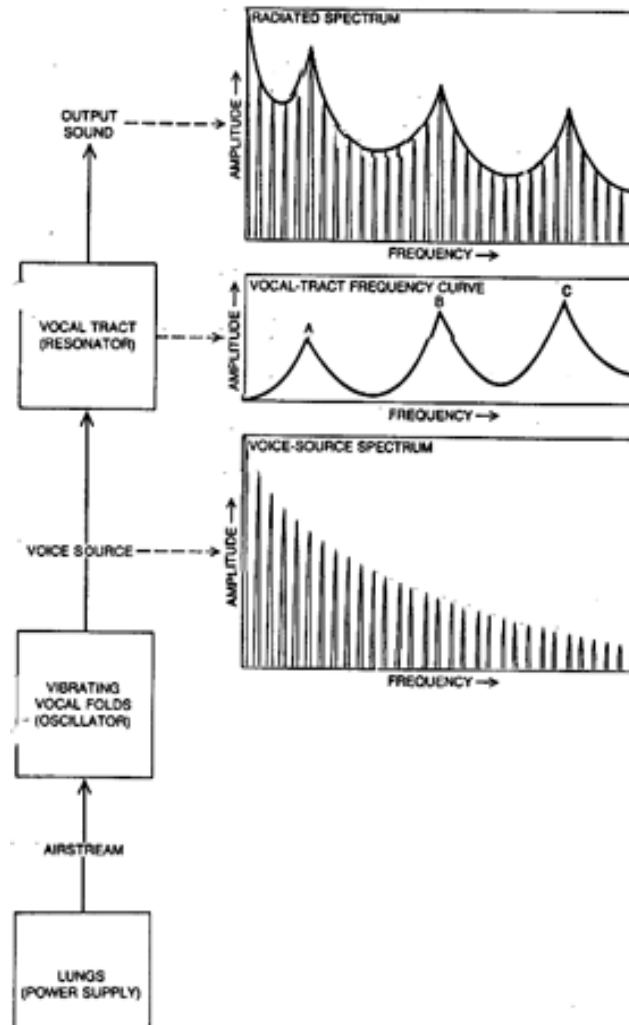
ΜΕΜΟΝΩΜΕΝΟΙ ΗΧΟΙ

- -Τονικό ύψος (f_0)-φωνητικές χορδές
- -Ενταση/ακουστότητα (a)--SPL
- -Χρόχρωμα (spectrum, formants)-----αντηχείο
- -διάρκεια (duration)

ΦΡΑΣΗ

- Μελωδική γραμμή
- Ρυθμική γραμμή
- Διακύμανση εντάσεων
- Υφή του λόγου

ΠΗΓΗ-ΦΙΛΤΡΟ



Pipelines Model

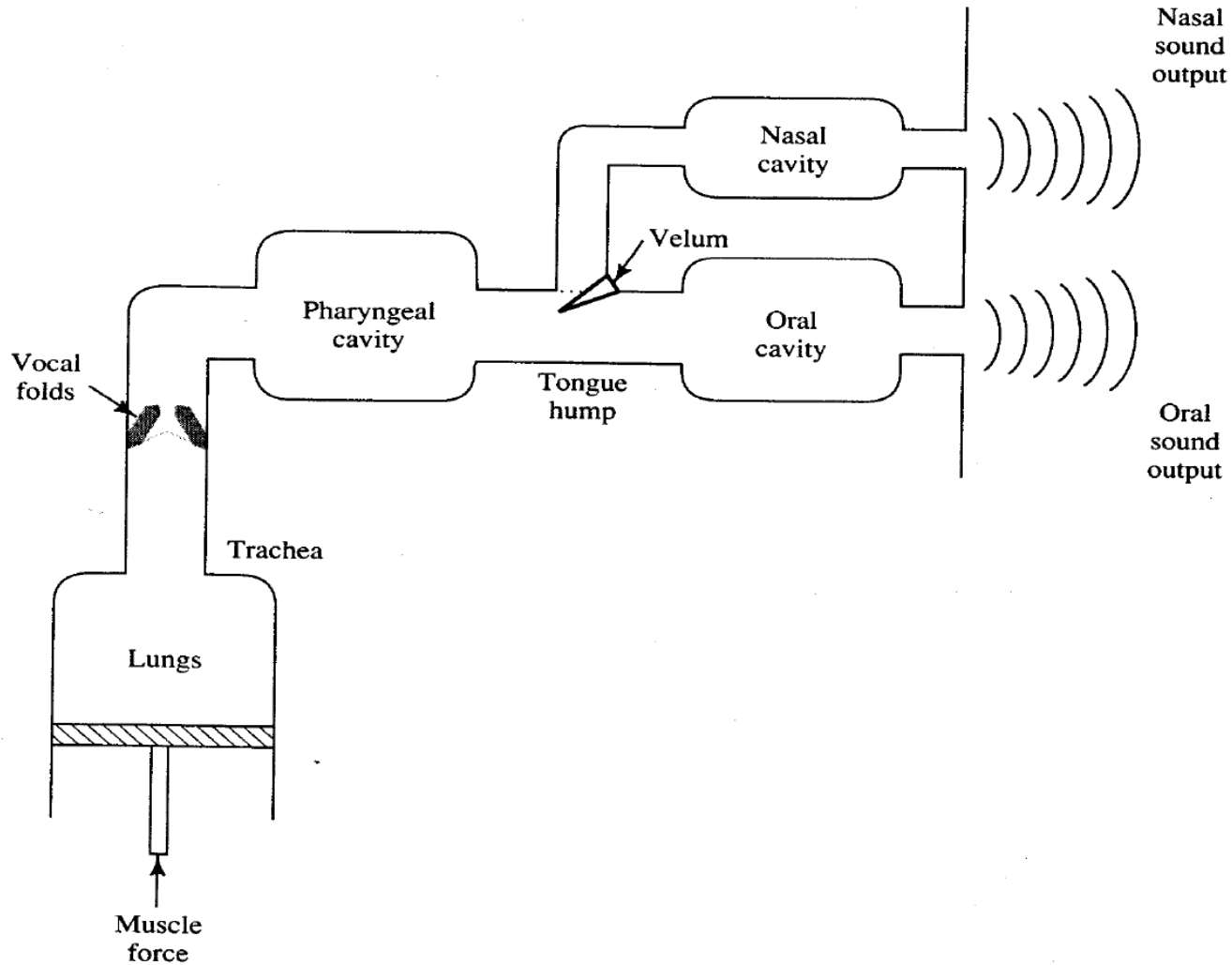
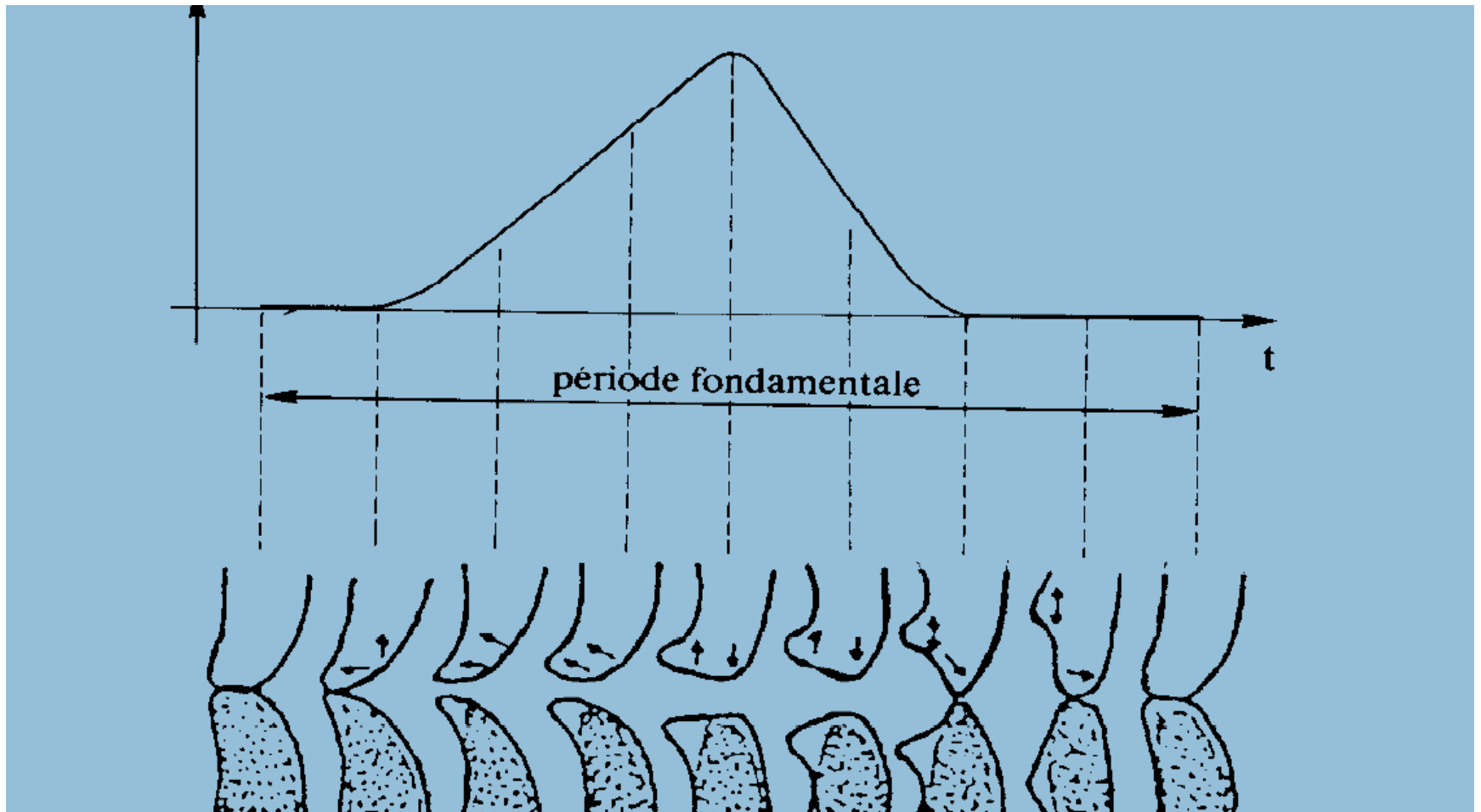
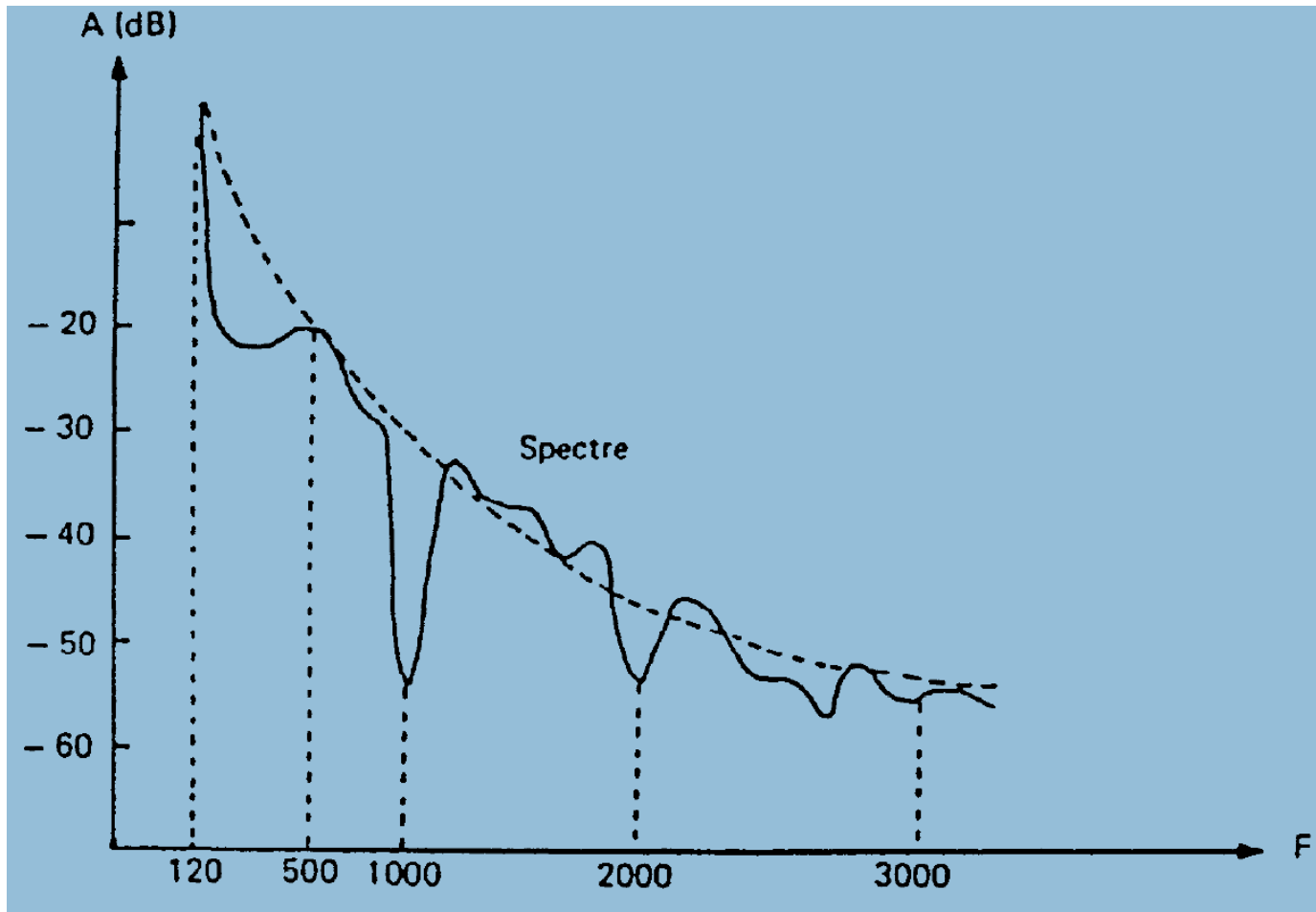


FIGURE 2.2. A block diagram of human speech production.

ΠΗΓΗ (Φων. Χορδές). ΦΩΝΗΣΗ



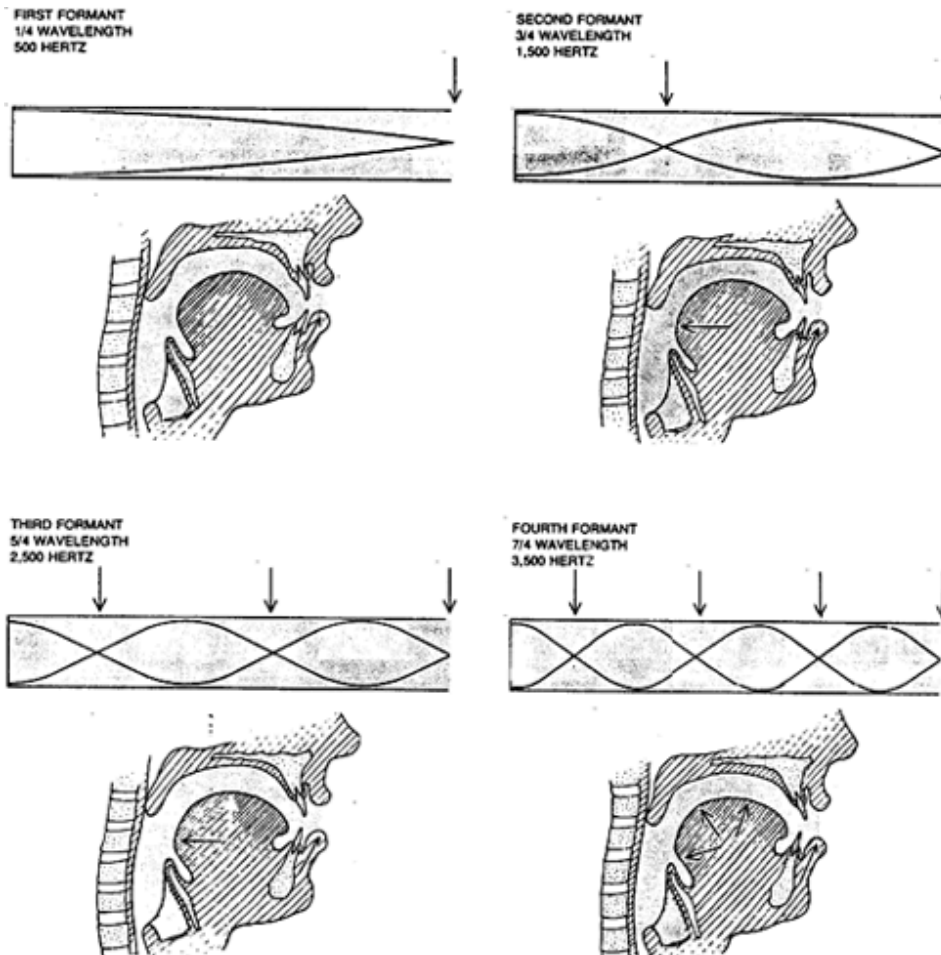
Τυπικό φάσμα (spectrum) του γλωττιδικού σήματος (πηγή, φων. χορδές) (Sundberg, 1989)



ΤΟ ΑΝΤΗΧΕΙΟ:

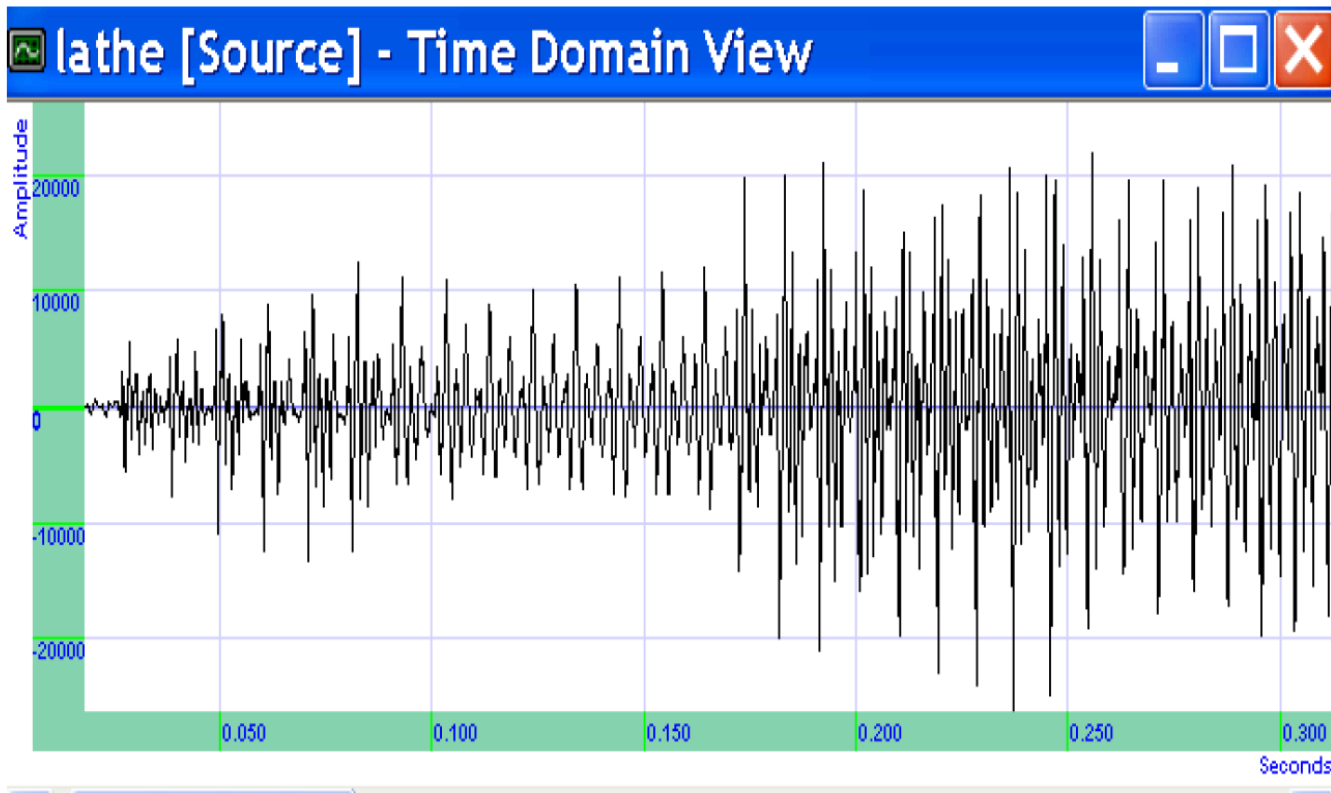
Φωνητικός σωλήνας δρα σαν φίλτρο που διαμορφώνει το σήμα που έρχεται από τις φωνητικές χορδές.

LTI (Linear Time invariant Filter) with a periodic impulse-like input



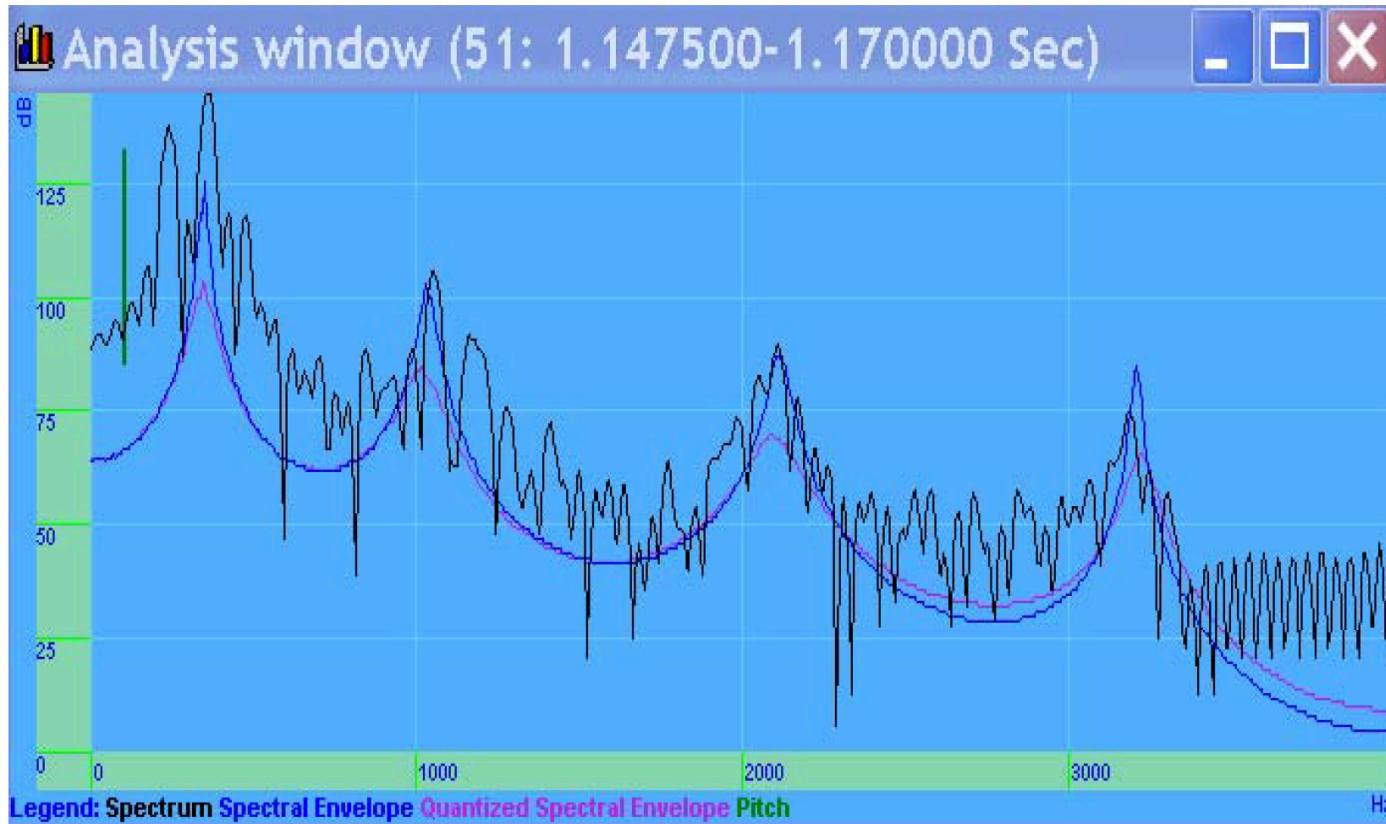
Typical Voiced Sound

1Sec, 10,000 Samples, 8bps, Voiced (“Ahhhhh”)



A Quasi-Periodic Signal

Power Spectrum: Voiced Speech



Είδη φωνητικών σημάτων

- Η πηγή (φων. χορδές) θέτει σε λειτουργία το φίλτρο (φωνητικός σωλήνας) και δημιουργεί:
- Φωνητικά σήματα (voiced) : περιοδικά σήματα-φωνήεντα
- Μη φωνητικά (unvoiced): μη περιοδικά σήματα
- ΦΩΝΗΕΝΤΑ
- ΣΥΜΦΩΝΑ

Speech Sounds Categories

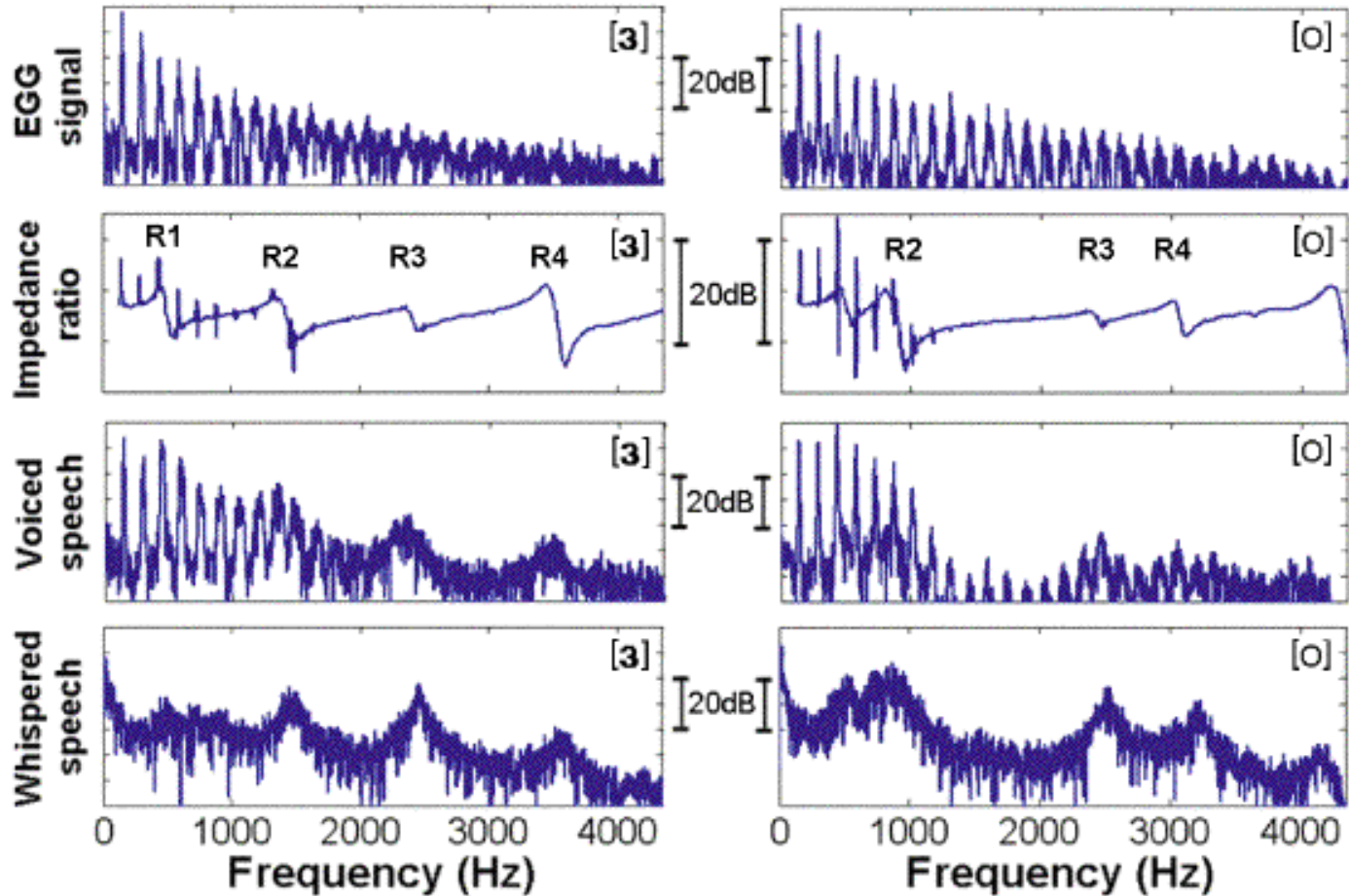
- *Periodic* (Sonorants, Voiced: קולי)
- *Noisy* (Fricatives, Un-Voiced: א-קולי)
- *Impulsive* (Plosive: פּוּצֵץ)

- Example:

In the word “shop,” the “sh,” “o,” and “p” are generated from a noisy, periodic, and impulsive source, respectively.

Διαφορές φωνητικών σημάτων /ε/ και /ο/

- α) σήμα πηγής (EGG)
- β) Transfer function αντηχείου
- γ) διαμορφωμένο σήμα
- δ) ψιθυριστό σήμα (χωρίς τη πηγή)



Εκτάσεις των φωνών -registers

- ΟΜΙΛΙΑ:
 - Άρρεν: 85-155Hz
 - Θήλυ: 165-255 Hz
- ΤΡΑΓΟΥΔΙ
 - Μπάσσοσ- Τενόροσ, Άλτο, Σοπράνο
 - 80 Hz-1100

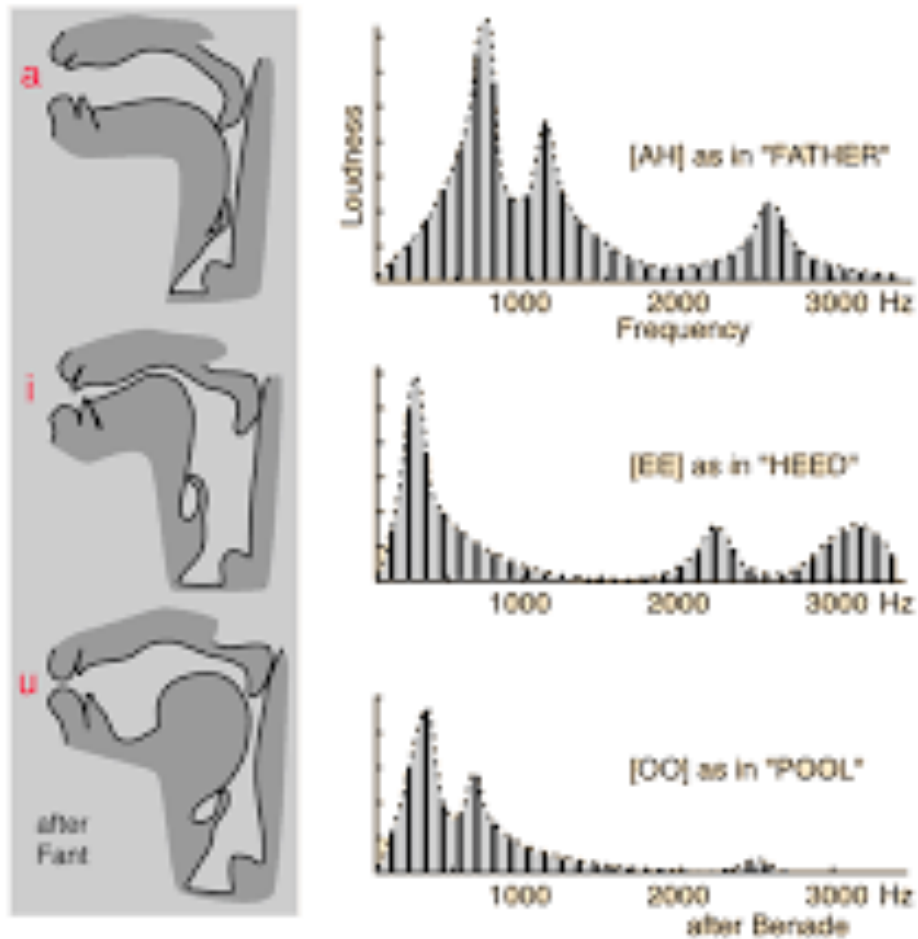
Ρετζίστρα στην άδουσα φωνή

COMMON VOCAL REGISTERS AND PASSAGGIOS BY VOICE TYPE

	Lowest Sung Note	1st Passaggio	2nd Passaggio	3rd Passaggio	Highest Sung Note
Coloratura Soprano	varies	Bb3	Gb4	E5	D6
Lyric Soprano	varies	Bb3	Gb4	E5	C6
Dramatic Soprano	varies	Ab3	F4	Eb5	C6
Mezzo Soprano	varies	Gb3	Eb4	Db5	Bb5
Lyric Tenor	C3	D4	G4	C5	varies
Dramatic Tenor	B2	Db4	Gb4	B4	varies
Baritenor	Ab2	B3	E4	Bb4	varies
Baritone	G2	Bb3	Eb4	Ab4	varies
Bass	F2	A3	D4	F4	varies

Table provided courtesy of Vocal Nebula

Formants



ΠΑΡΑΓΩΓΗ ΤΩΝ ΦΩΝΗΕΝΤΩΝ

Typical Vowels

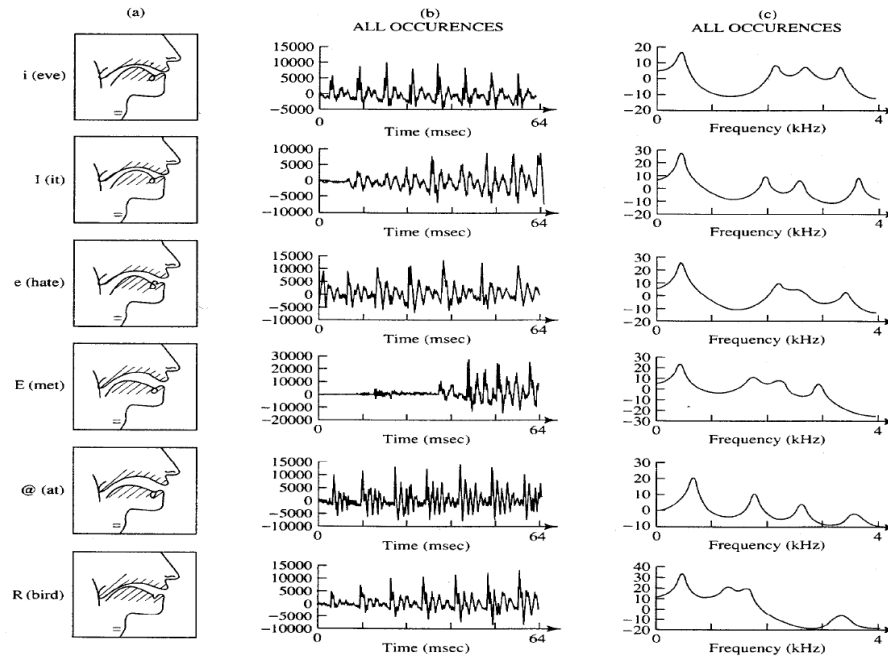
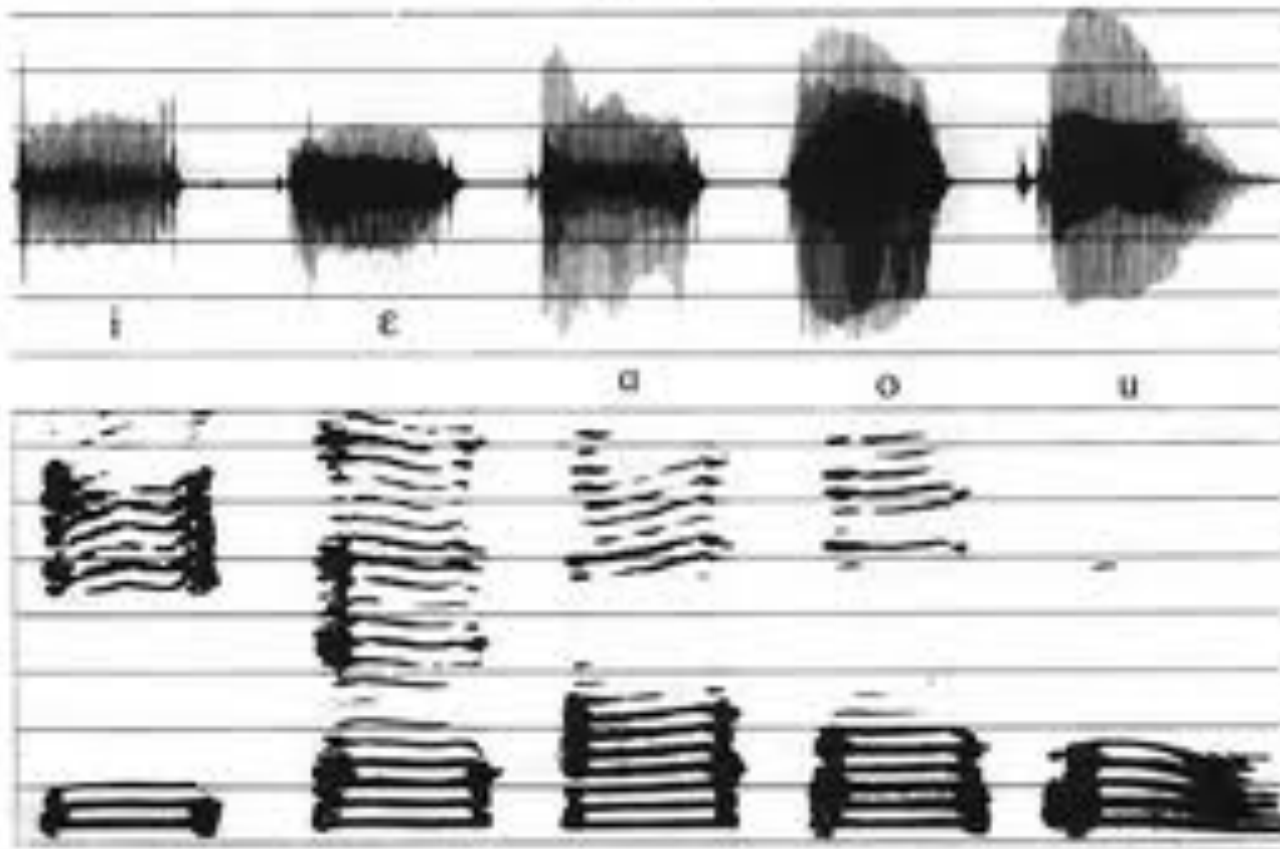


FIGURE 2.10. A collection of features for vowels in American English. Column (a) represents schematic vocal-tract profiles, (b) typical acoustic waveforms, and (c) the corresponding vocal-tract magnitude spectrum for each vowel.

ΦΩΝΗΕΝΤΑ

Time waveform and narrowband spectrogram of [i e a o u]



Μέσες τιμές των formants

Average Formant Locations

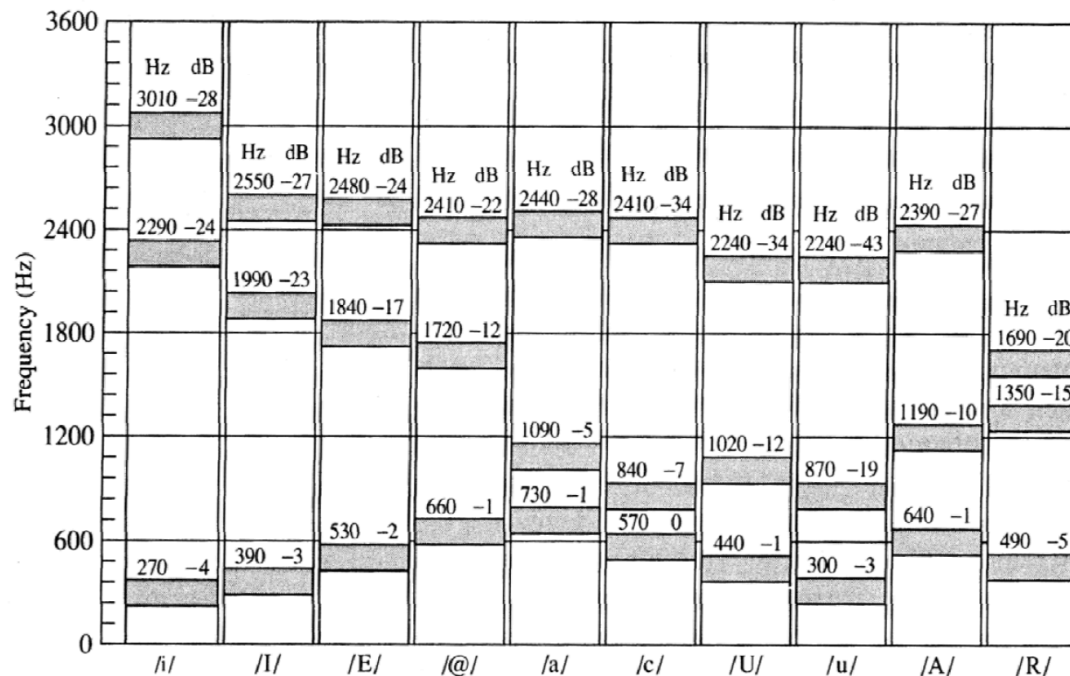


FIGURE 2.11. Average formant locations for vowels in American English (Peterson and Barney, 1952).

Μέση τιμές των formants

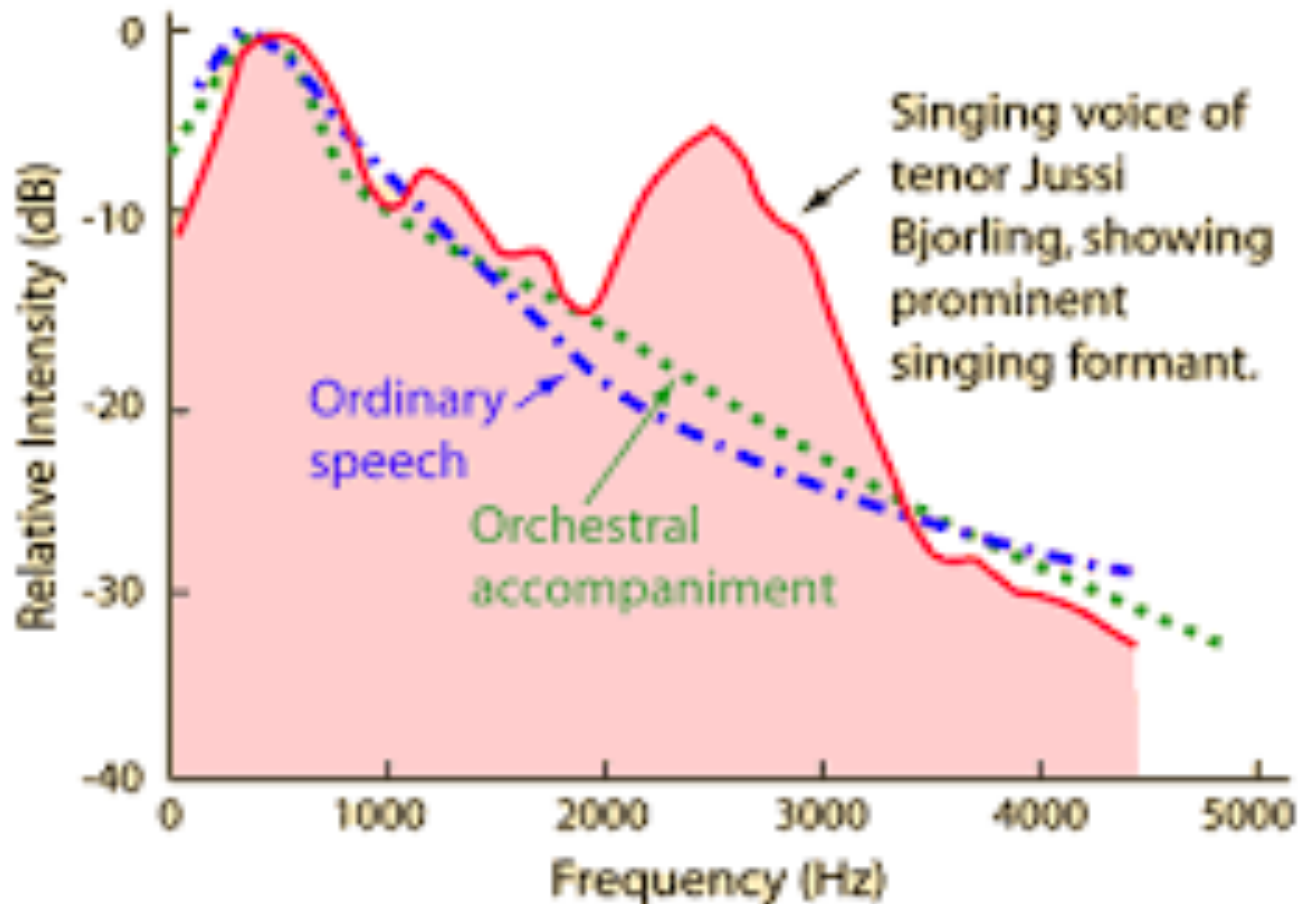
Table 3.2 Average Formant Frequencies for the Vowels. (After Peterson and Barney [11].)

FORMANT FREQUENCIES FOR THE VOWELS					
Typewritten Symbol for Vowel	IPA Symbol	Typical Word	F ₁	F ₂	F ₃
IY	i	(beet)	270	2290	3010
I	ɪ	(bit)	390	1990	2550
E	ɛ	(bet)	530	1840	2480
AE	æ	(bat)	660	1720	2410
UH	ʌ	(but)	520	1190	2390
A	ɑ	(hot)	730	1090	2440
OW	ɔ	(bought)	570	840	2410
U	u	(foot)	440	1020	2240
OO	u	(boot)	300	870	2240
ER	ɜ	(bird)	490	1350	1690

IPA	Worldbet	Example
ɪ:	i:	beet
ɪ	ɪ	bit
ɛ	ɛ	bet
æ	@	bat
ʌ	^	above
u	u	boot
U	U	book
ə	&	above
ɑ	A	father
ɜ	3r	bird
ɝ	&r	butter

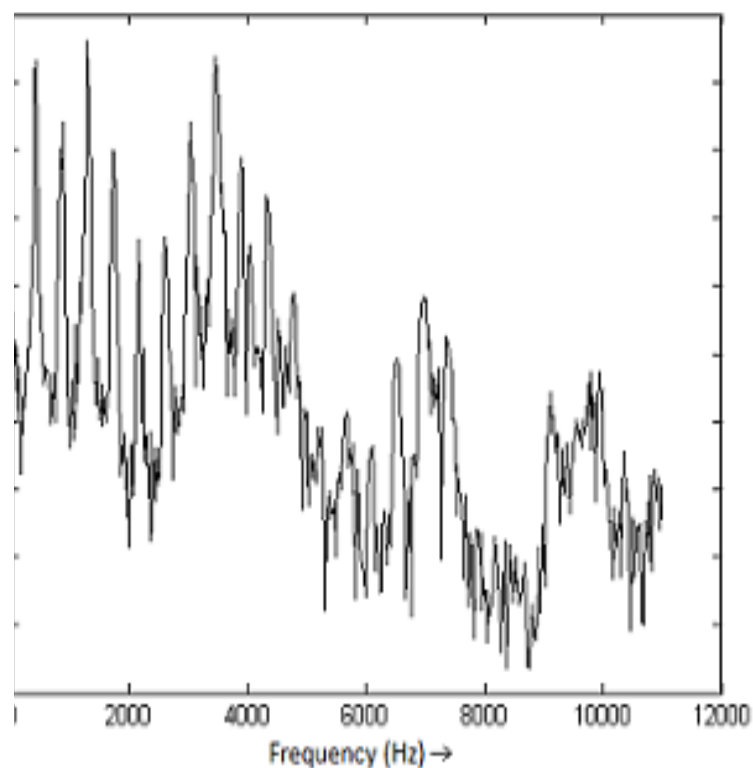
From: Rabiner & Schafer, Digital Processing of Speech Signals

Singer's formant

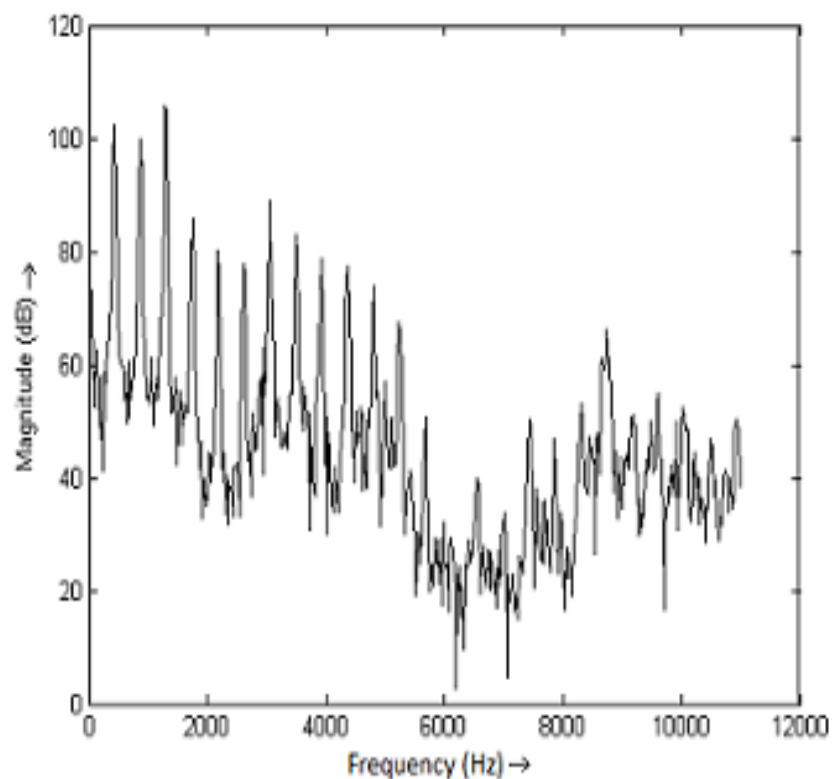


After Sundberg, *The Acoustics of the Singing Voice*

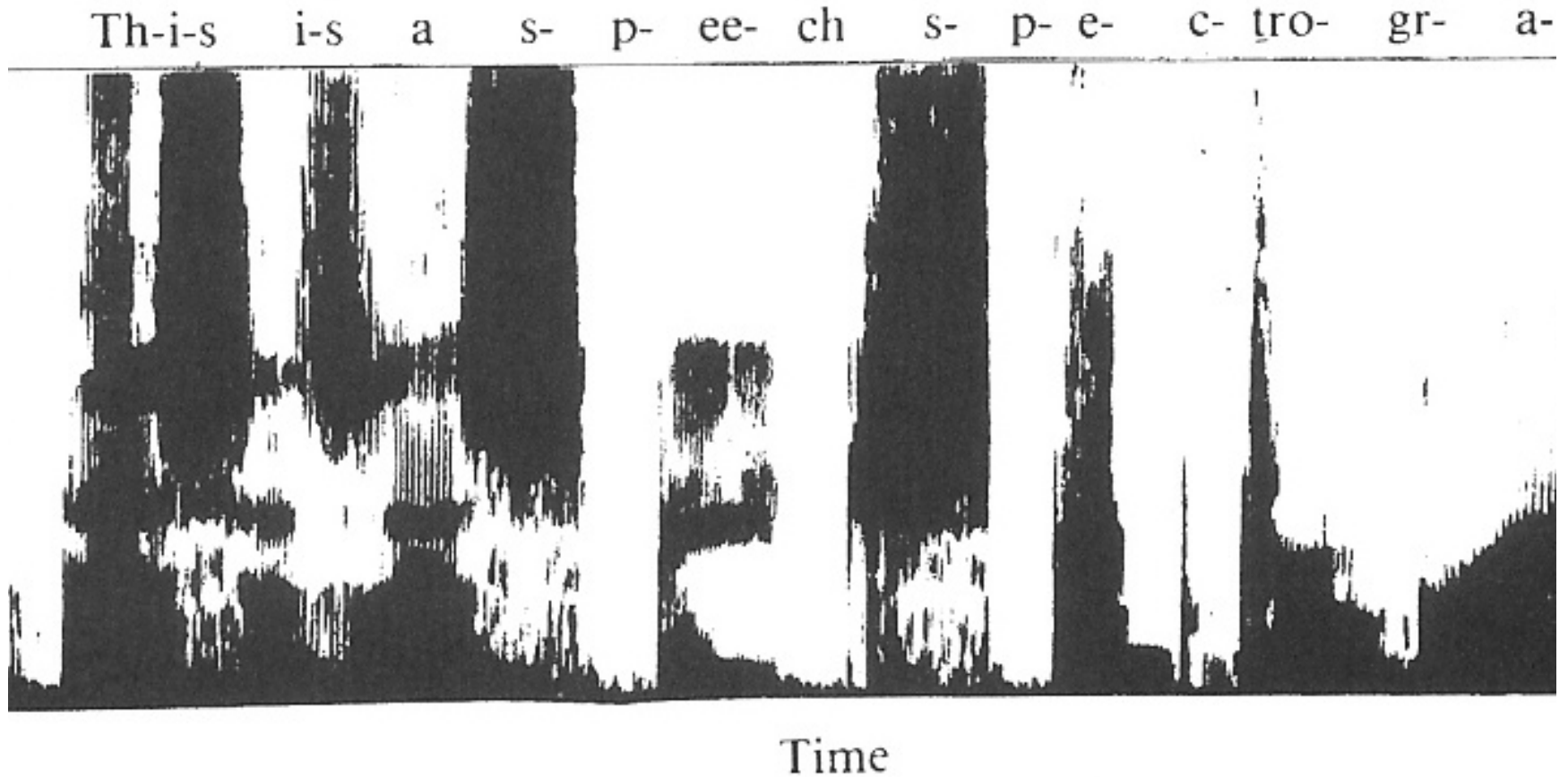
Vith Singer's Formant
(female voice)



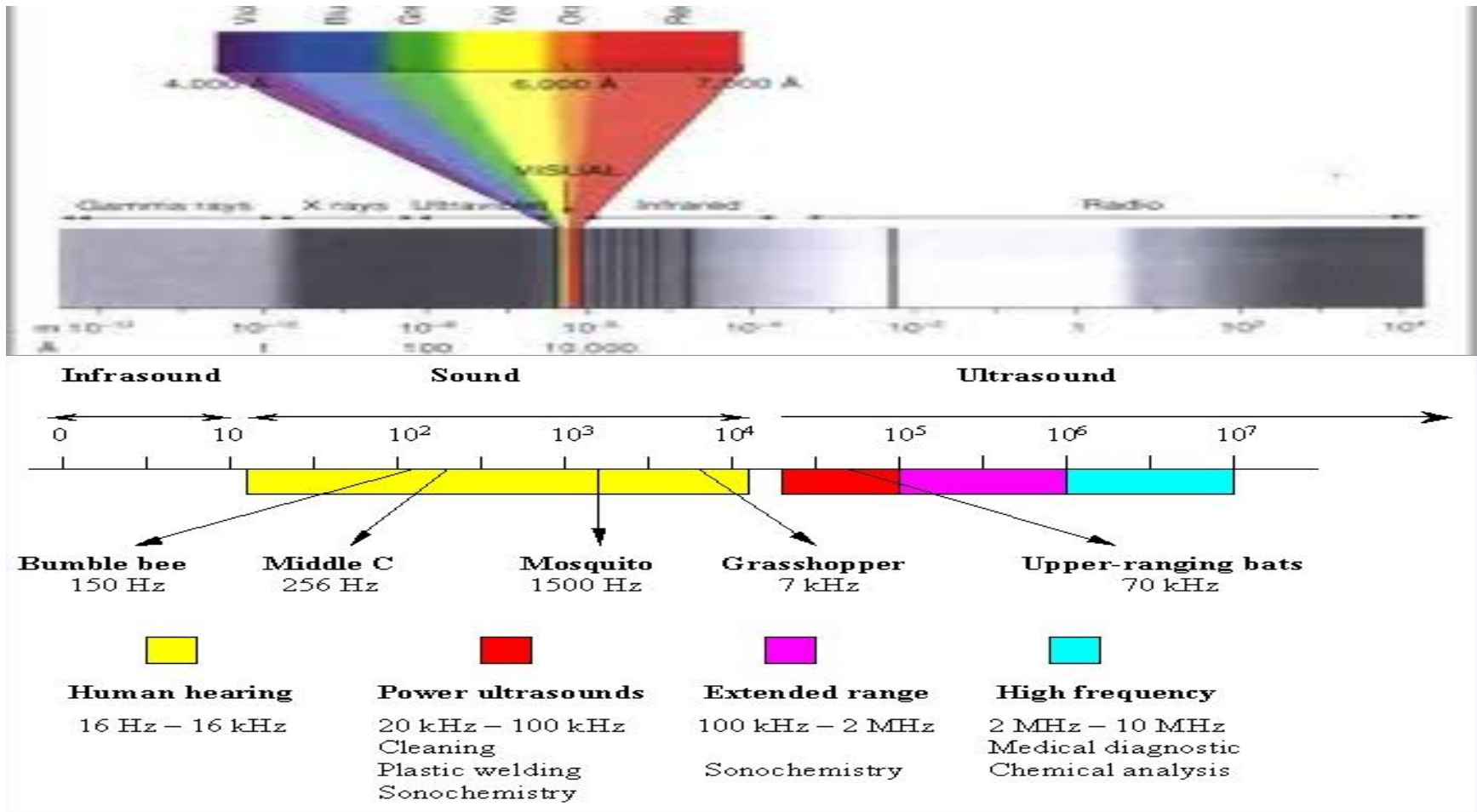
Without Singer's Formant
(female voice)



Φασματογράφημα (spectrogram) φράσης

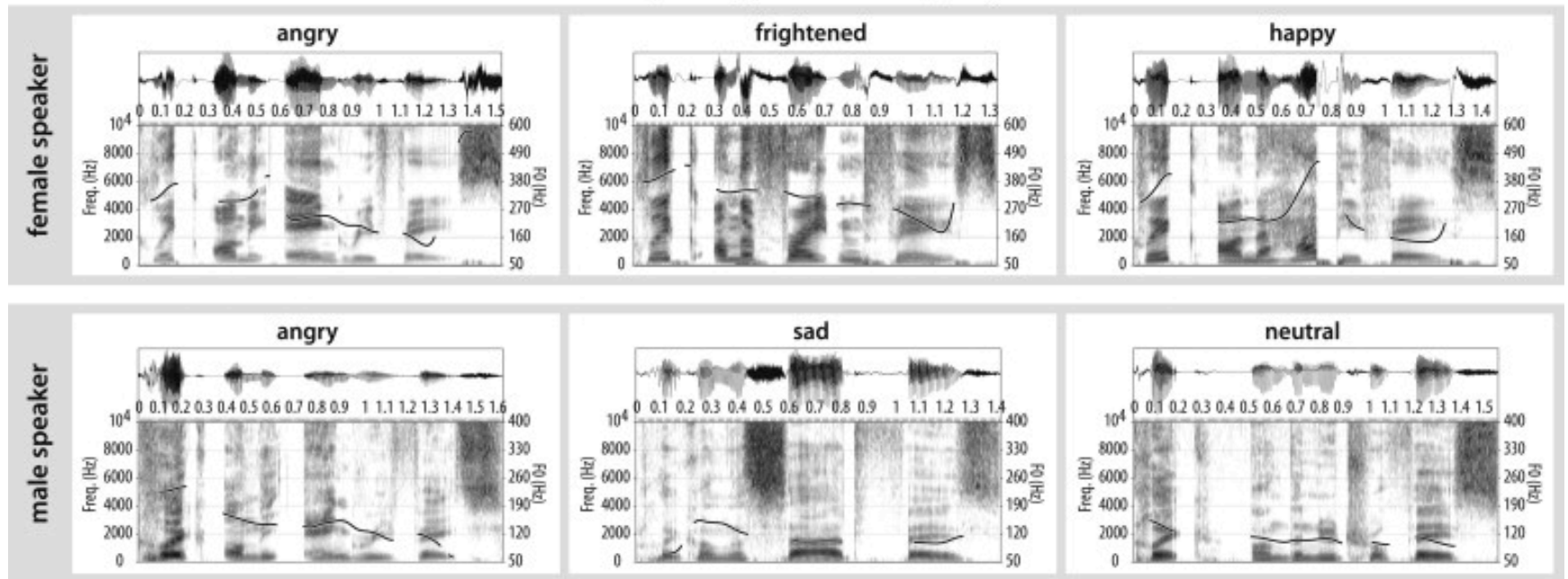


Ηχητικό φάσμα

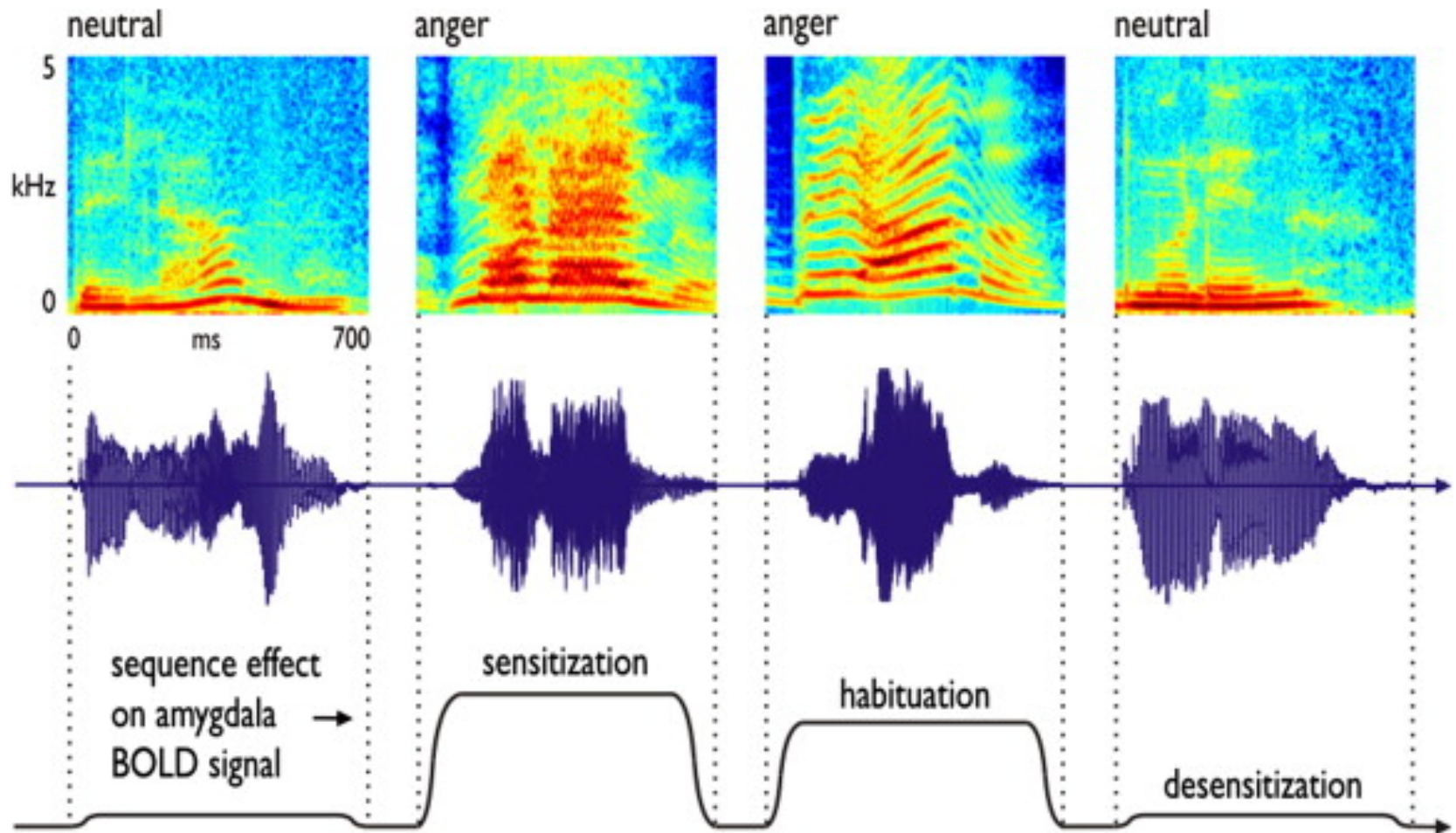


Voice and emotion

waveforms and spectrograms of exemplary utterances



Voice and emotion



Μουσικολογία+εθνομουσικολογία

Φασματογράφος εργαλείο

- Α) για την αισθητική αξιολόγηση της φωνής (ένρινη, πλούσια, βραχνή, λαρυγγώδης, κτλ)
- Β) επεξήγηση παραγωγής χαρακτηριστικών ηχοχρωμάτων της παράδοσης που χάνονται
- Γ) για την παιδαγωγική της φωνής (singer's formant, ακρίβεια τονικού ύψους, μίμηση χαρακτηριστικών φωνών, κ.α.)

Αποκωδικοποίηση

- 1. Εύρεση αρμονικού και μη αρμονικού περιεχομένου (θόρυβος)
- 2. Κατανόηση του ρυθμού και δομής των λεξεων και των φράσεων
- 2. Εύρεση της θεμελιώδου συχνότητας
- 3. Ανάλυση των θέσεων των αρμονικών συχνοτήτων (πλάτος, συχνότητα, περιβάλλουσα, φάση)
- 4. Εύρεση της συνολικής ενέργειας
- 5. Vibrato, jitter, shimmer

Table 6
Definition and Measurement of Acoustic Cues in Vocal Expression and Music Performance

Acoustic cues	Perceived correlate	Definition and measurement
Vocal expression		
Pitch		
Fundamental frequency (F0)	Pitch	F0 represents the rate at which the vocal folds open and close across the glottis. Acoustically, F0 is defined as the lowest periodic cycle component of the acoustic waveform, and it is extracted by computerized tracking algorithms (Scherer, 1982).
F0 contour	Intonation contour	The F0 contour is the sequence of F0 values across an utterance. Besides changes in pitch, the F0 contour also contains temporal information. The F0 contour is hard to operationalize, and most studies report only qualitative classifications (Cowie et al., 2001).
Jitter	Pitch perturbations	Jitter is small-scale perturbations in F0 related to rapid and random fluctuations of the time of the opening and closing of the vocal folds from one vocal cycle to the next. Extracted by computerized tracking algorithms (Scherer, 1989).
Intensity		
Intensity	Loudness of speech	Intensity is a measure of energy in the acoustic signal, and it reflects the effort required to produce the speech. Usually measured from the amplitude acoustic waveform. The standard unit used to quantify intensity is a logarithmic transform of the amplitude called the <i>decibel</i> (dB; Scherer, 1982).
Attack	Rapidity of voice onsets	The attack refers to the rise time or rate of rise of amplitude for voiced speech segments. It is usually measured from the amplitude acoustic waveform (Scherer, 1989).
Temporal aspects		
Speech rate	Velocity of speech	The rate can be measured as overall duration or as units per duration (e.g., words per min). It may include either complete utterances or only the voiced segments of speech (Scherer, 1982).
Pauses	Amount of silence in speech	Pauses are usually measured as number or duration of silences in the acoustic waveform (Scherer, 1982).
Voice quality		
High-frequency energy	Voice quality	High-frequency energy refers to the relative proportion of total acoustic energy above versus below a certain cut-off frequency (e.g., Scherer et al., 1991). As the amount of high-frequency energy in the spectrum increases, the voice sounds more sharp and less soft (Von Bismarck, 1974). It is obtained by measuring the long-term average spectrum, which is the distribution of energy over a range of frequencies, averaged over an extended time period.
Formant frequencies	Voice quality	Formant frequencies are frequency regions in which the amplitude of acoustic energy in the speech signal is high, reflecting natural resonances in the vocal tract. The first two formants largely determine vowel quality, whereas the higher formants may be speaker dependent (Laver, 1980). The mean frequency and the width of the spectral band containing significant formant energy are extracted from the acoustic waveform by computerized tracking algorithms (Scherer, 1989).
Precision of articulation	Articulatory effort	The vowel quality tends to move toward the formant structure of the neutral schwa vowel (e.g., as in <i>sofa</i>) under strong emotional arousal (Tolkmitt & Scherer, 1986). The precision of articulation can be measured as the deviation of the formant frequencies from the neutral formant frequencies.
Glottal waveform	Voice quality	The glottal flow waveform represents the time air is flowing between the vocal folds (abduction and adduction) and the time the glottis is closed for each vibrational cycle. The shape of the waveform helps to determine the loudness of the sound generated and its timbre. A jagged waveform represents sudden changes in airflow that produce more high frequencies than a soft waveform. The glottal waveform can be inferred from the acoustical signal using inverse filtering (Laukkanen et al., 1996).

Διαφορές μεταξύ φωνητικού σήματος στην ομιλία και τραγούδι

- Καταληπτότητα του μηνύματος
 - Διαφοροποίηση τονικών υψών, δυναμικών, ηχοχρωμάτων και διαρκειών
 - Εμφαση στο εύρος της δυναμικής
 - Εμφαση στο εύρος των συχνοτήτων
 - Χρήση του αέρα στα πνευμόνια (ζωτική ενέργεια)
 - Διάρκεια/επανάληψη φωνηέντων
 - Συνάρθρωση των συμφώνων με τα φωνήεντα ανάλογα με το είδος της μουσικής
 - Άλλη χρήση του αντηχείου (pitch extremes and passaggio)
 - Κατευθυντικότητα της φωνής (So if you cut the distance between you and the listener in half, the listener gets a 6 dB increase in SPL; if you double your distance from them, they lose 6 dB)
-
- ΤΕΧΝΑΣΜΑΤΑ ΤΩΝ ΤΡΑΓΟΥΔΣΤΙΩΝ
 - Αυξημένη υπογλωττιδική πείση
 - Διαμόρφωση του αντηχείου
 - Διαμόρφωση της προσαγωγής των φωνητικών χορδών * καλύτερη ενέργεια στους ψηλούς αμρονικούς)
 - Αρθρωση

3d sound spectrogram

Overtone Analyser

<http://www.youtube.com/watch?v=r3kLKi6MhPo>

