universidade de aveiro
theoria poiesis praxis

inet[MD]
instituto de etnomusicologia

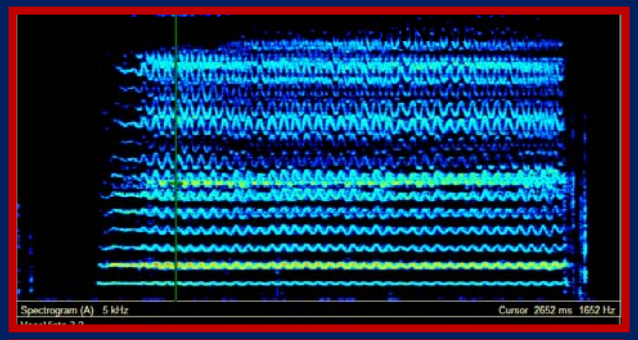
EVTA SESSION HELSINKI

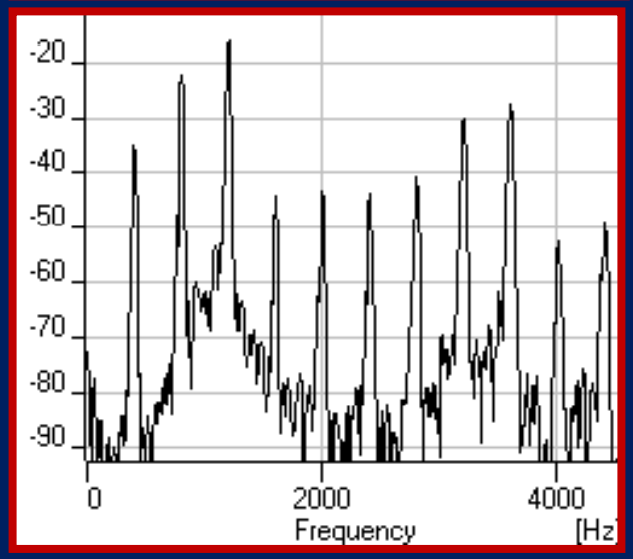
JUNE 06 – 10, 2012

evta
european voice teachers association

Education and Culture
Lifelong Learning Programme

Reading Spectrograms





Filipa Lã (PhD)

Department of Communication and Arts

University of Aveiro, INET-MD - Portugal

Department of Communication and Arts

University of Aveiro
Campus Universitário de Santiago
3810-193 Aveiro

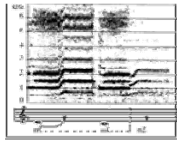
Portugal

filipa.la@ua.pt

What information does a spectrum display reveals?

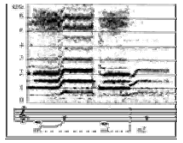
SPECTRUM DISPLAY

UNDERLYING INFORMATION

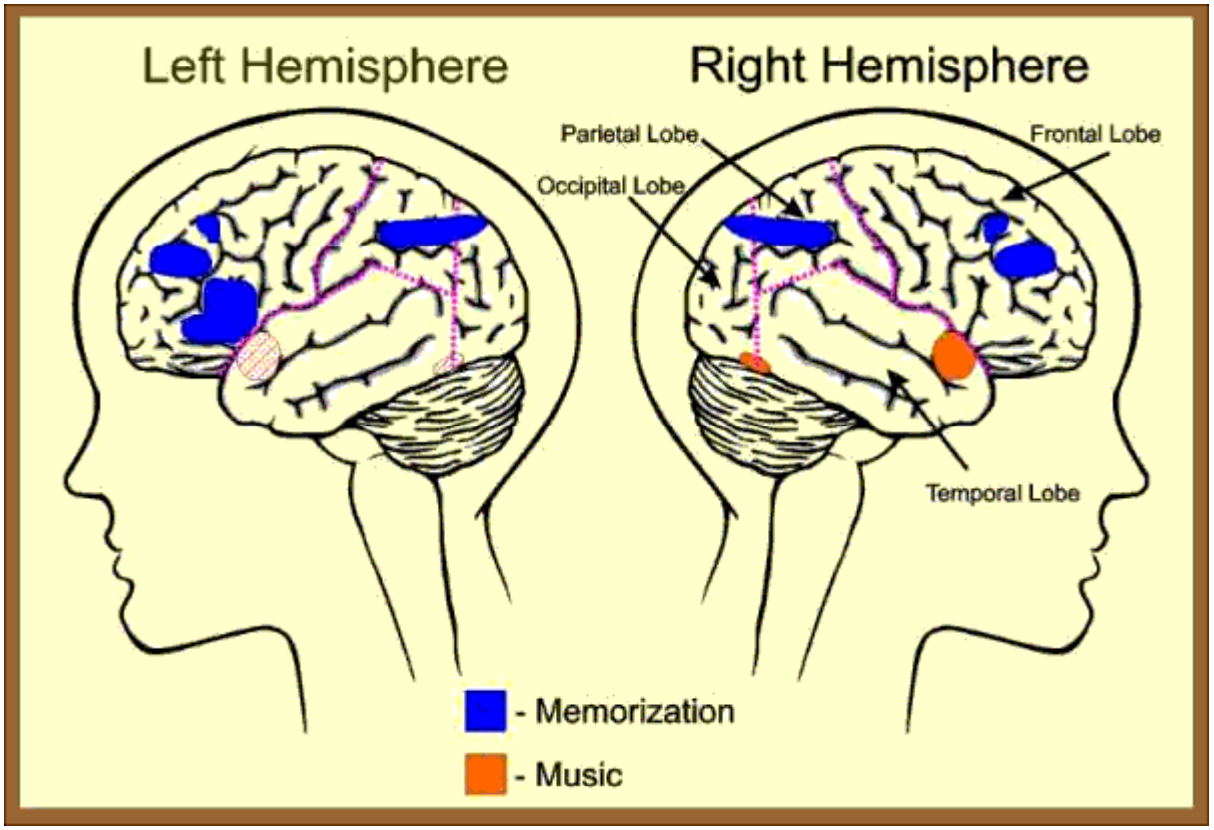


❑ Any musical sound is characterised by four interrelated properties

- ❖ Duration
- ❖ Frequency (objective measure of pitch)
- ❖ Amplitude (objective measure of loudness)
- ❖ Spectral envelope (objective measure of timbre)



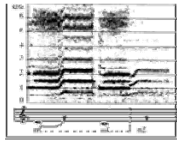
- **Duration**: the periodicity of a musical sound needs adequate duration to be perceived by the brain as pitch



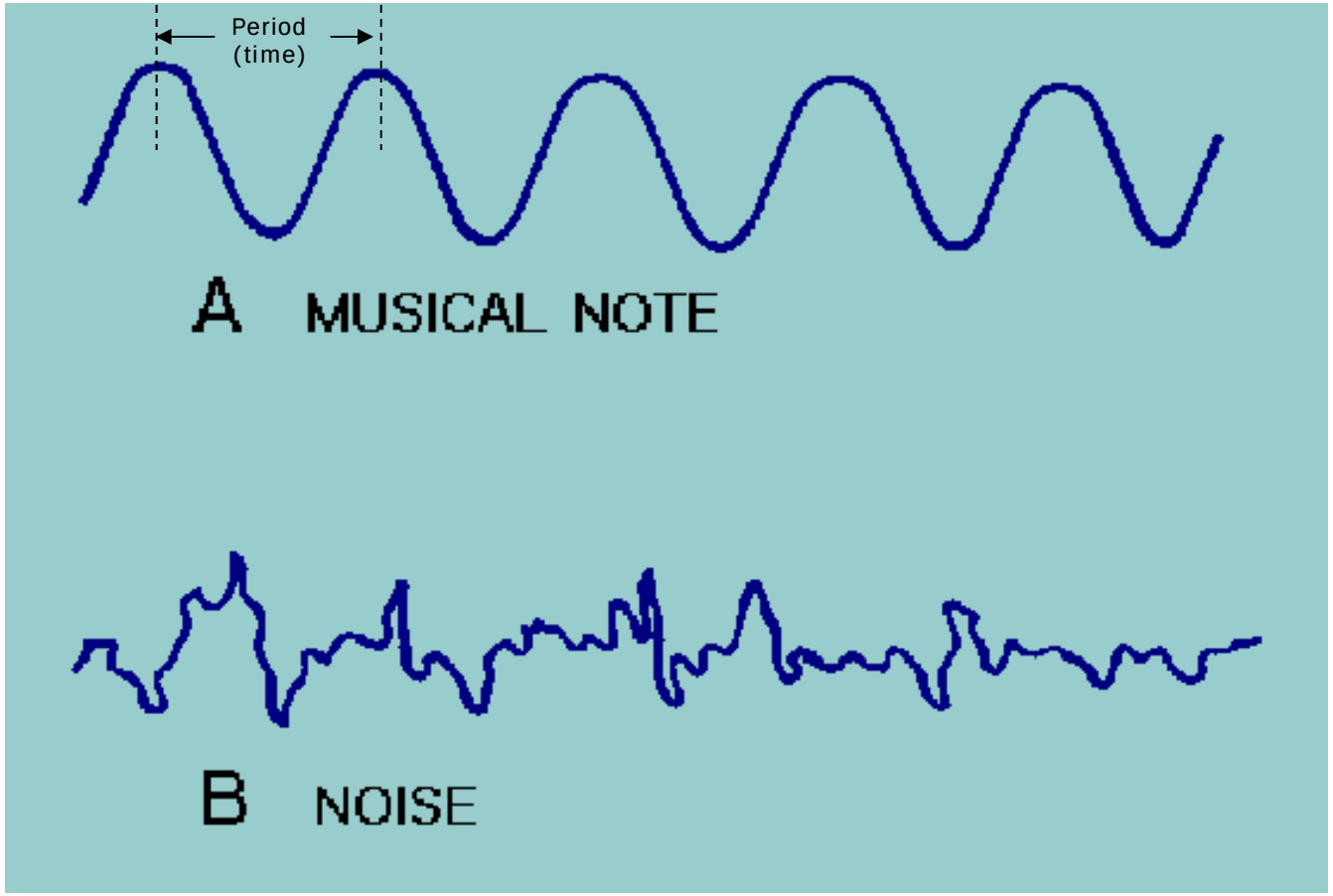
Brain areas invlolced in music perception and memorization [adapted from http://images.google.pt/imgres?imgurl, in 02/05/2010]

SPECTRUM DISPLAY

UNDERLYING INFORMATION

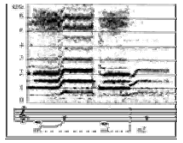


- **Frequency**: resulting from periodicity; distinguishes noise from musical sounds

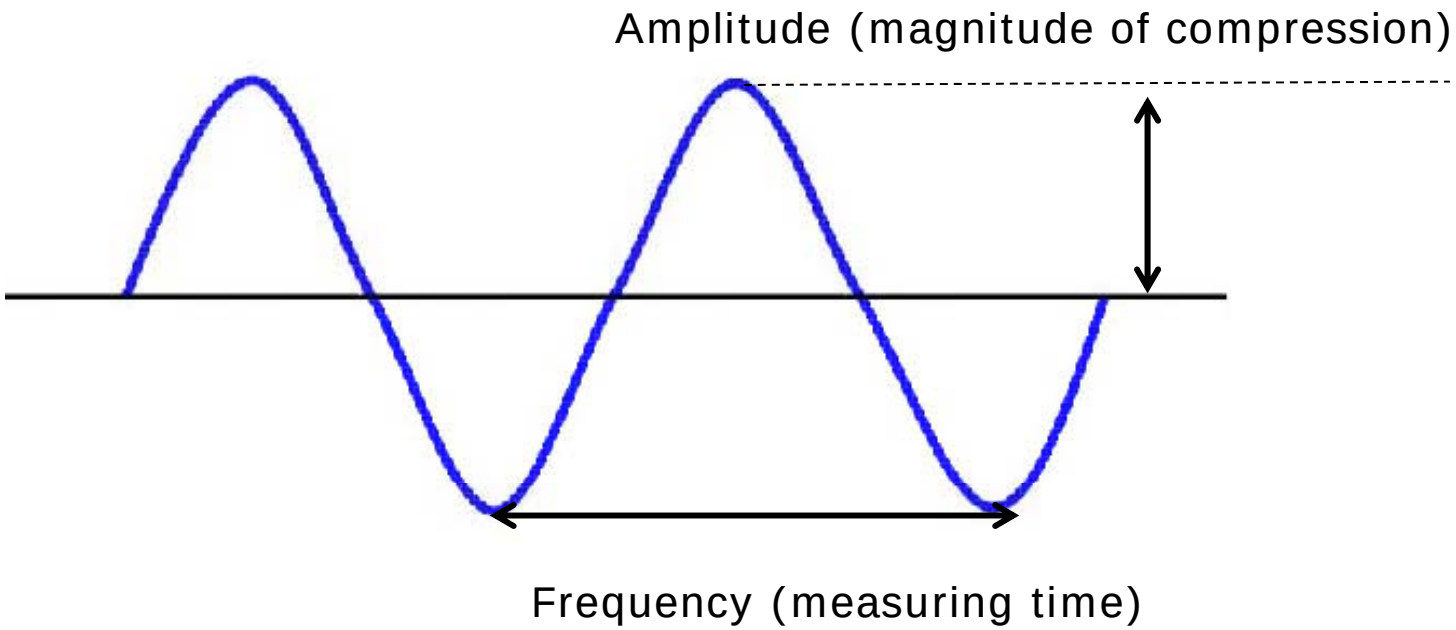


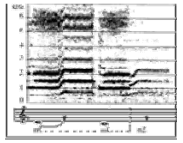
SPECTRUM DISPLAY

UNDERLYING INFORMATION

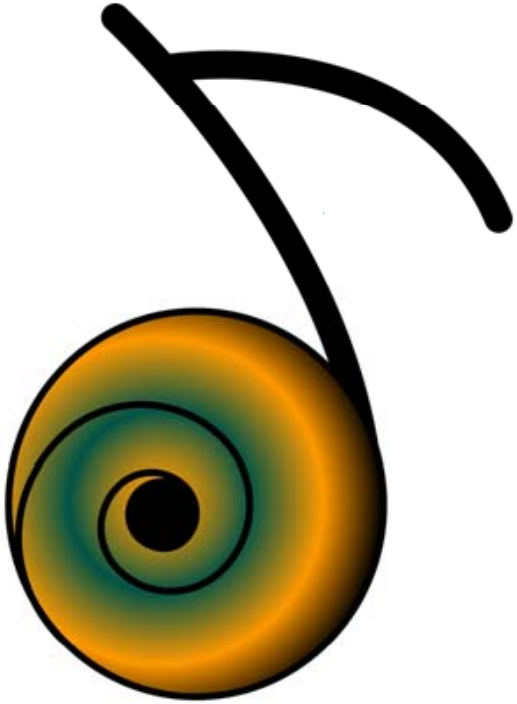


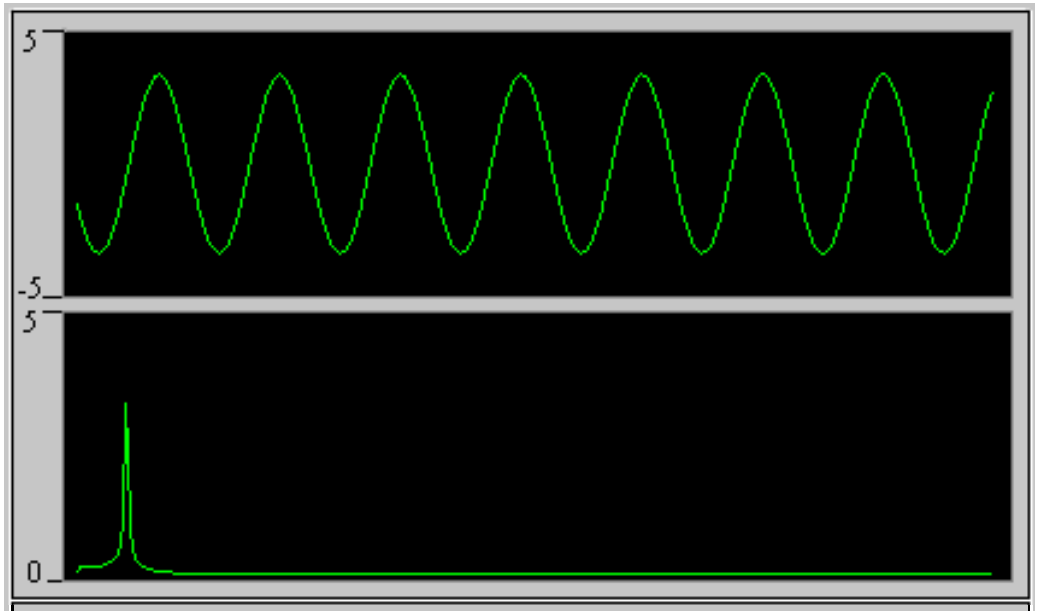
- **Amplitude**: magnitude of the compressions within a sound wave





- **Spectral envelope**: assessing the timbre (i.e. tone color) of a given sound

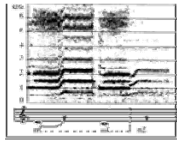




Pure tone (adapted from McCoy, 2004)

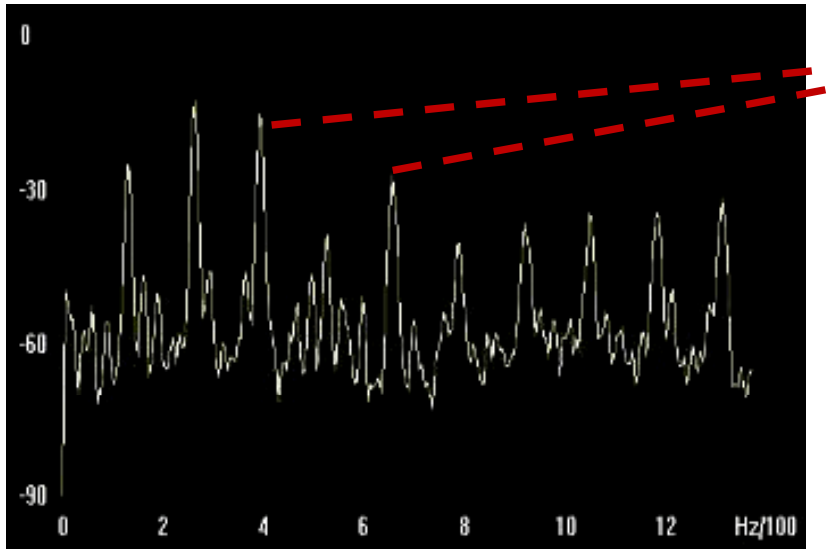
SPECTRUM DISPLAY

UNDERLYING INFORMATION



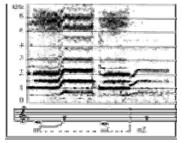
READING SPECTROGRAMS Filipa Lã, 2012

- **Musical sounds**: result from many simultaneously-occurring, interrelated frequencies at different amplitudes



Harmonics = additional vibratory frequencies that are *whole-number multiples of F0*

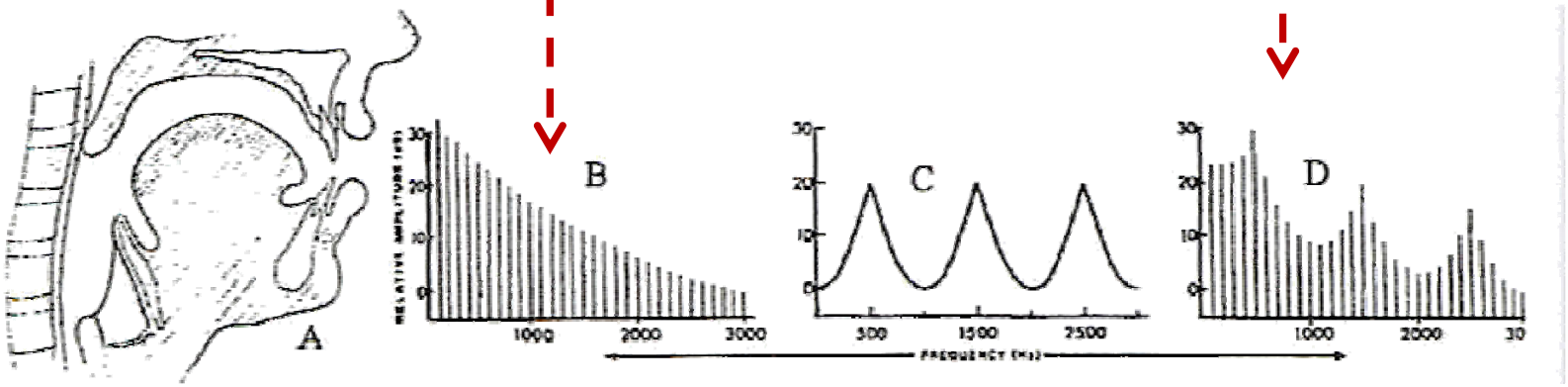
Spectral envelop of a cello (power-spectrum) (adapted from McCoy, 2004)



❑ The human voice: source/ filter theory

The glottal sound, before being filtered by the vocal tract resonances, have harmonics which diminish in amplitude as they increase in frequency

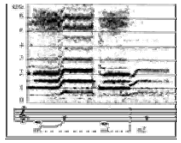
Harmonics do not diminish at a constant rate as they increase in frequency: some are strongly amplified whereas others are dampened



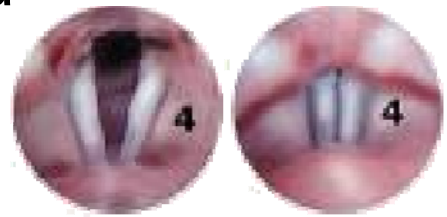
SOURCE FILTER THEORY: Vocal tract subareas (A); Primary sound spectrum (B) modified by the resonances of the vocal tract (C). The result is a radiated sound spectrum that has certain partials enhanced and other dampened (adapted from Urrutia & González, 1996: 80)

SPECTRUM DISPLAY

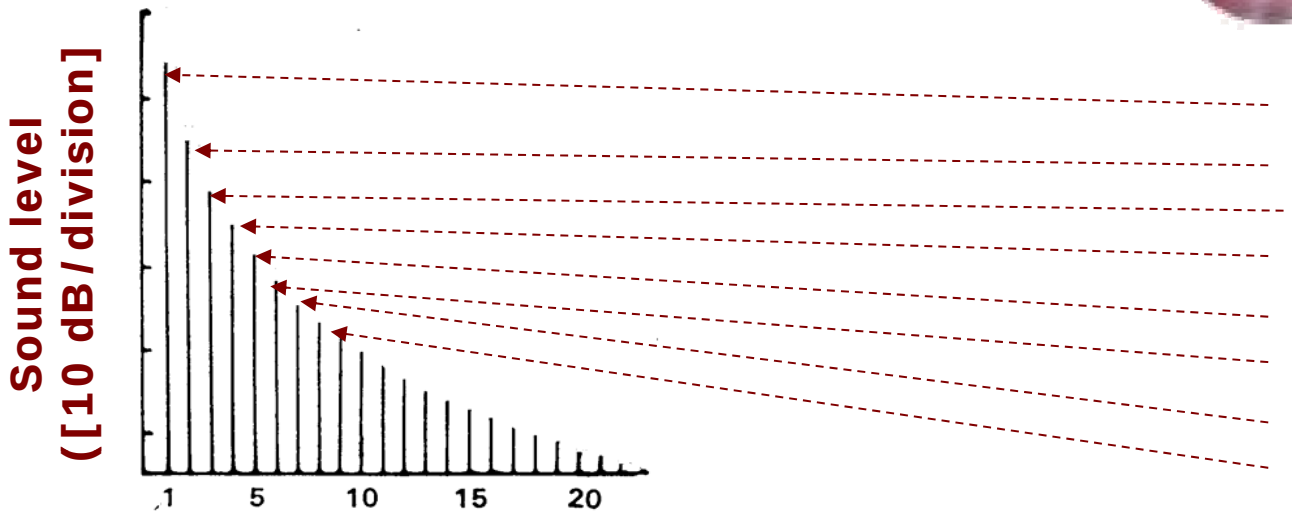
UNDERLYING INFORMATION

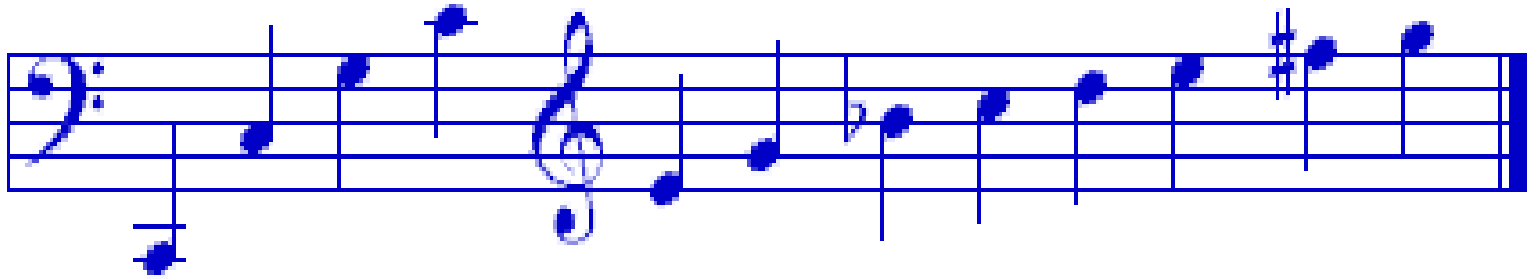


❑ The human voice: primary sound



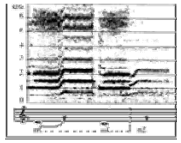
VOICE SOURCE SPECTRUM





SPECTRUM DISPLAY

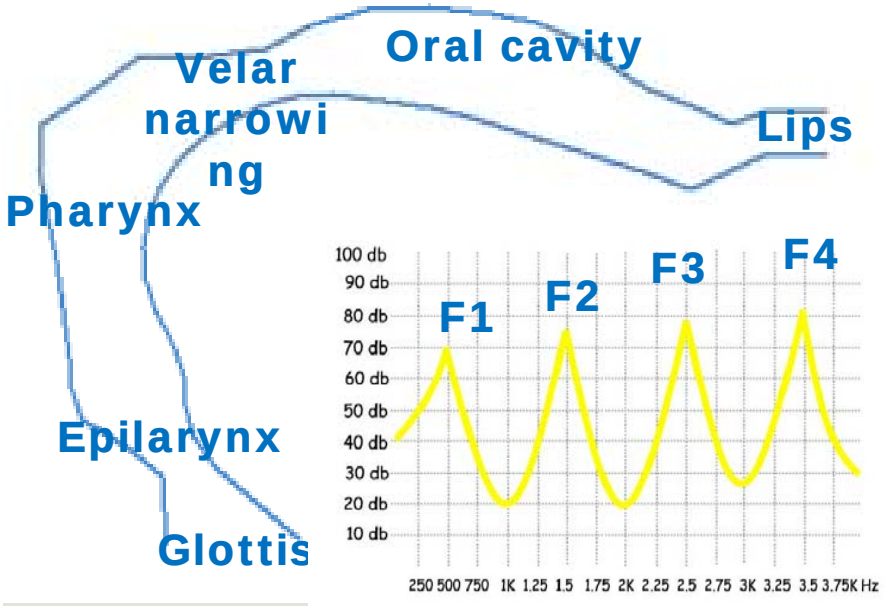
UNDERLYING INFORMATION

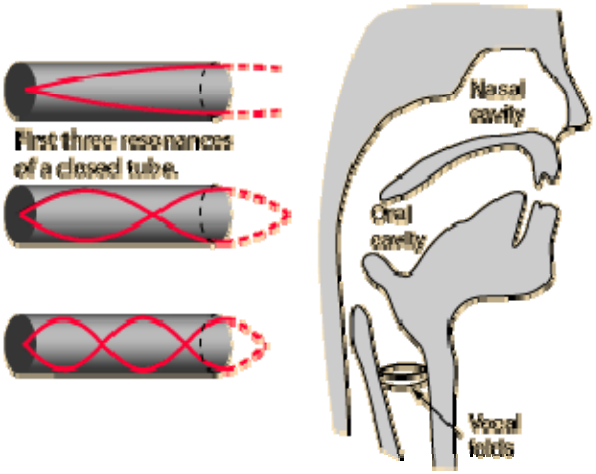


READING SPECTROGRAMS

Filipa Lã, 2012

❑ The human voice: modified sound





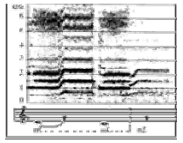
Subareas of the vocal tract

A formant is a resonance of the vocal tract

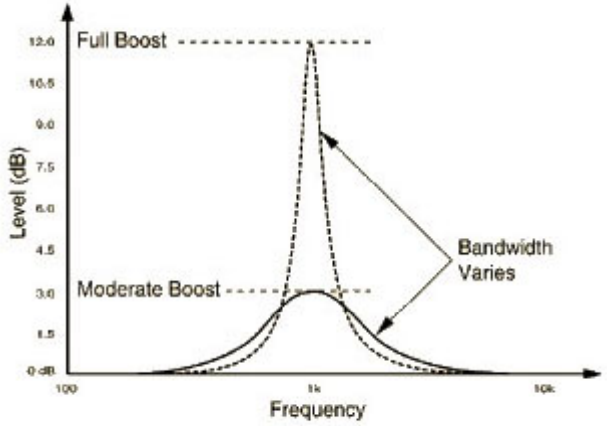
Vocal tract – A resonator with its own resonance frequencies (formants)

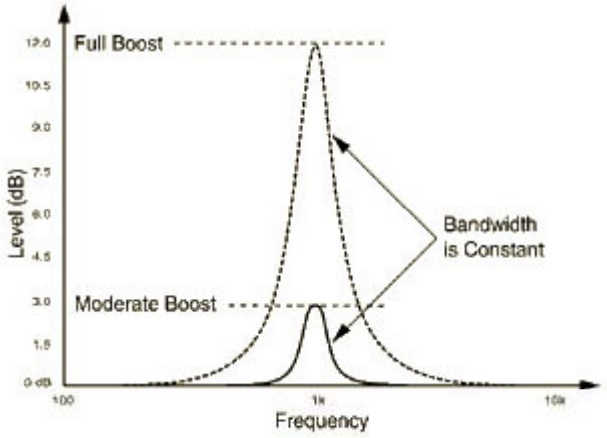
F1,2,3,4, 5 = (V/L) x ¼ (V/L) x ¾ (V/L) x 5/4 (V/L) x 7/4 (V/L)

V = 350 ms (sound velocity)
L = tube length



❑ **Formants of the human voice:** equivalent to the amplifier and tone controls in a stereo





Formant's frequency peak (adapted from McCoy, 2004: 41)

❖ Frequency

✓ Measured at the centre, top of the peak

✓ Provide vowel accuracy (F1 and F2)

✓ Provide individual timbre and may assist in vocal projection (F3, F4 and F5)

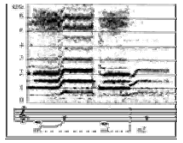
❖ Bandwith

✓ Measures the with of the peak at a specific distance from the top (eg. 10dB)

✓ An harmonic near the vicinity of a formant will be amplified

SPECTRUM DISPLAY

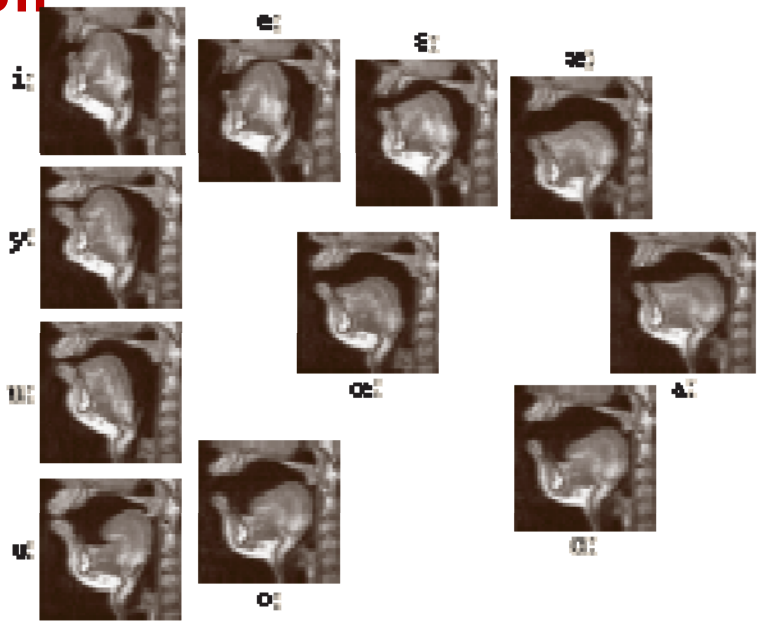
UNDERLYING INFORMATION



- **The human voice:** unique musical instrumental in which articulation also affects resonance

Formant frequencies change when altering the shape of the vocal tract through articulation

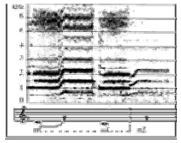
- A constriction in the front of the vocal tract (e.g. palatal moving of the tongue) lowers F1 and raises F2
- A constriction in the back of the vocal tract (e.g. pharyngeal movement of the tongue) raises F1 and lowers F2



Tongue position captured by magnetic resonance for different vowels showing the effects on vocal tract constrictions (adapted from Lindblom & Sundberg, 2007: 683)

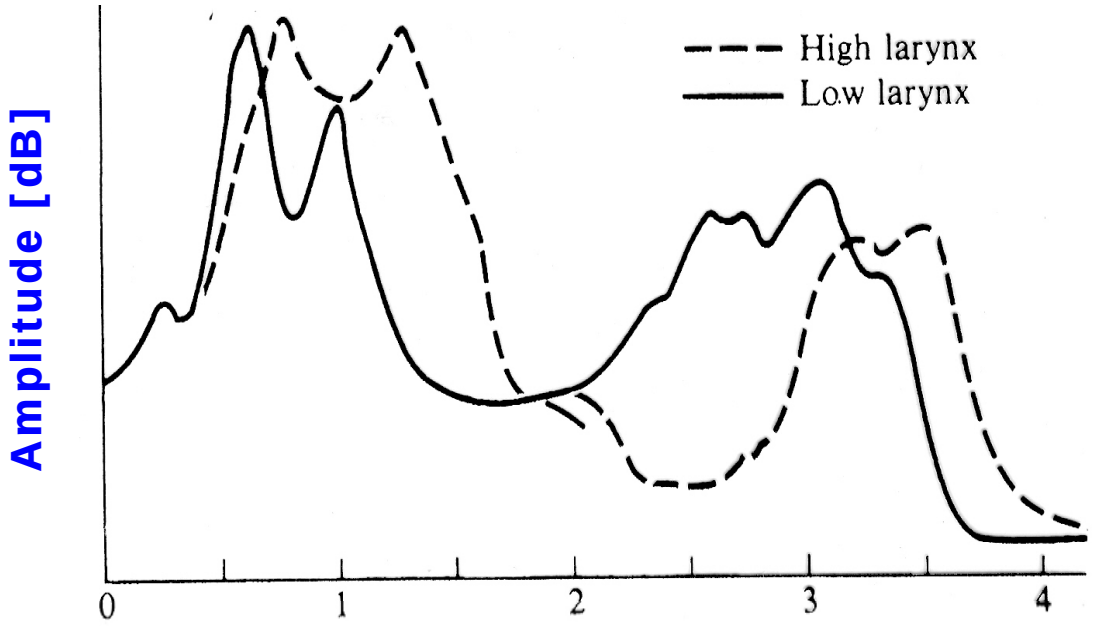
SPECTRUM DISPLAY

UNDERLYING INFORMATION



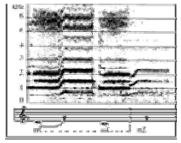
- All formant frequencies lower uniformly when the vocal tract is lengthened (e.g. lowering the larynx and protruding the lips)
- All formant frequencies rise uniformly when the vocal tract is shortened (e.g. raising the larynx and spreading the lips)

Vowel /a/ spoken and sung

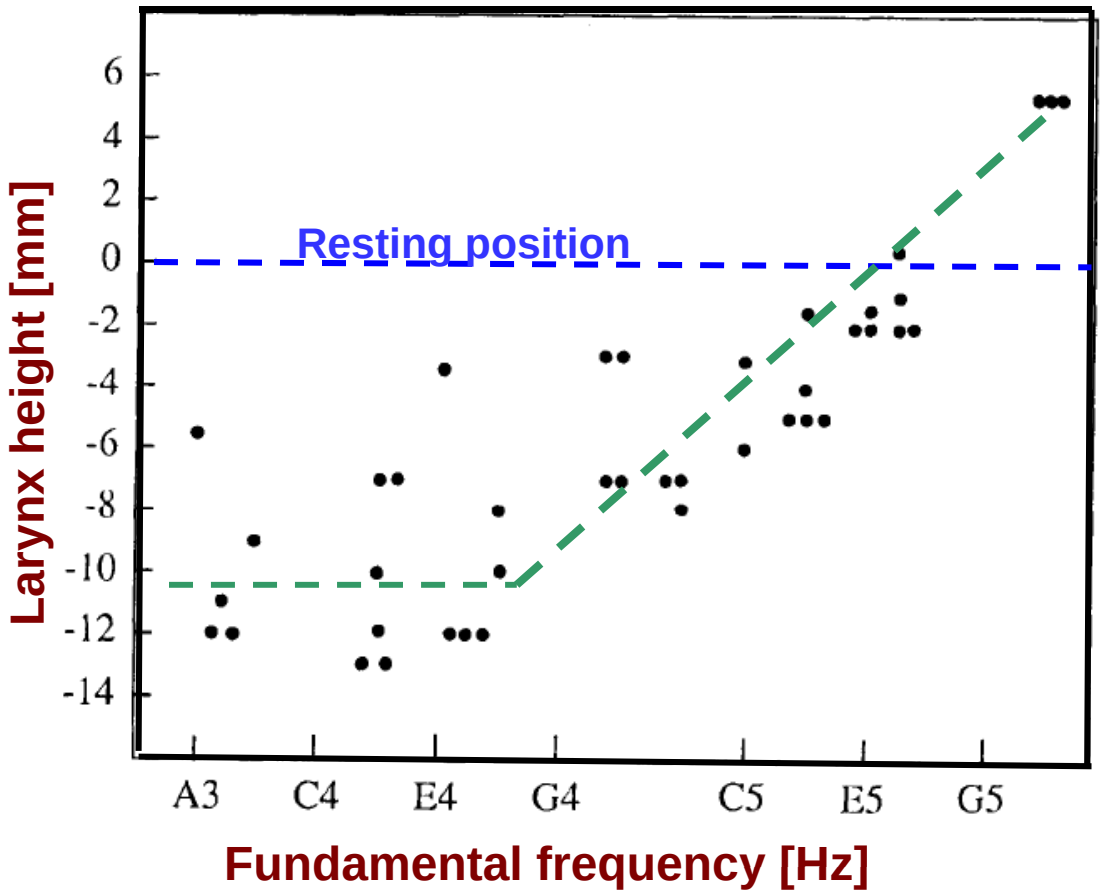


SPECTRUM DISPLAY

UNDERLYING INFORMATION



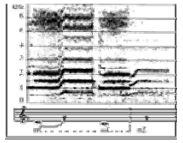
❖ Larynx height increases with increasing pitch for a soprano



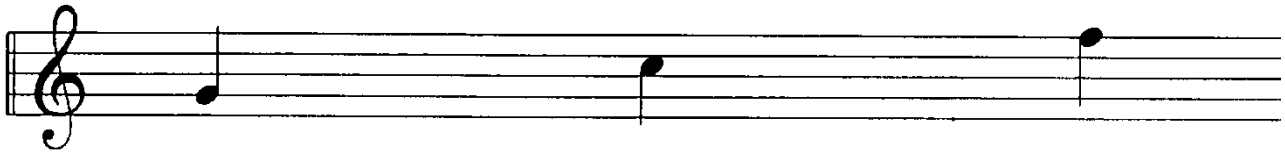
Change in larynx position with raising F0 (adapted from Sundberg, 1987)

SPECTRUM DISPLAY

UNDERLYING INFORMATION



- Lowering the jaw (opening the mouth) increases F1
- This is needed to avoid F0 > F1 (which causes instabilities in the vibration of the vocal folds)



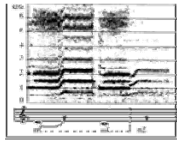
Vowel [i]

Vowel [u]

Visualisation of jaw opening for vowels /i/ and /u/ with raising F0 (adapted from Sundberg, 1987)

SPECTRUM DISPLAY

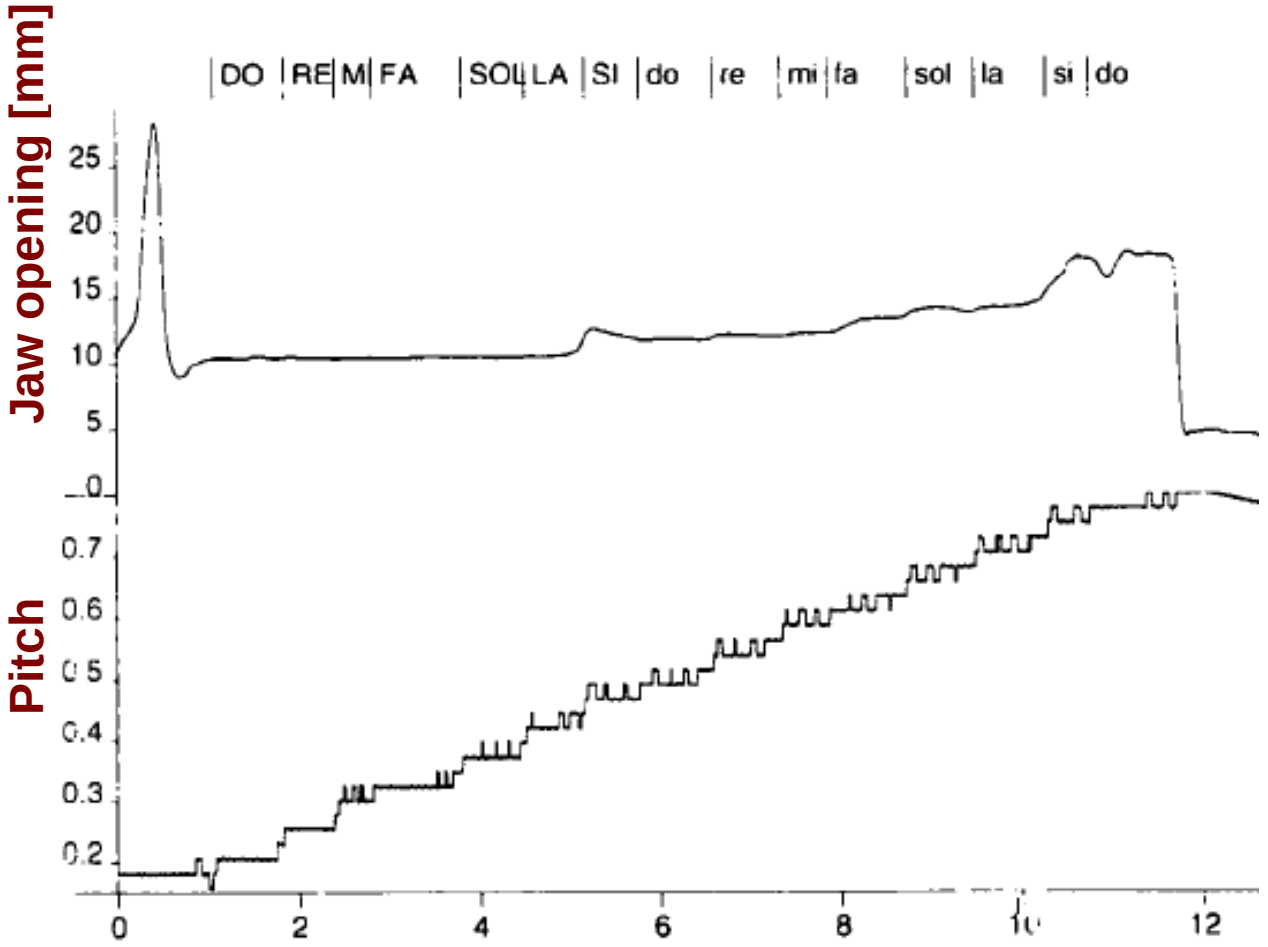
UNDERLYING INFORMATION



READING SPECTROGRAMS

Filipa Lã, 2012

Jaw opening during two octave scale

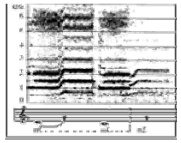


Jaw opening during a two octave scale singing

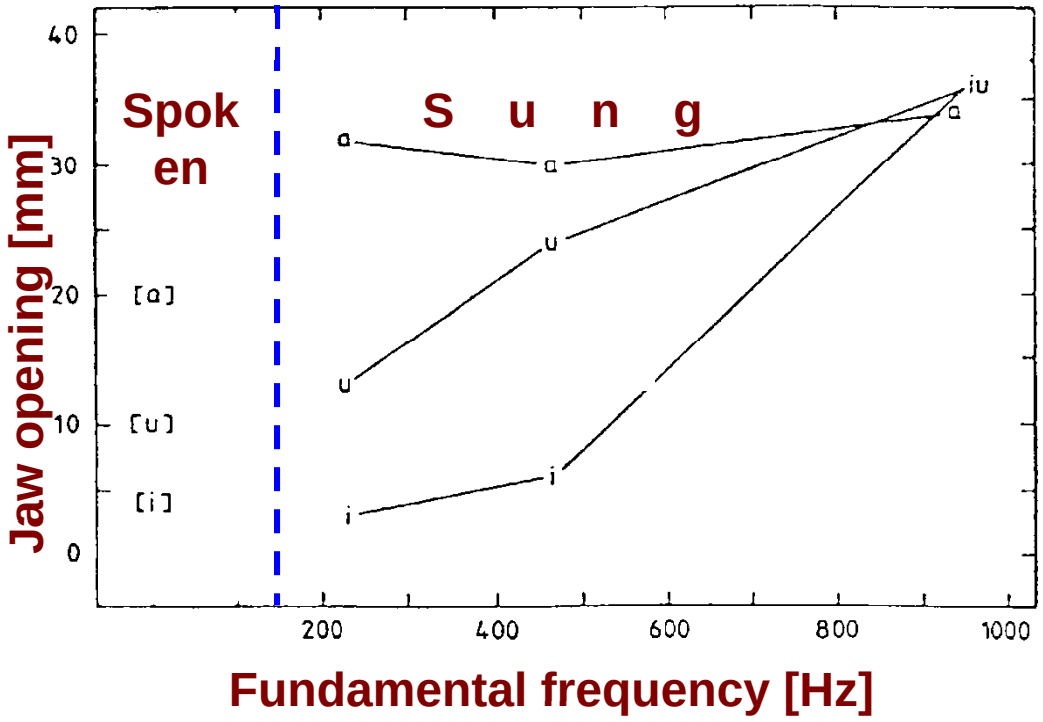
(Courtesy from Sundberg, 2010: Summer course)

SPECTRUM DISPLAY

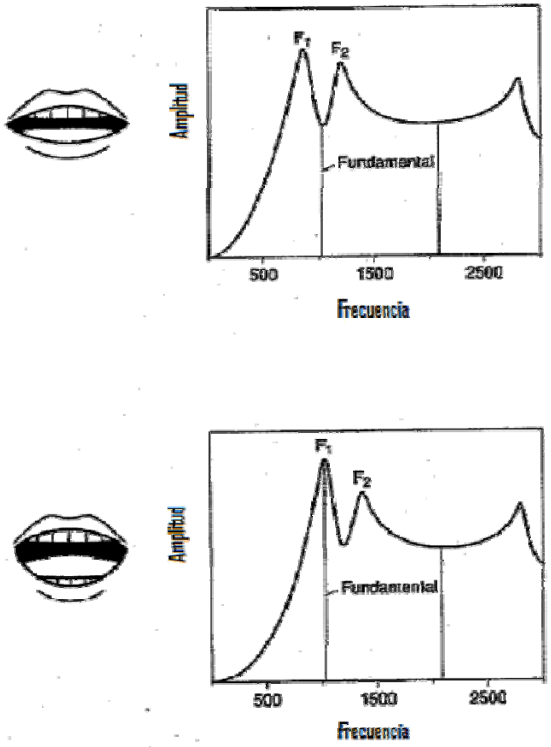
UNDERLYING INFORMATION



- High voices need to open their jaw as F0 increase



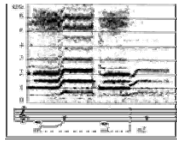
Jaw opening with increasing F0 (Courtesy from Sundberg, 2010: Summer course)



Opening of the jaw constricts the pharynx and thus raises F1 to track F0 (adapted from Thurman & Welch, 2000: 481)

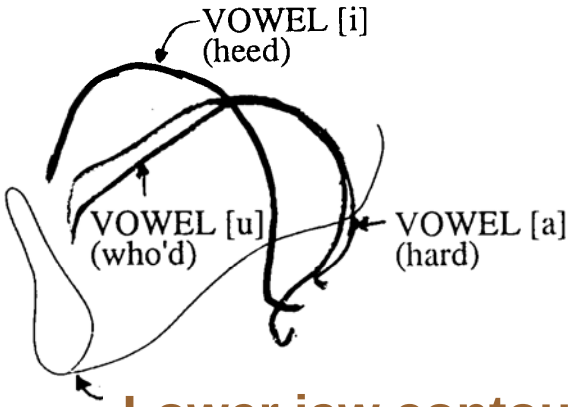
SPECTRUM DISPLAY

UNDERLYING INFORMATION



Tongue shape also varied

Spoken



Lower jaw contour

Sung

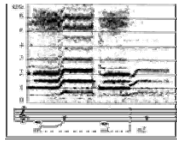
Fundamental frequency [Hz]

230 *465* *940*

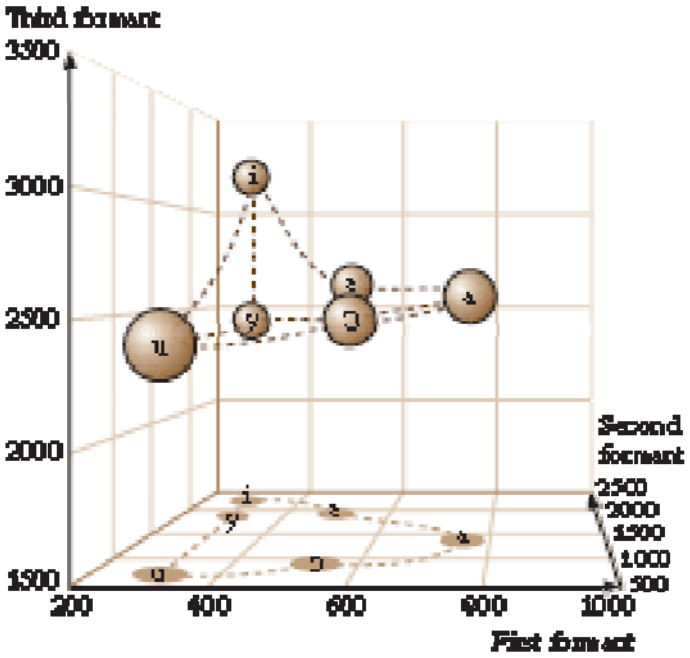
Jaw opening with increasing F0 (Courtesy from Sundberg, 2010: Summer course)

SPECTRUM DISPLAY

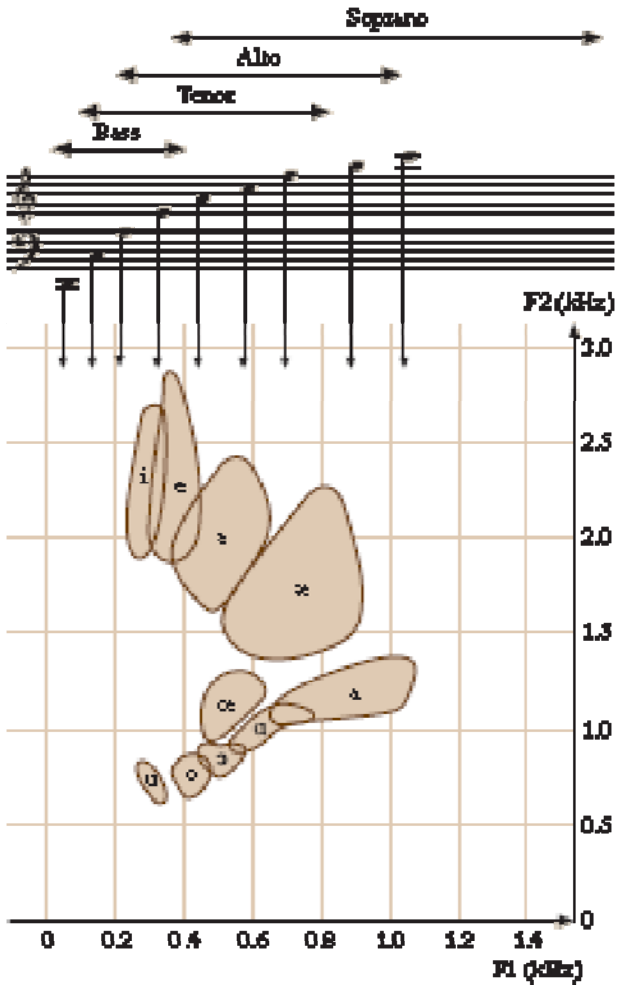
UNDERLYING INFORMATION



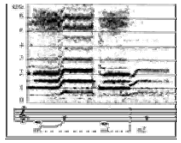
- The consequence is vowel modification!
  - Vowel accuracy is defined by F1 and F2



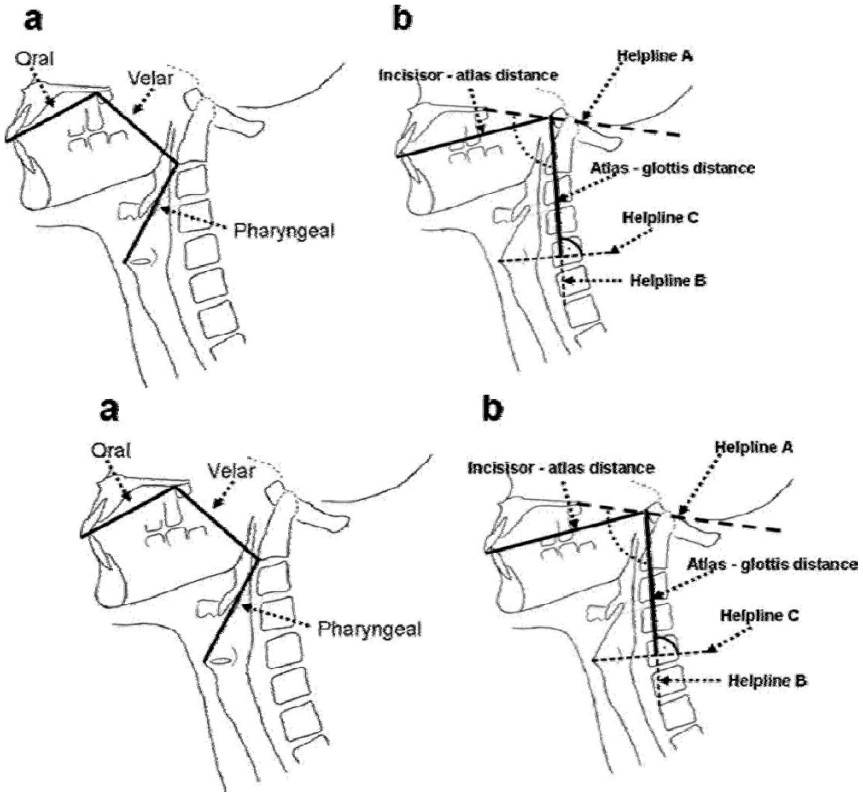
Vowel F1, F2 and F3 according with APEX model (adapted from Lindblom & Sundberg, 2007: 686)

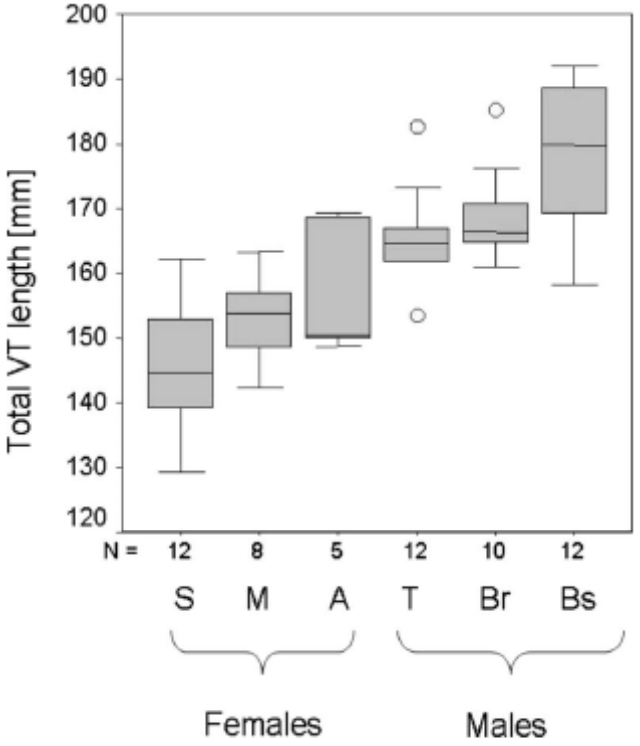


Vowel chart and corresponding vocal ranges of different singers (adapted from Lindblom & Sundberg, 2007: 686)



❑ Female and male vocal tract and vocal folds length

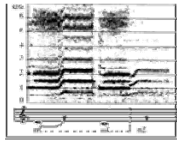




Different vocal tract lenghts (left) and vocal folds length (right)
(adapted from Roers, Mürbe & Sundberg, 2008)

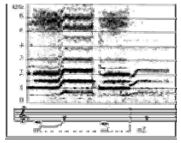
VoceVista 3.lnk

Spectrum displays



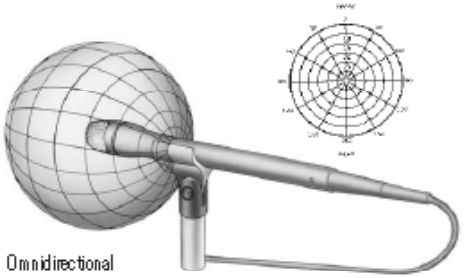
❑ Fast Fourier Transform

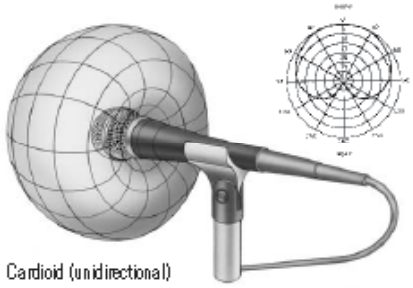
- ❖ mathematical formula which converts sound into its component parts
- ❖ it is used to perform:
  - ✓ power spectrum
  - ✓ spectrogram
- ❖ data collected by means of a microphone is:
  - ✓ converted from an electric sound to a digital signal (computer sound card or external device)
  - ✓ the signal is processed applying FFT analysis
  - ✓ results are plotted in a graph displayed on the computer screen

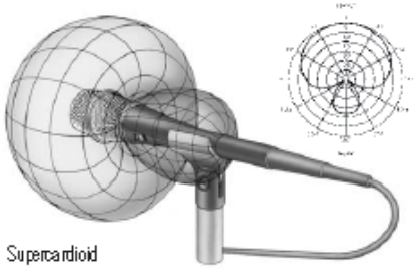


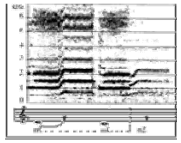
❑ Effects of microphone placement

- ❖ placing the microphone near to the mouth increases energy of all partials
- ❖ the effect we see is not related with the student's changing behaviours
- ❖ best to measure distance to the mouth with the student and use the same distance always with that student
- ❖ best to use omnidirectional microphones with a flat response over a wide range of frequency band (e.g. Audio-Technica ATR-3350 Lavalier Omnidirectional Condenser Microphone)





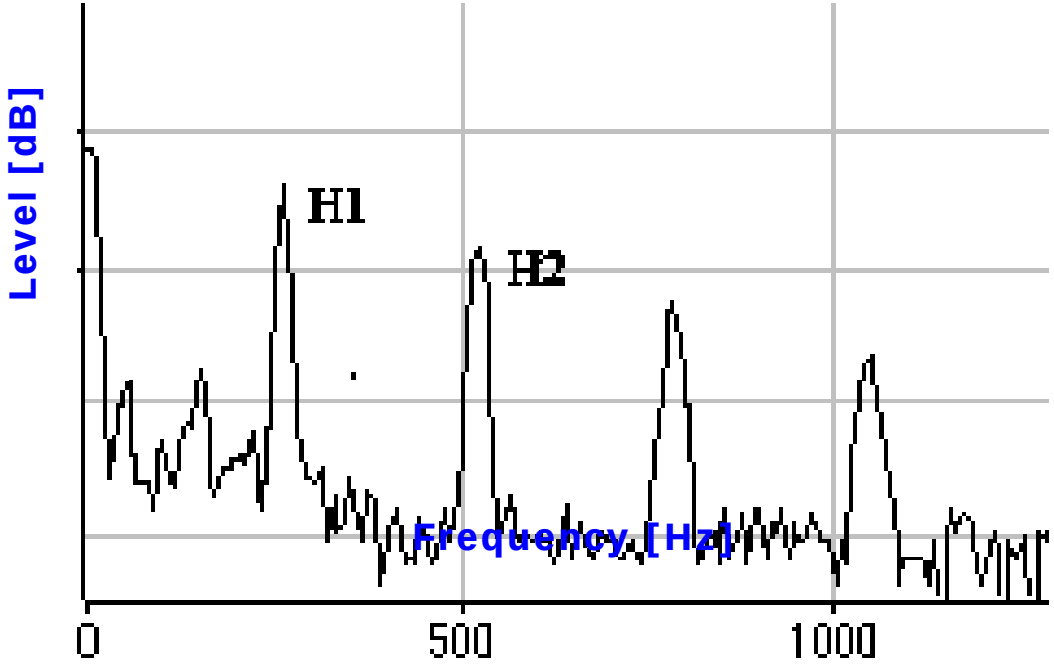




Power Spectrum

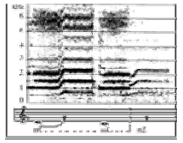
Amplitude of the F0 and partials on the vertical axis
Frequency along vertical axis

- It shows sound components during a single moment in time (some ms)



REAL-TIME FEEDBACK

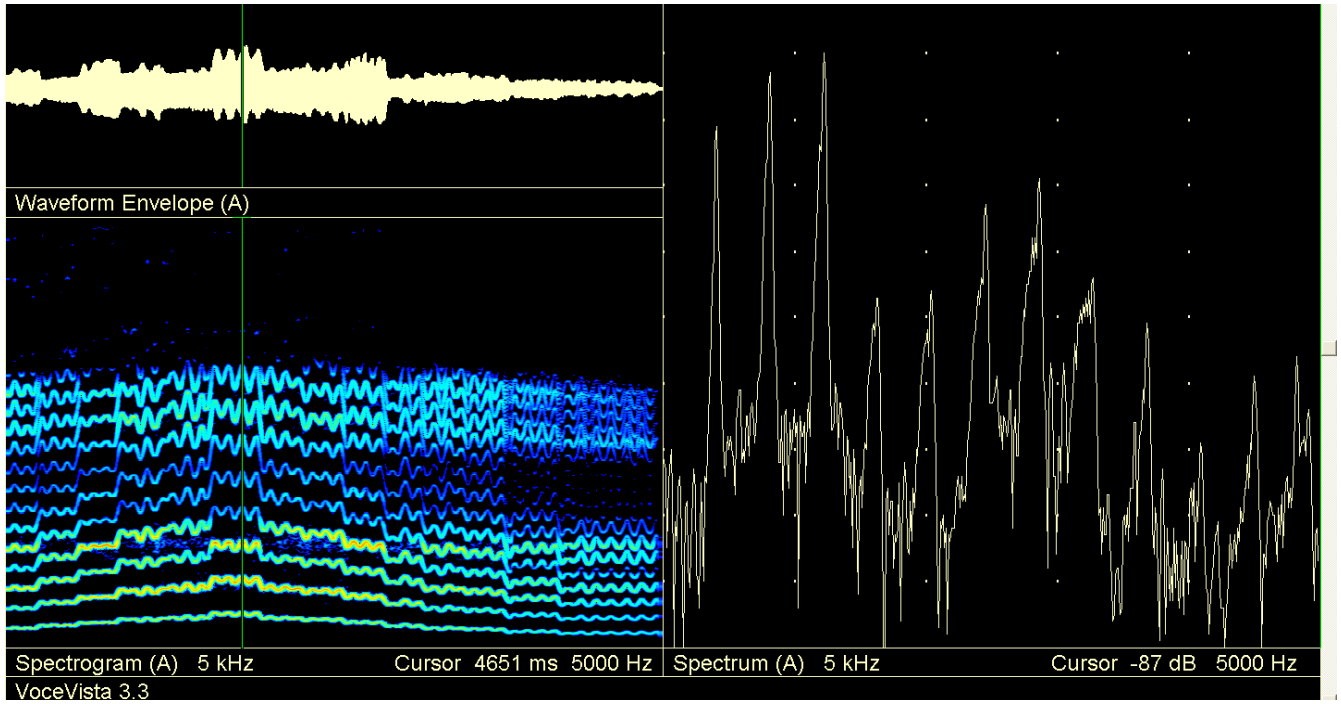
APPLICATIONS IN SINGING LESSONS



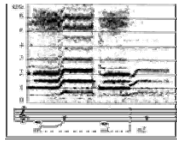
READING SPECTROGRAMS

Filipa Lã, 2012

- useful to display formant tuning strategies, although it cannot show location of formants



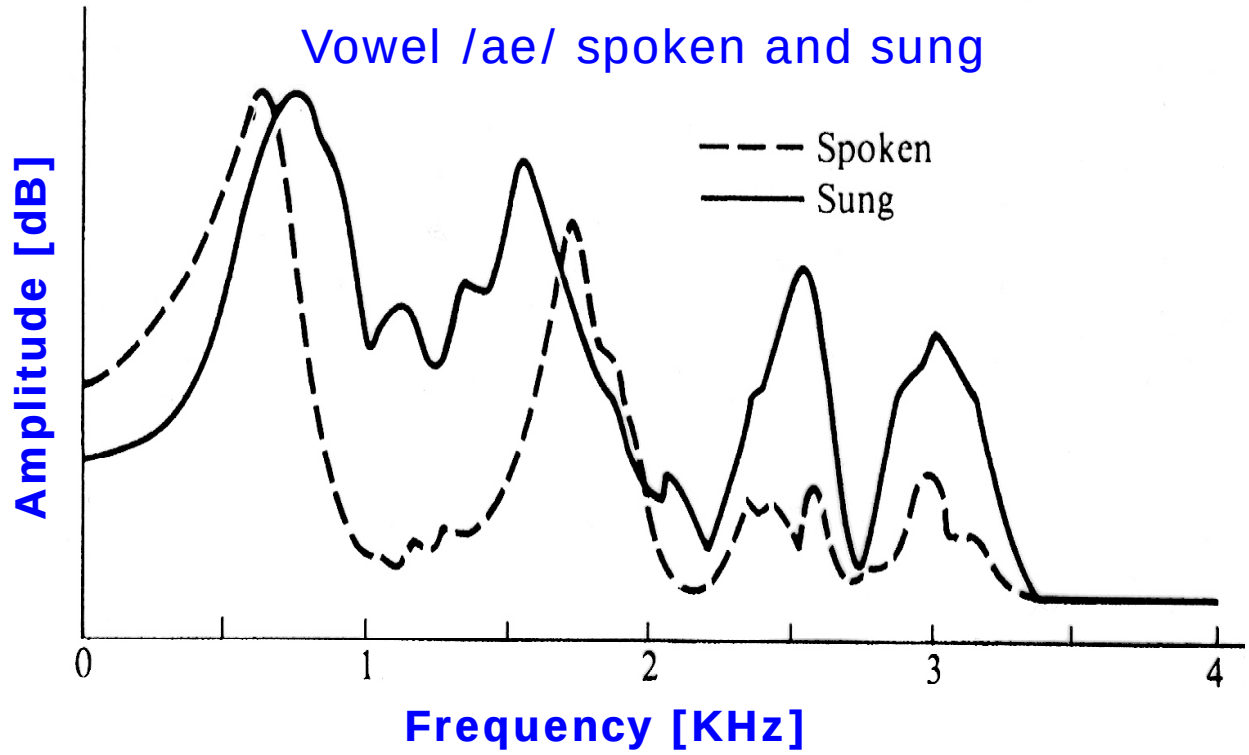
Spectrogram of a baritone voice singing a scale up to G4, vowel [a], with power spectrum at G4 showing different harmonic partials (using Don Miller's VoceVista program, version 3.3)



Long-Term Average Spectrum (LTAS)

Amplitude of the F0 and partials on the vertical axis
Frequency along vertical axis

- It shows sound components during several cycles of vibration; it is therefore an average acoustic output

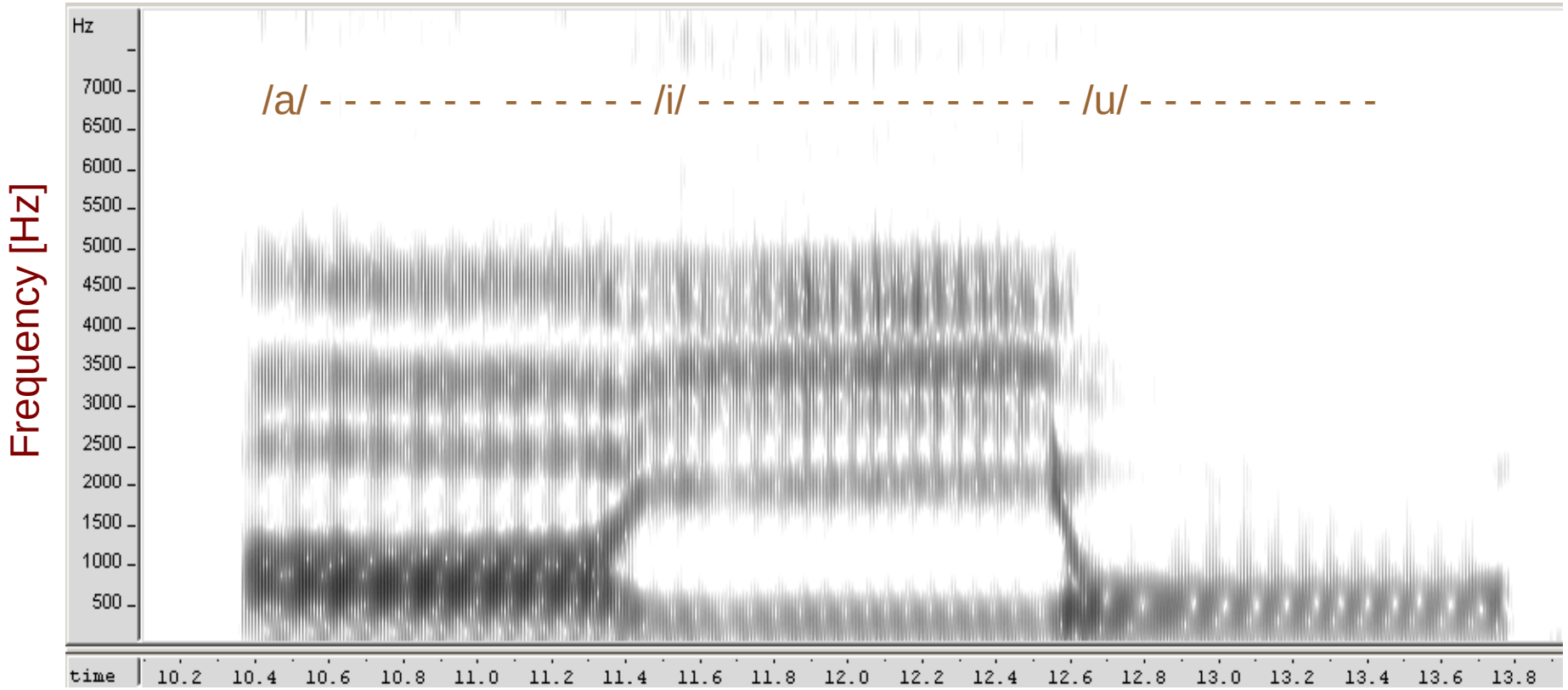


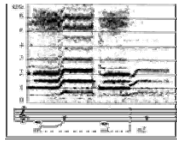
universidade de aveiro
theoria poiesis praxis

inet

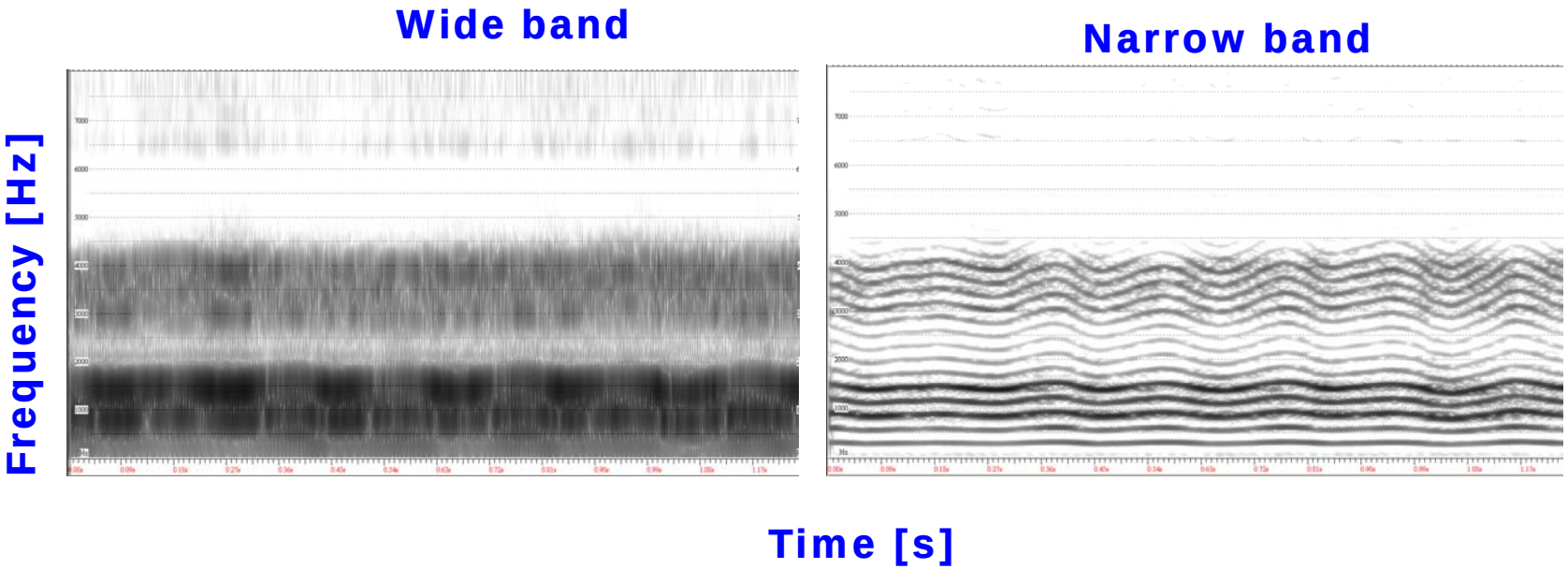
Spectrogram

Time along horisontal axis
Frequency along vertical axis
Dark for formant frequencies of different vowels





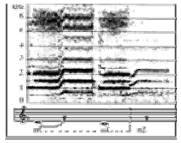
Different types of spectrograms



Wide (left) and arrow (right) band spectrograms of a vowel /a/ sung in chest register D4

- **Wide band:** used to display formant frequencies with clarity (because it divides the frequency spectrum into broad swaths)
- **Narrow band:** clear visualization of harmonics and vocal elements (because it divides the frequency spectrum into narrow segments)

Applicability of spectrum displays in teaching



❑ Important considerations

❑ Spectral displays are easier to understand for male than for female voices (frequency range that allows the display of many harmonics)

❑ Spectral displays are quite useful for working with the student in learning to control:

- ❖ Phonation types
- ❖ Vocal timbre (singer's spectrum peak; chiaroscuro tone)
- ❖ Vibrato
- ❖ Legato
- ❖ Staccato (voice onset and offset)
- ❖ Vowel accuracy
- ❖ Synchronicity with accompaniment
- ❖ Acoustical efficiency

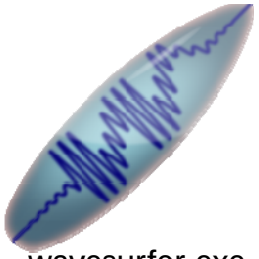
SPECTRUM DISPLAY

APPLICABILITY IN TEACHING

universidade de aveiro
theoria poiesis praxis

inet

Free

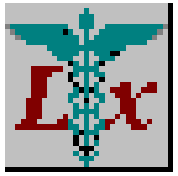


wavesurfer.exe

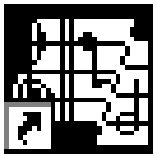
praat

RTSect.lnk

Paid



Spead3.exe

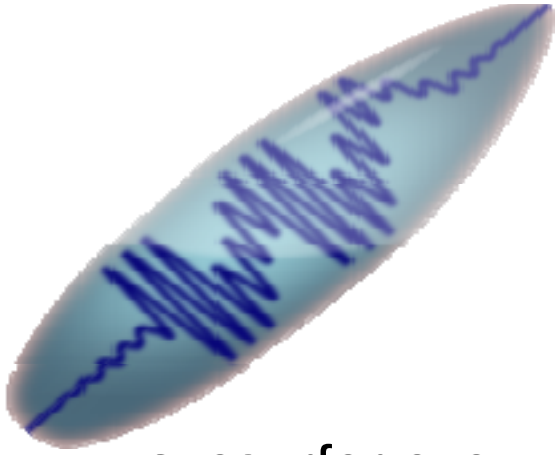


VoceVista 3.lnk

OVERTONE ANALYZER

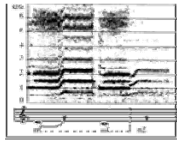
Sing & See
See your voice
Hear the results

Let us try Wavesurfer!



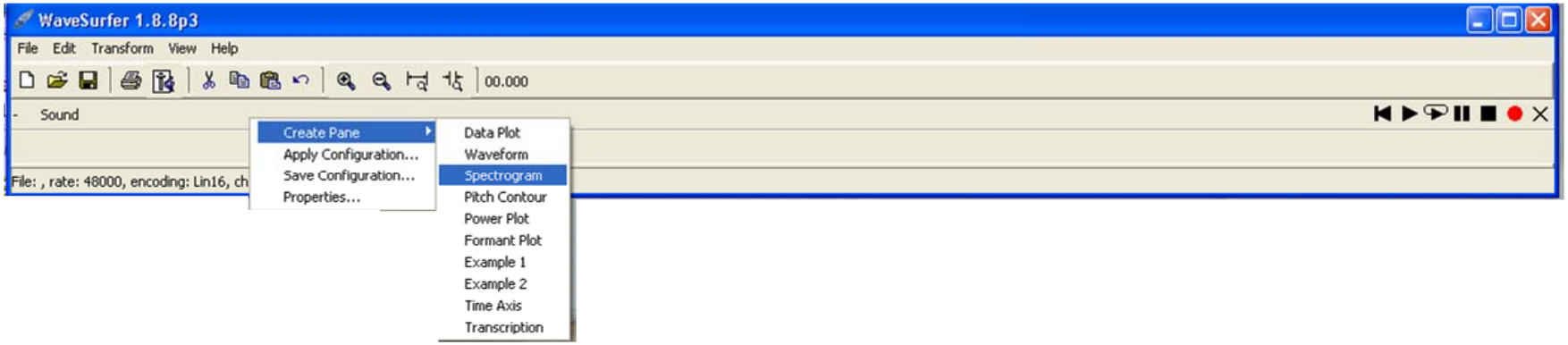
wavesurfer.exe

http://www.speech.kth.se/wavesurfer/



❑ Procedures

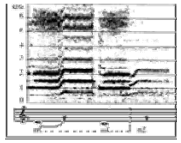
- Open the program
- Create a pane



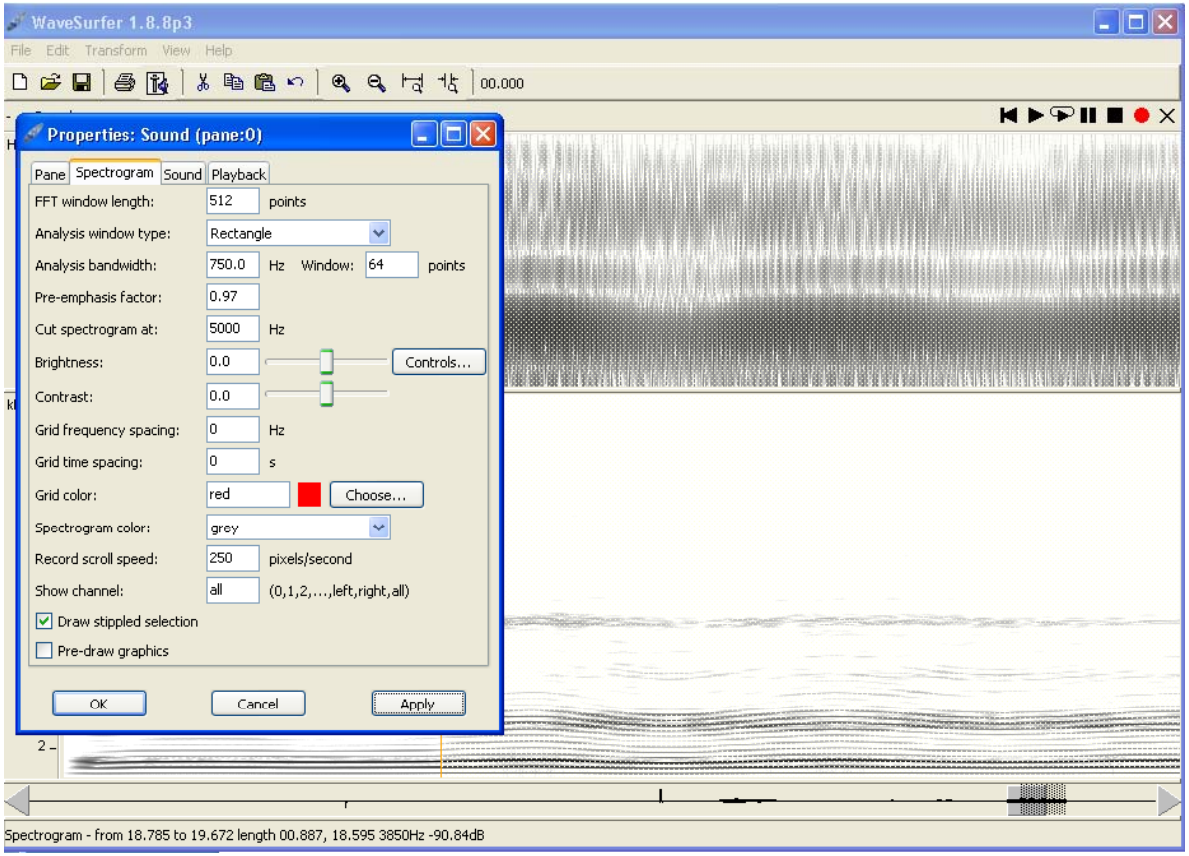
- Enlarge the window size
- Define type of spectrogram
  - ✓ **Wide band:**
    - ➢ FFT = 1024 points
    - ➢ Window = 1024 points
    - ➢ Cut spectrum at 5000Hz

SPECTRUM DISPLAY

APPLICABILITY IN TEACHING

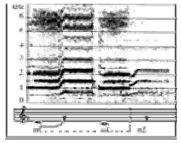


- Define type of spectrogram
  - ✓ **Narrow band:**
    - FFT = 512 points
    - Window = 64 points
    - Analysis window type: rectangle
    - Cut spectrum at 5000Hz



universidade de aveiro
theoria poiesis praxis

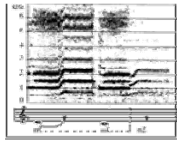
inet



Applicability of wide band spectrograms

SPECTRUM DISPLAY

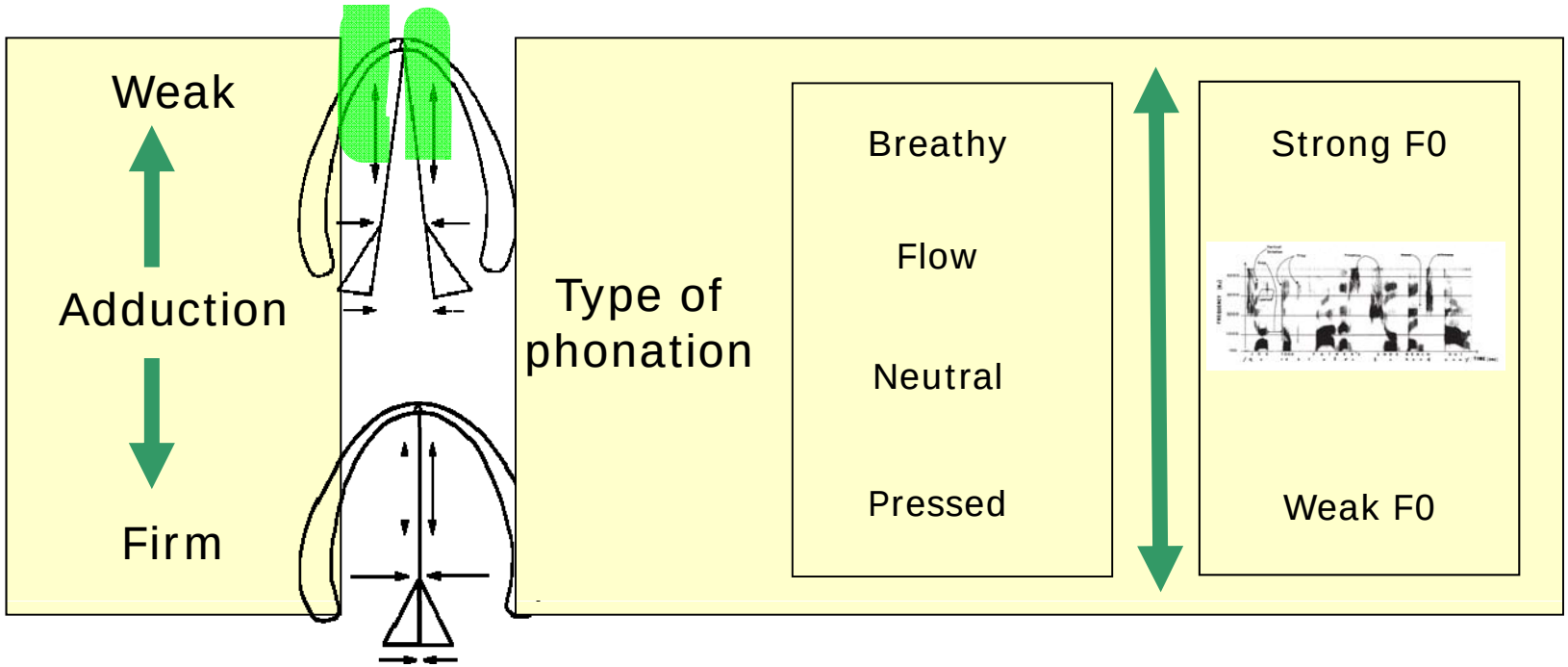
APPLICABILITY IN TEACHING



1. **Phonation types:** determine voice timbre and correspond to different levels of vocal effort can be demonstrated using a synthesizer software (Madde, by Svante Granqvist)

madde(3).exe.lnk

Courtesy of Professor Johan Sundberg

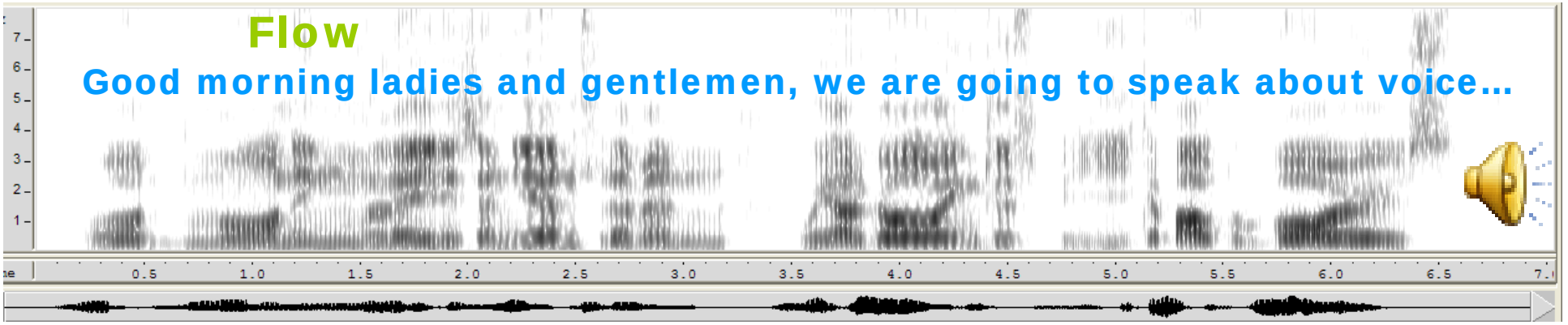


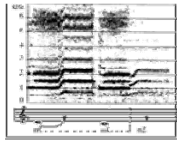
SPECTRUM DISPLAY

APPLICABILITY IN TEACHING

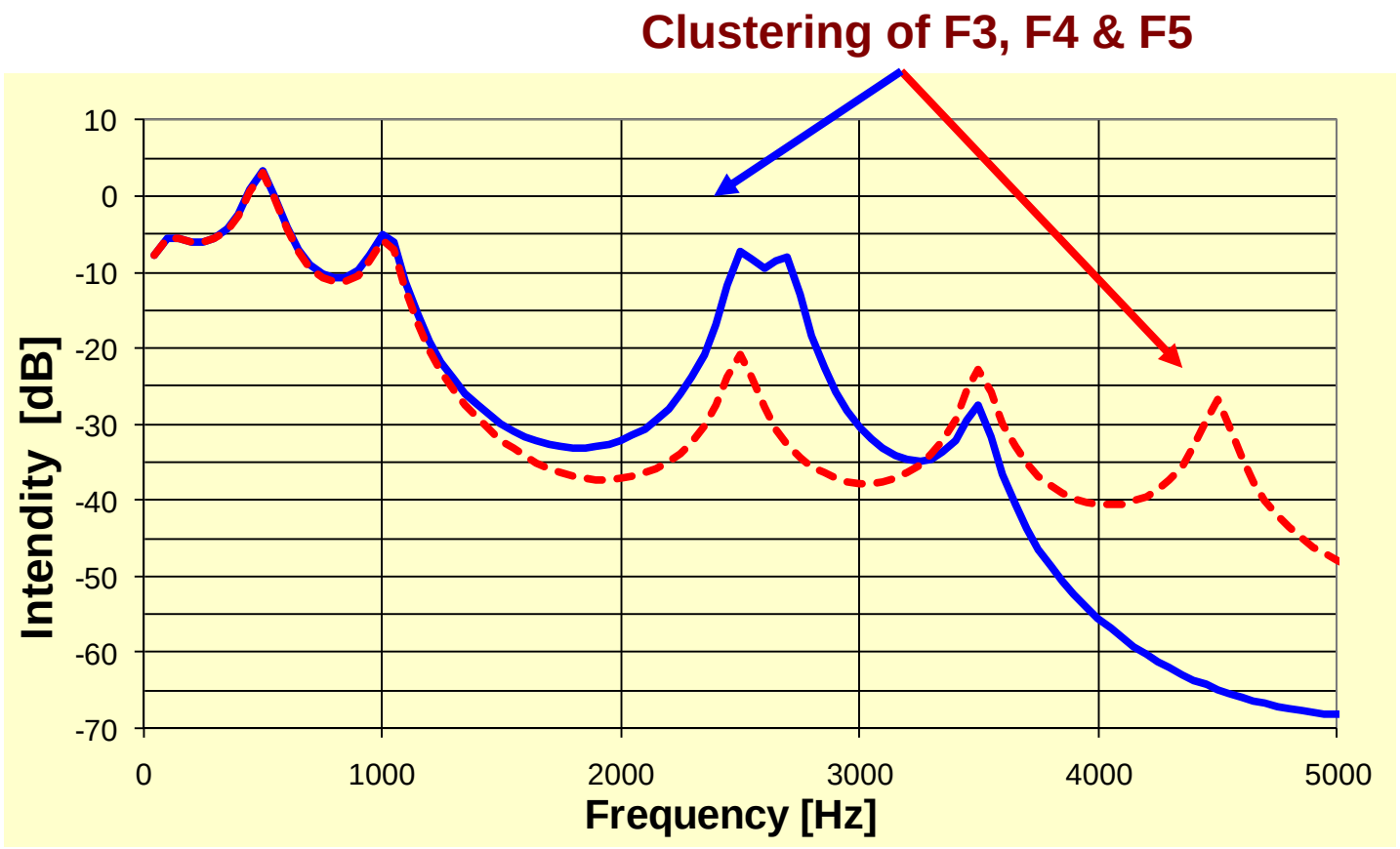
universidade de aveiro
theoria poiesis praxis

inet





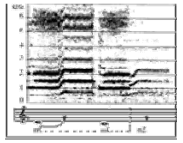
2. Singer's spectrum peak: the fine art of clustering resonances by male classically trained singers can be displayed in wide band spectrograms



(Courtesy of Professor Sundberg, "Distinguished Lecturer", CIRMMT, 2009)

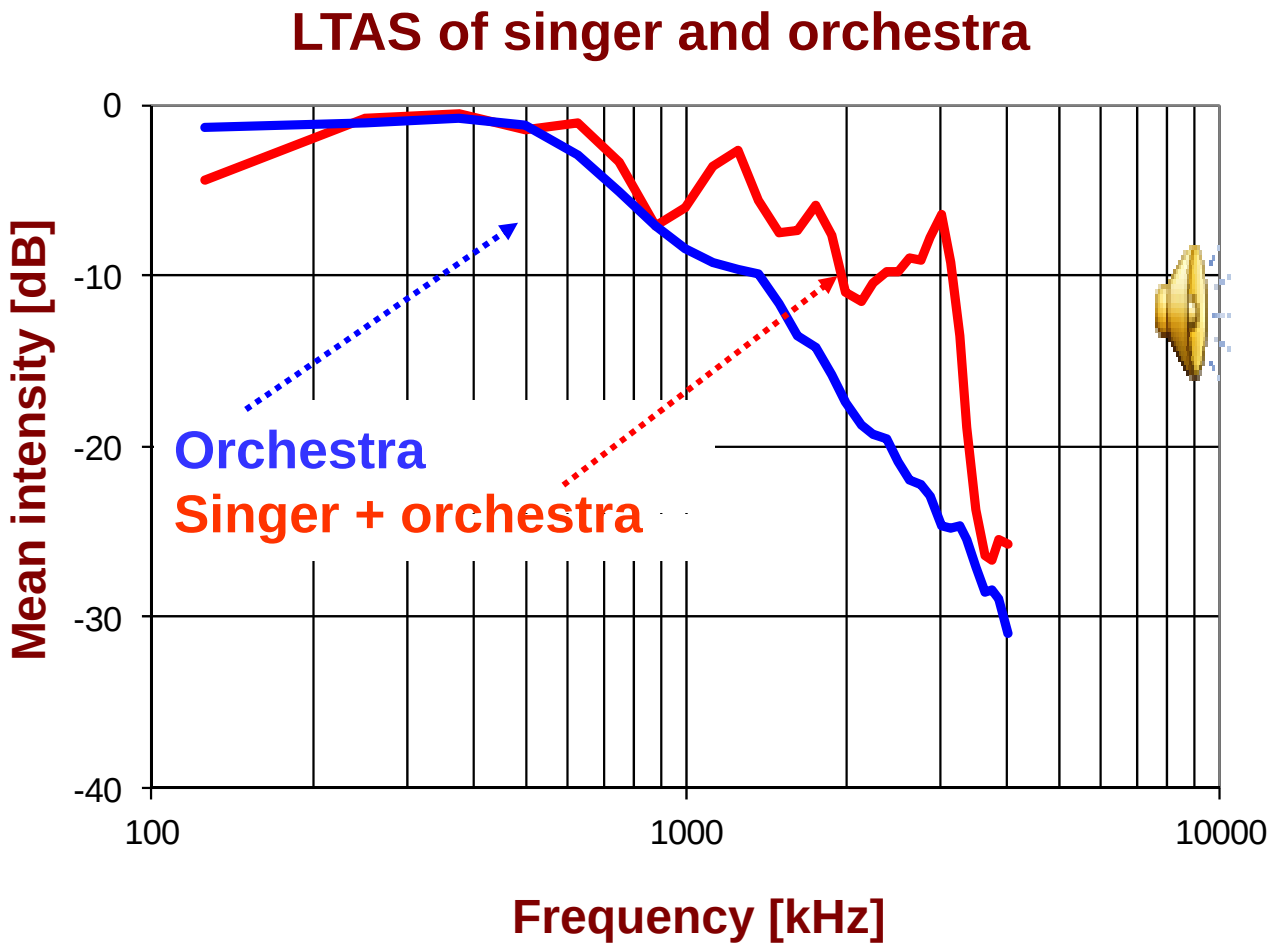
FEEDBACK EM TEMPO REAL

SISTEMA RESSOADOR



READING SPECTROGRAMS Filipa Lã, 2012

- **Singer's spectrum peak:** one strategy that male singers use to be heard over a loud accompaniment



1º - Sound corresponding to the level of an orchestra

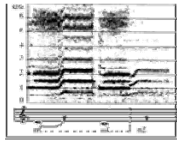
2º - Singing without and with singer's spectrum peak

3º - First and second examples together

(Cortesy from Professor Sundberg, "Distinguished Lecturer", CIRMMT, 2009)

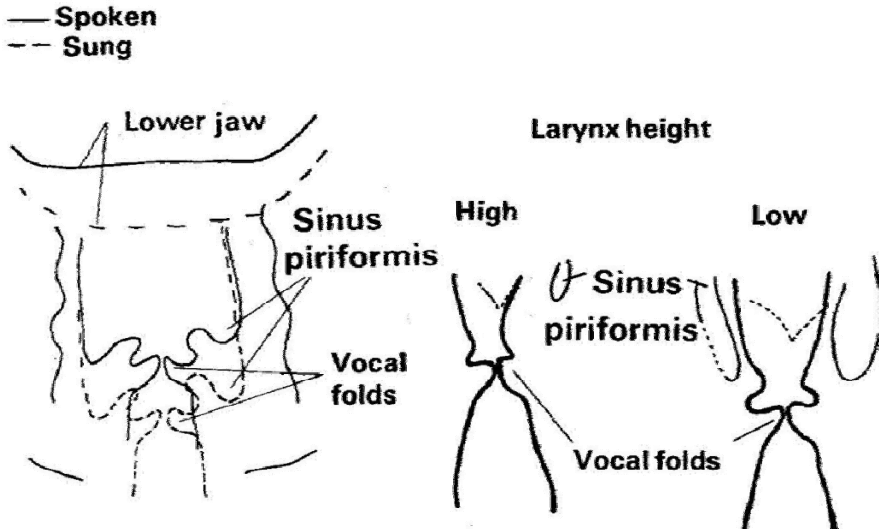
SPECTRUM DISPLAY

APPLICABILITY IN TEACHING



- **Physiologically**: lowering of the larynx and thus widening of the pharynx

- widening of laryngeal ventricle and the *sinus piriformes* (i.e. Bottom part of the vocal tract surrounding the larynx tube)



Comparison between speaking and singing of a male x-ray frontal tracings (adapted from Sundberg 1987: 121)

SPECTRUM DISPLAY

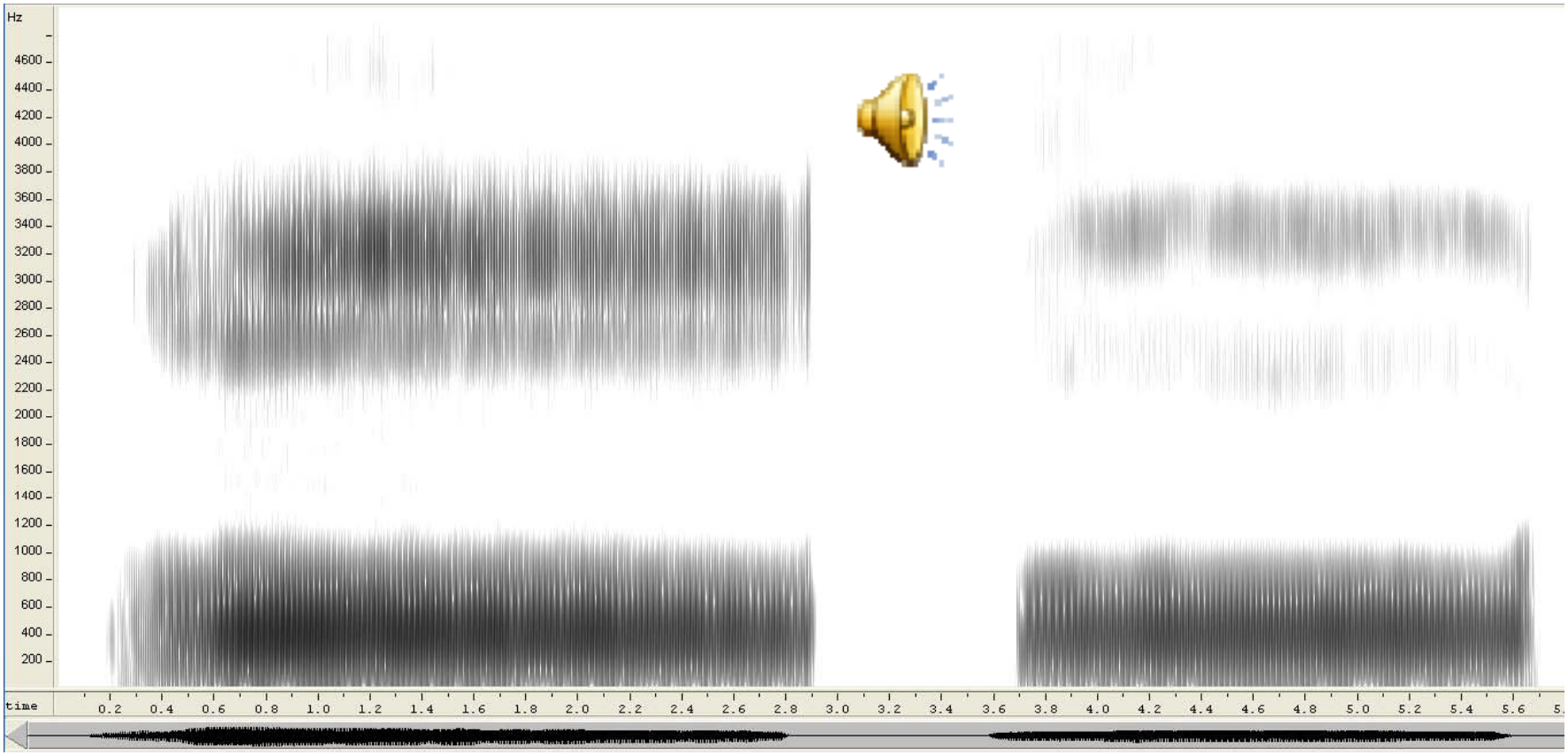
APPLICABILITY IN TEACHING

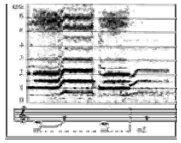
universidade de aveiro
theoria poiesis praxis

inet

/u/ vowel sung with singer's spectrum peak

/u/ vowel sung without singer's spectrum peak





❑ Aspects of language accuracy

- ❖ Intelligible pronunciation of lyrics
- ❖ Correct pronunciation of different languages
- ❖ Accuracy of vowel production in each language

SPECTRUM DISPLAY

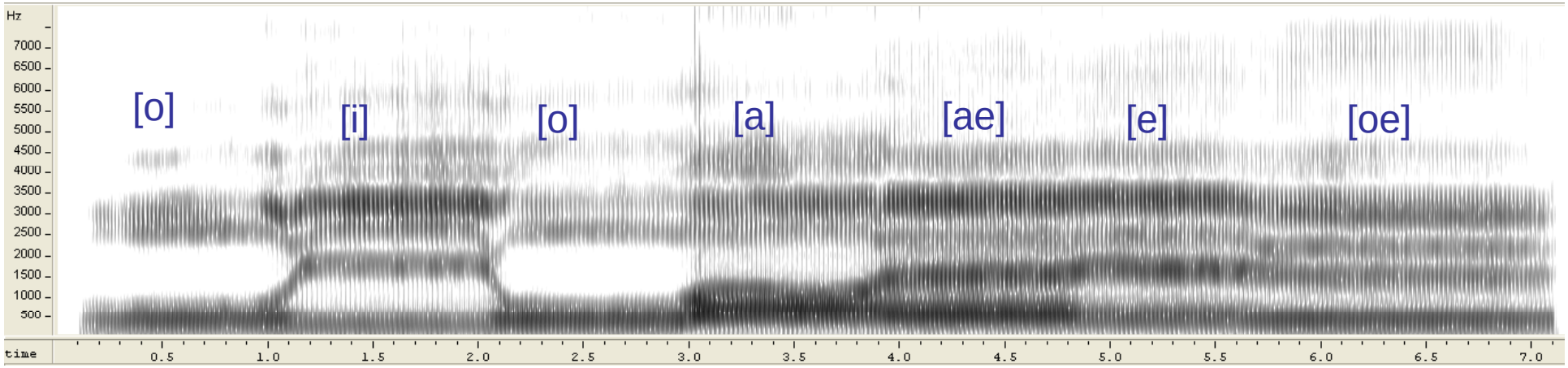
APPLICABILITY IN TEACHING

universidade de aveiro
theoria poiesis praxis

inet

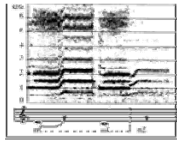
❑ Aspects of language accuracy

- ❖ Intelligible pronunciation of lyrics
- ❖ Correct pronunciation of different languages
- ❖ Accuracy of vowel production in each language



SPECTRUM DISPLAY

APPLICABILITY IN TEACHING

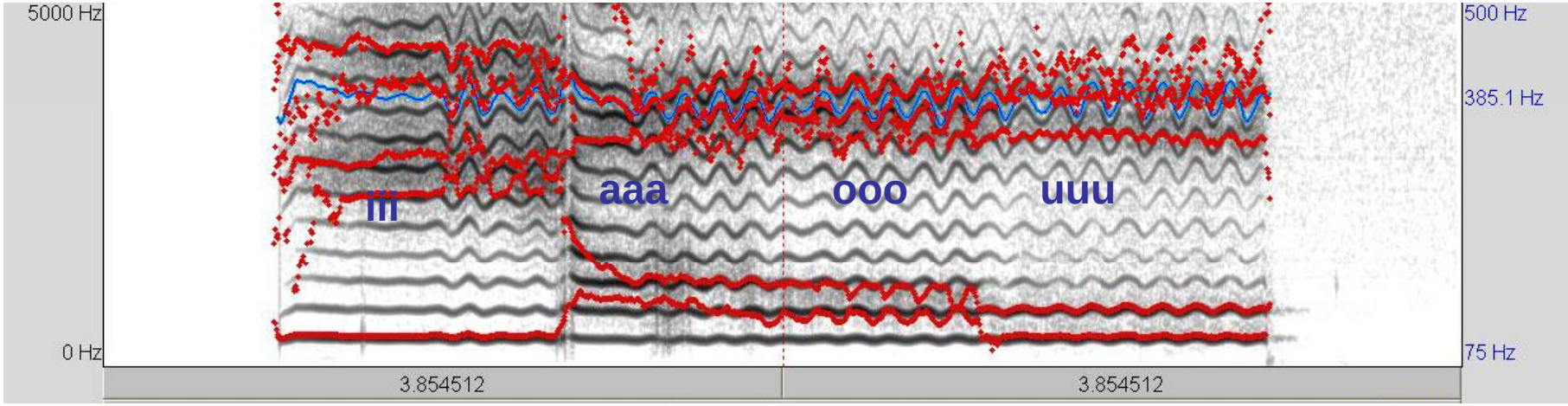


READING SPECTROGRAMS

Filipa Lã, 2012

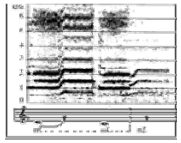
❑ Formant frequencies of vowels

Narrow band spectrogram



universidade de aveiro
theoria poiesis praxis

inet



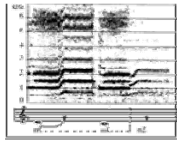
Applicability of narrow band spectrograms

SPECTRUM DISPLAY

APPLICABILITY IN TEACHING

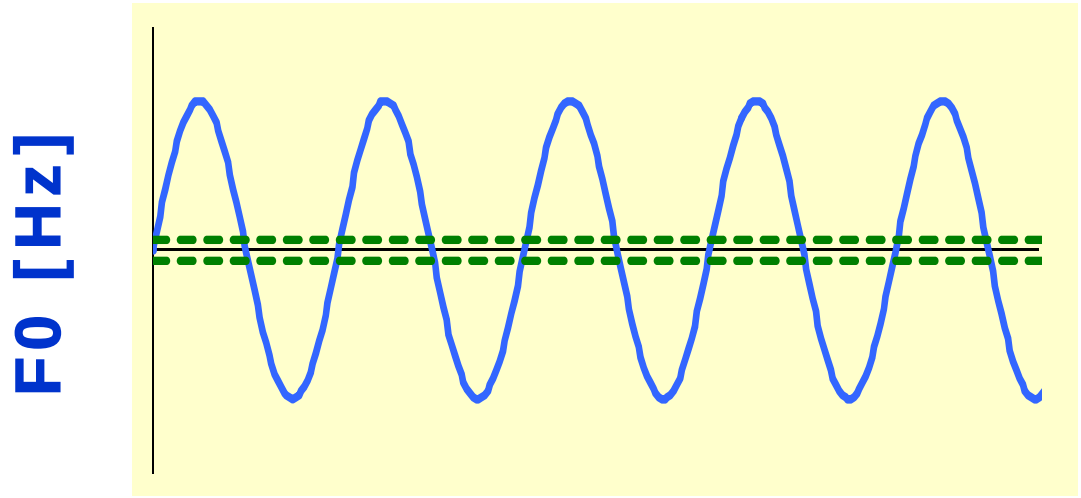
universidade de aveiro
theoria poiesis praxis

inet



READING SPECTROGRAMS
Filipa Lã, 2012

- **Vibrato:** regular variation of F0
  - the perceived frequency corresponds to the mean F0 variation
  - the mean must fall into a tolerated rate (usually between 5 and 6 Hz)

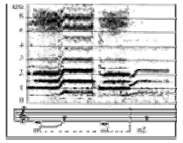


madde(3).exe.lnk

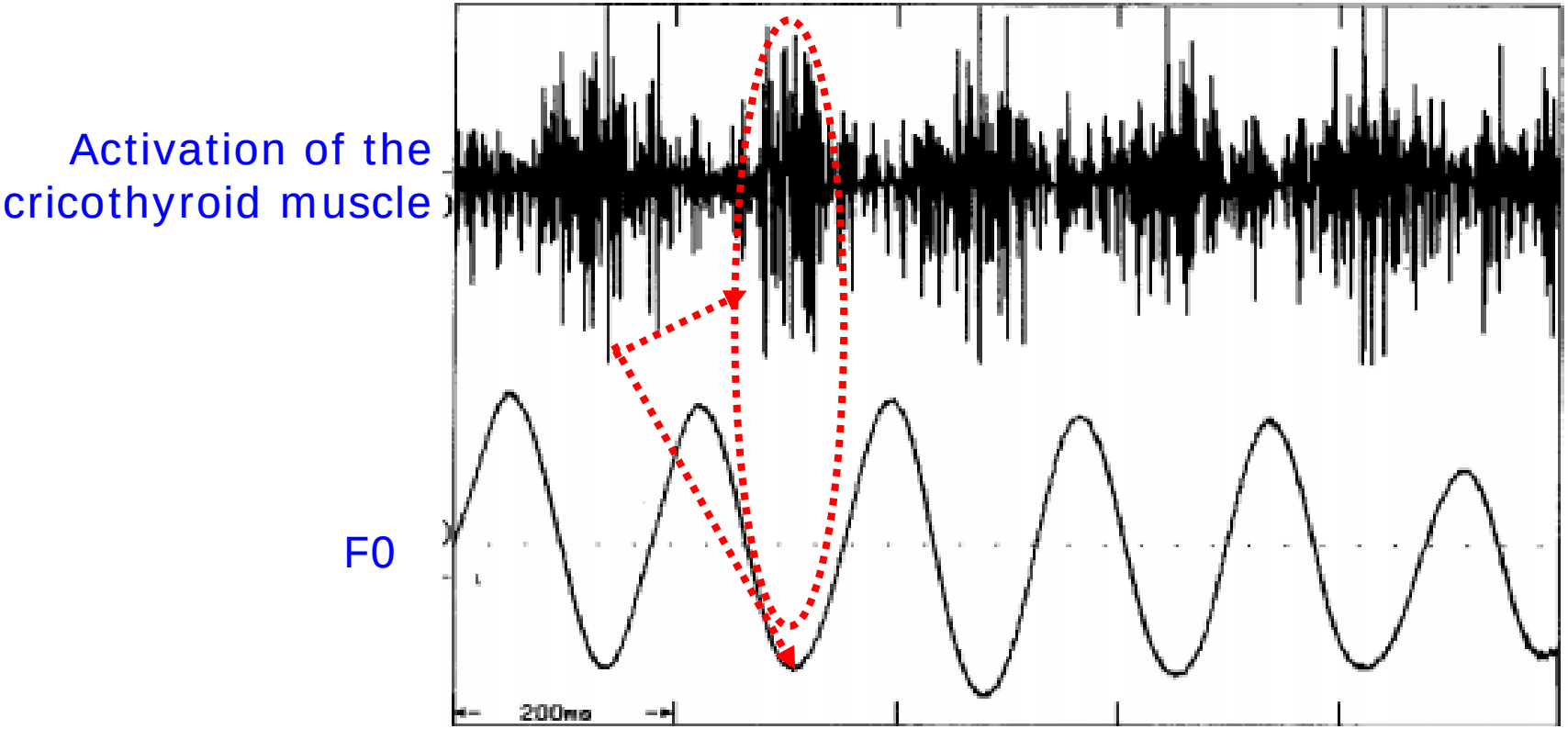
Representation of vibrato (adapted from Lã & Sundberg, 2010: Perceptual evaluations of Summer course "**Function of the Singing Voice**", 2010)

SPECTRUM DISPLAY

APPLICABILITY IN TEACHING



- **Vibrato:** regular variation of F0



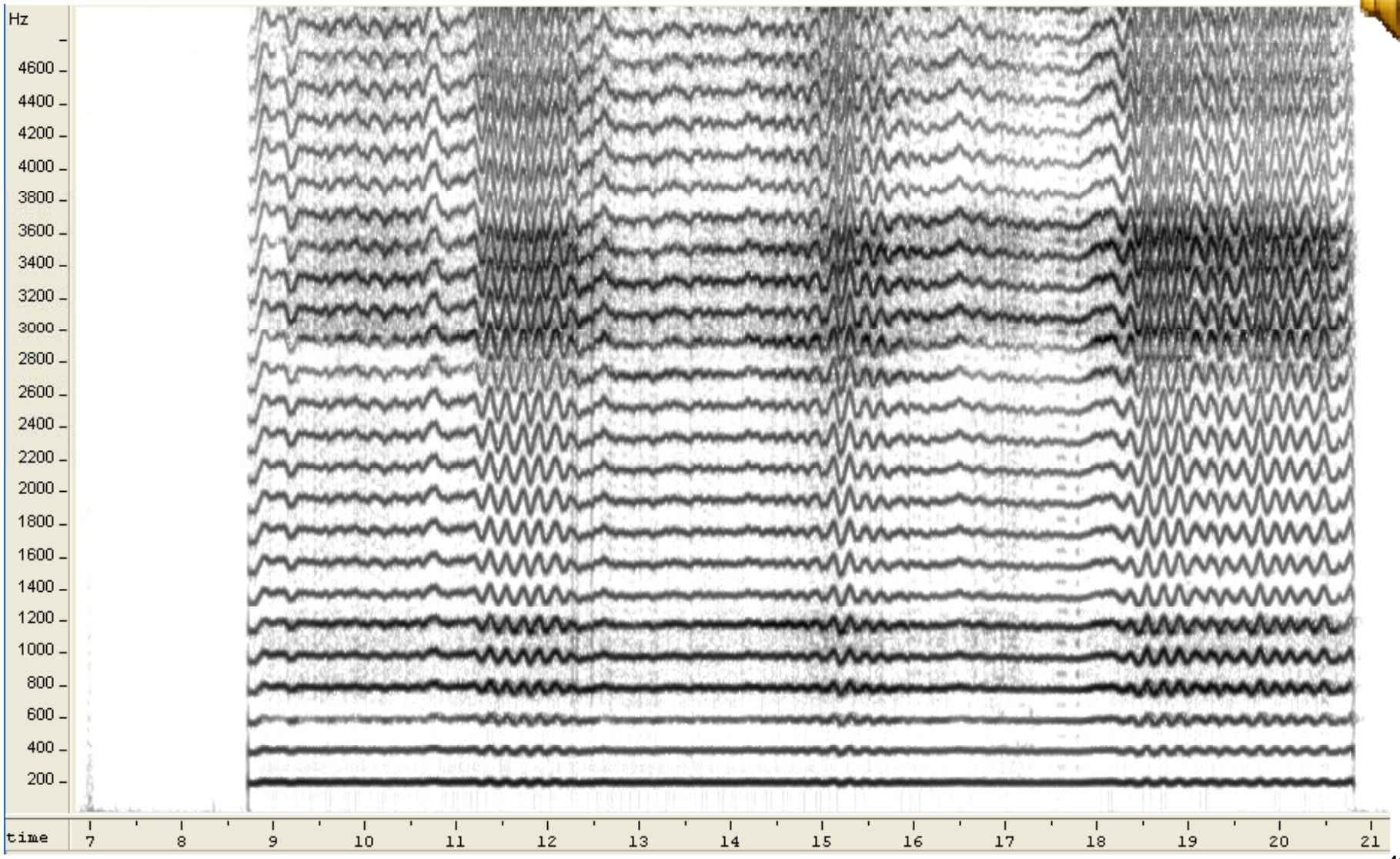
SPECTRUM DISPLAY

APPLICABILITY IN TEACHING

universidade de aveiro
theoria poiesis praxis

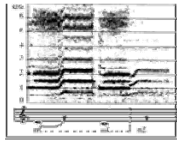
inet

- **Vibrato:** regular variation of F0

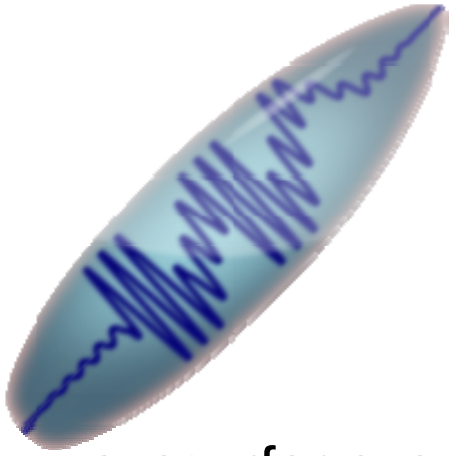


SPECTRUM DISPLAY

APPLICABILITY IN TEACHING



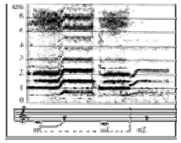
- **Legato:** visualised in the spectrogram through the continuity of the signal



wavesurfer.exe

SPECTRUM DISPLAY

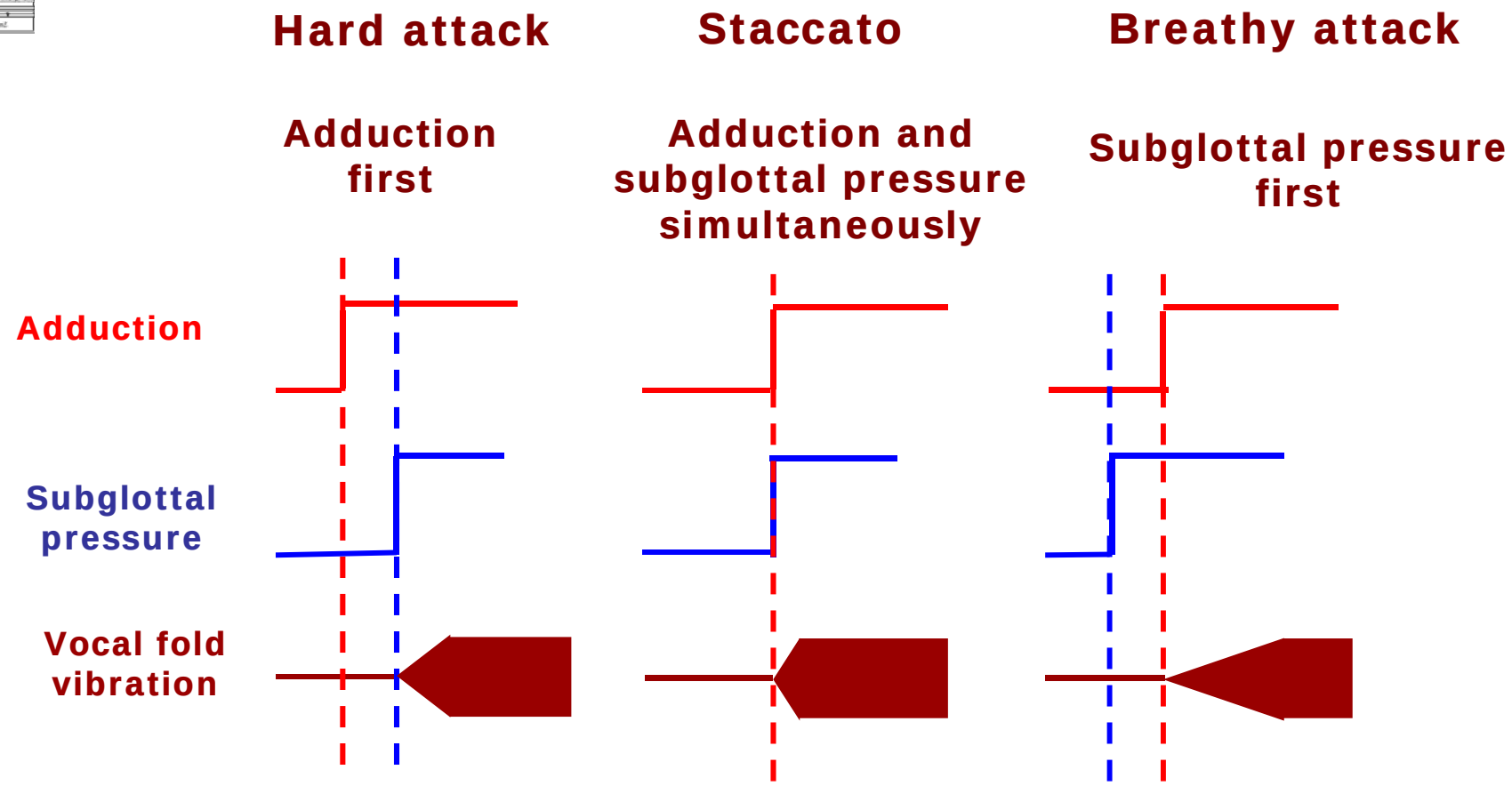
APPLICABILITY IN TEACHING



READING SPECTROGRAMS

Filipa Lã, 2012

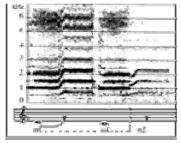
❑ Voice onset and offset



(Adapted from Lã & Sundberg, 2010: Perceptual evaluations of voices- Summer course "**Function of the Singing Voice**", 2010)

SPECTRUM DISPLAY

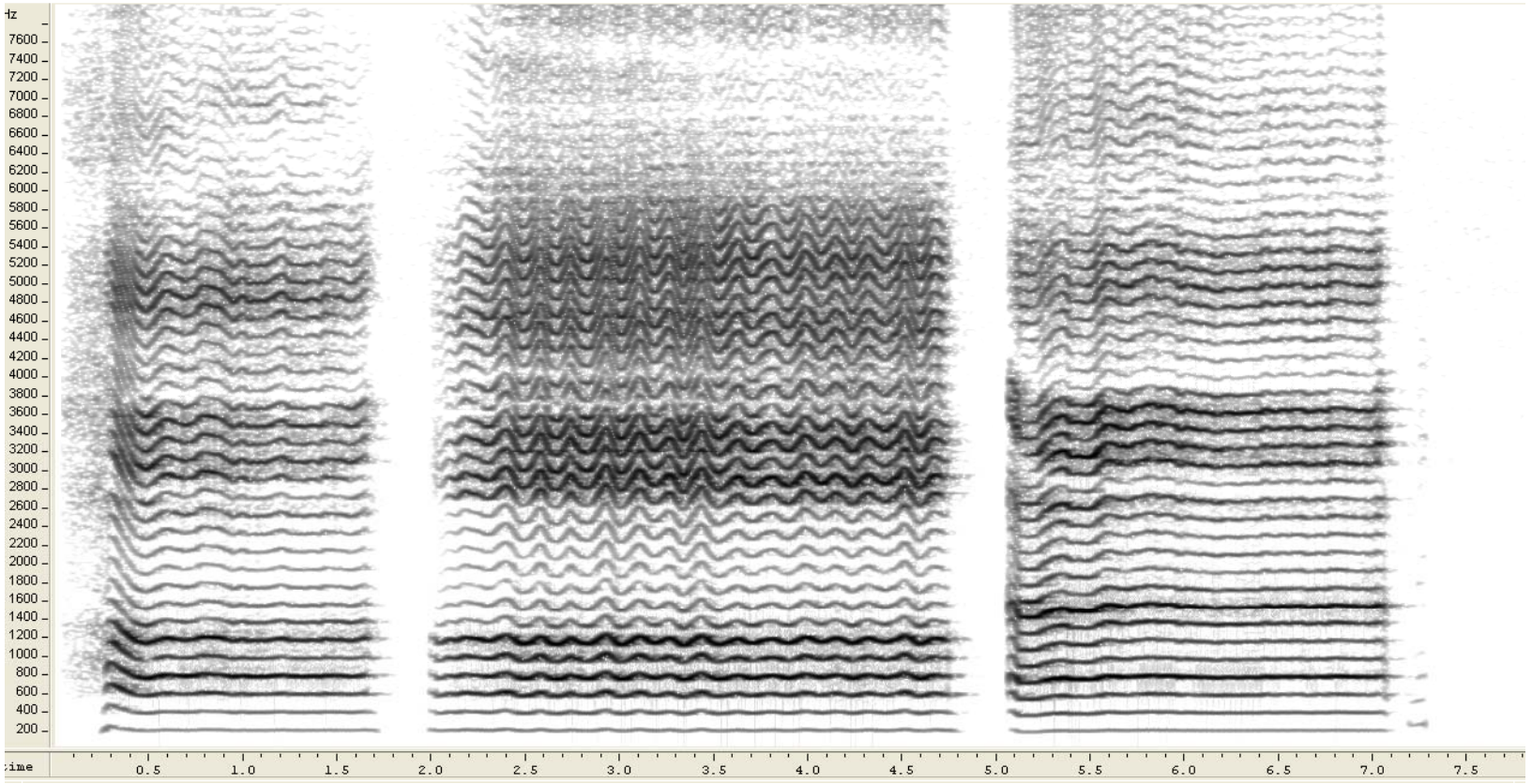
APPLICABILITY IN TEACHING



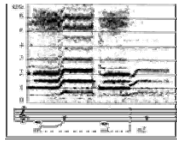
READING SPECTROGRAMS

Filipa Lã, 2012

❑ Different vocal onsets



Aspirated **Staccato** **Hard**



❑ Synchrony with accompaniment

- ❖ Realising intended timing of tones

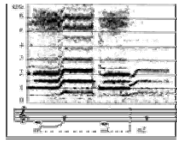
- ❖ Requires perfect control of respiratory apparatus, laryngeal function and vocal tract articulation

SPECTRUM DISPLAY

APPLICABILITY IN TEACHING

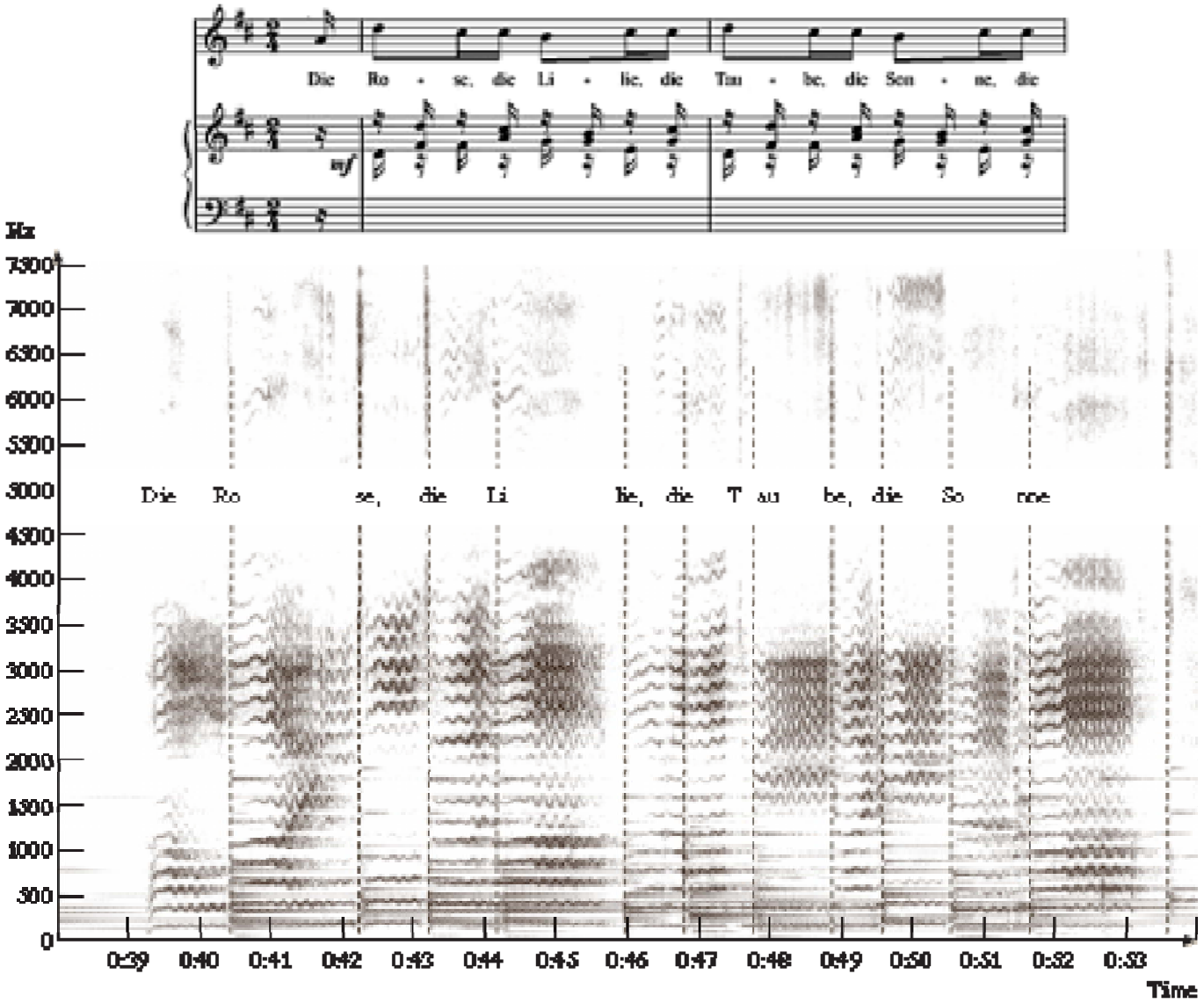
universidade de aveiro
theoria poiesis praxis

inet

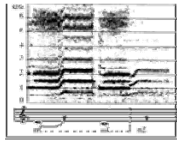


READING SPECTROGRAMS

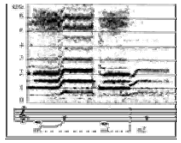
Filipa Lã, 2012



Almost perfect synchrony between vowel onset and the piano by Fischer-Dieskau's recording of Dichterliebe by Schumann (adapted from Lindblom & Sundberg, 2007: 694)



Applicability of power spectrum



❑ Acoustical efficiency

❑ Singers use register equalization to avoid discontinuities in voice quality between registers (e.g. Schutte & Miller, 1990; Titze, 1988;Schutte et al., 2005; Neuman et al., 2005)

❑ Previous studies suggest different formant tuning strategies around the male´s *passaggio* for open vowels:

- ❖ Falling of F1 below second partial (H2) (Miller & Schutte, 1994; Hertegård, Gaufin & Sundberg 1990; Neuman et al., 2005)
- ❖ F2 tracking H4 below the *passaggio* ('chest') and F2 tracking H3 at the level of the *passaggio* ('head') (Neuman et al., 2005)

❑ Vocal fry has been used to measure the frequencies of the two lowest formants in relation to the frequencies of the spectrum partials (Miller, 2000)

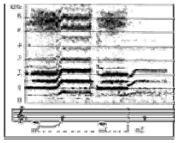
SPECTRUM DISPLAY

APPLICABILITY IN TEACHING

universidade de aveiro
theoria poiesis praxis

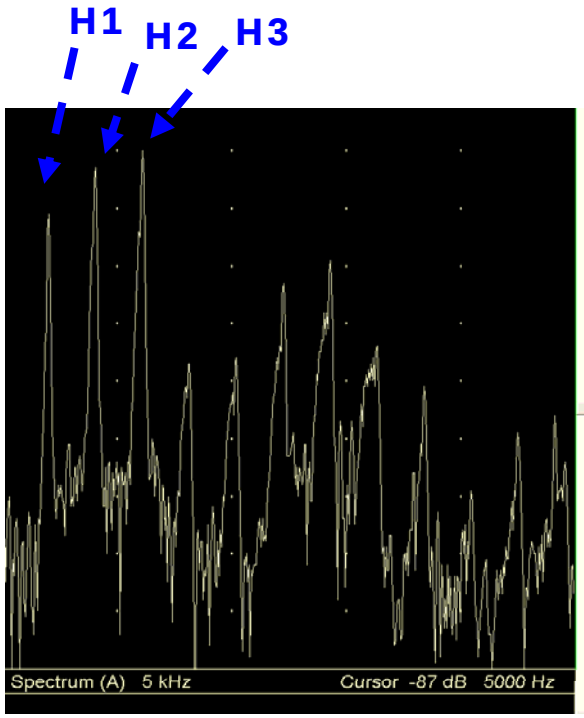
inet

(Sundberg, Lã & Gill, 2001)

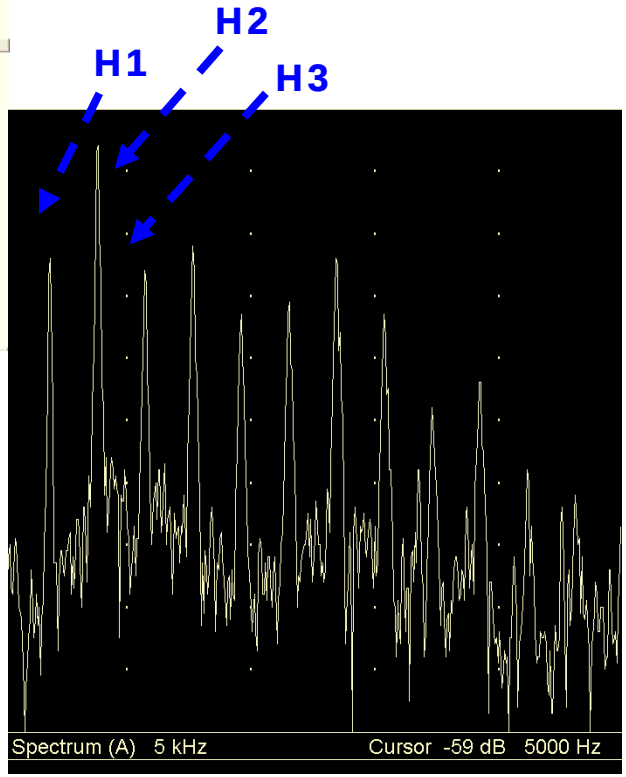


READING SPECTROGRAMS

Filipa Lã, 2012

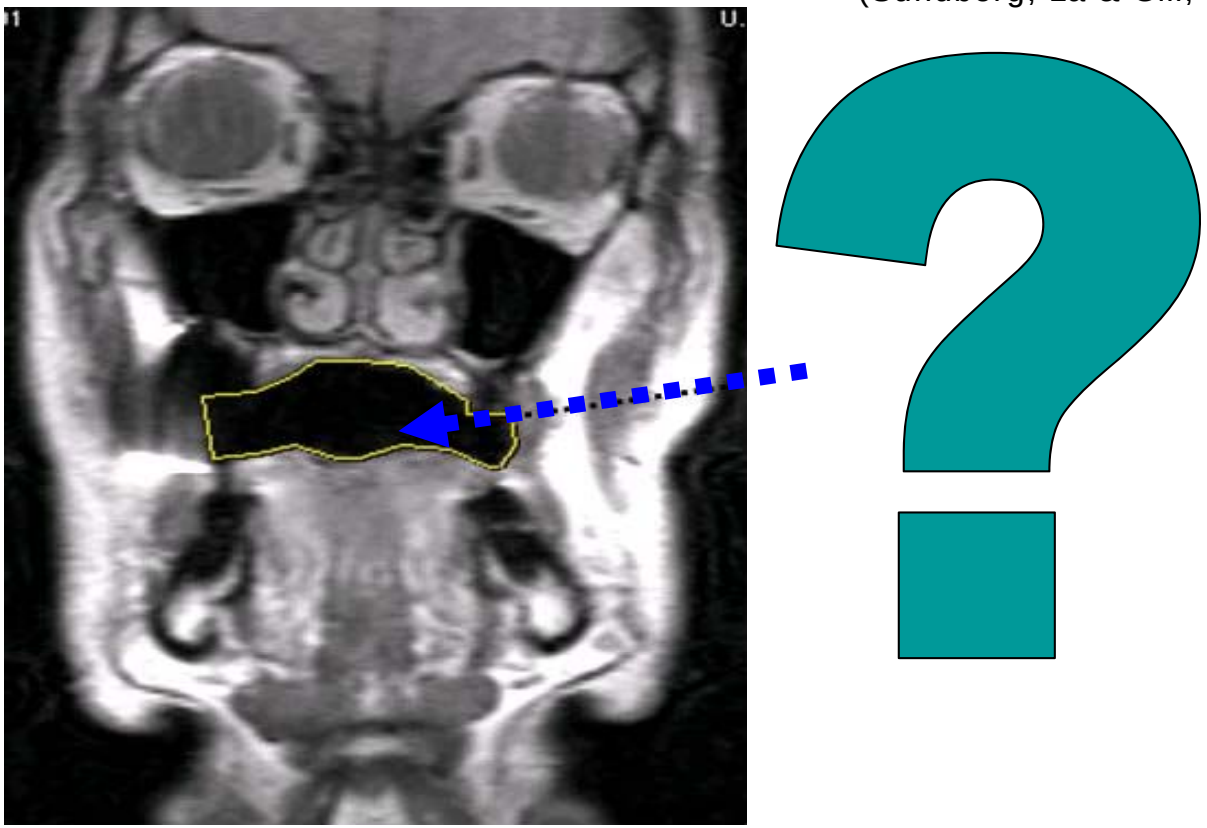


Using a “correct” tuning strategy for the *passaggio* note in a baritone (G4≅392 Hz)



Using an “incorrect” tuning strategy for the *passaggio* note (G4≅392 Hz)

(Sundberg, Lã & Gill, 2001)

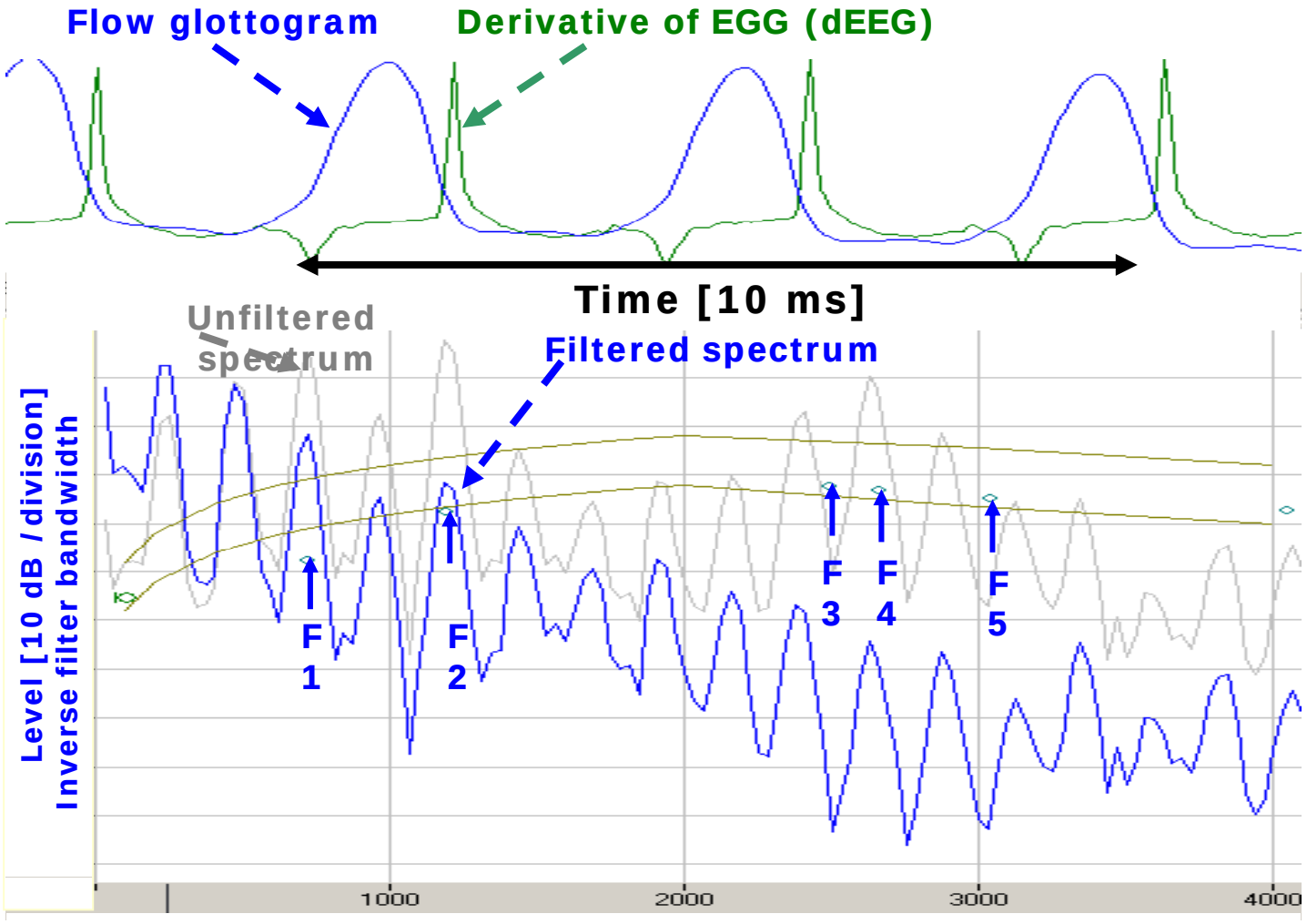


Where are the formants?

SPECTRUM DISPLAY

APPLICABILITY IN TEACHING

Inverse filtering by Decap (Sundberg, Lã & Gill, 2001)



Criteria for setting the filters:

- ❑ Ripple free closed phase
- ❑ Continuously falling voice source spectrum envelop
- ❑ Synchrony between dEEG peak and the flow discontinuity at glottal closure

SPECTRUM DISPLAY

APPLICABILITY IN TEACHING VOWEL /a/

(Sundberg, Lã & Gill, 2001)

- F1 and F2 are typically lower in the **Classical** than in the **Non-classical** tuning, especially in the *passaggio*
- Formants do not change systematically between scale tones
- For the top pitches, F1 falls below H2 in the **Classical** tuning
- For the top pitches, F2 is just above, just below or right on H3 in **Classical** tuning