Speech Science

WiSe 2023

Exercise 5: Acoustic analyses I

Dec 04, 2023

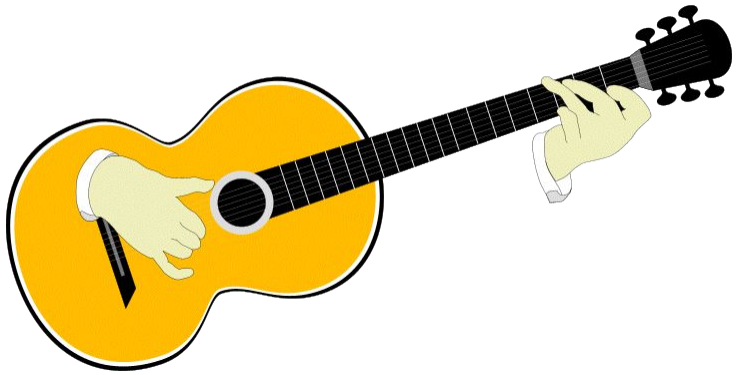
Bernd Möbius & Omnia Ibrahim

Language Science and Technology
Saarland University

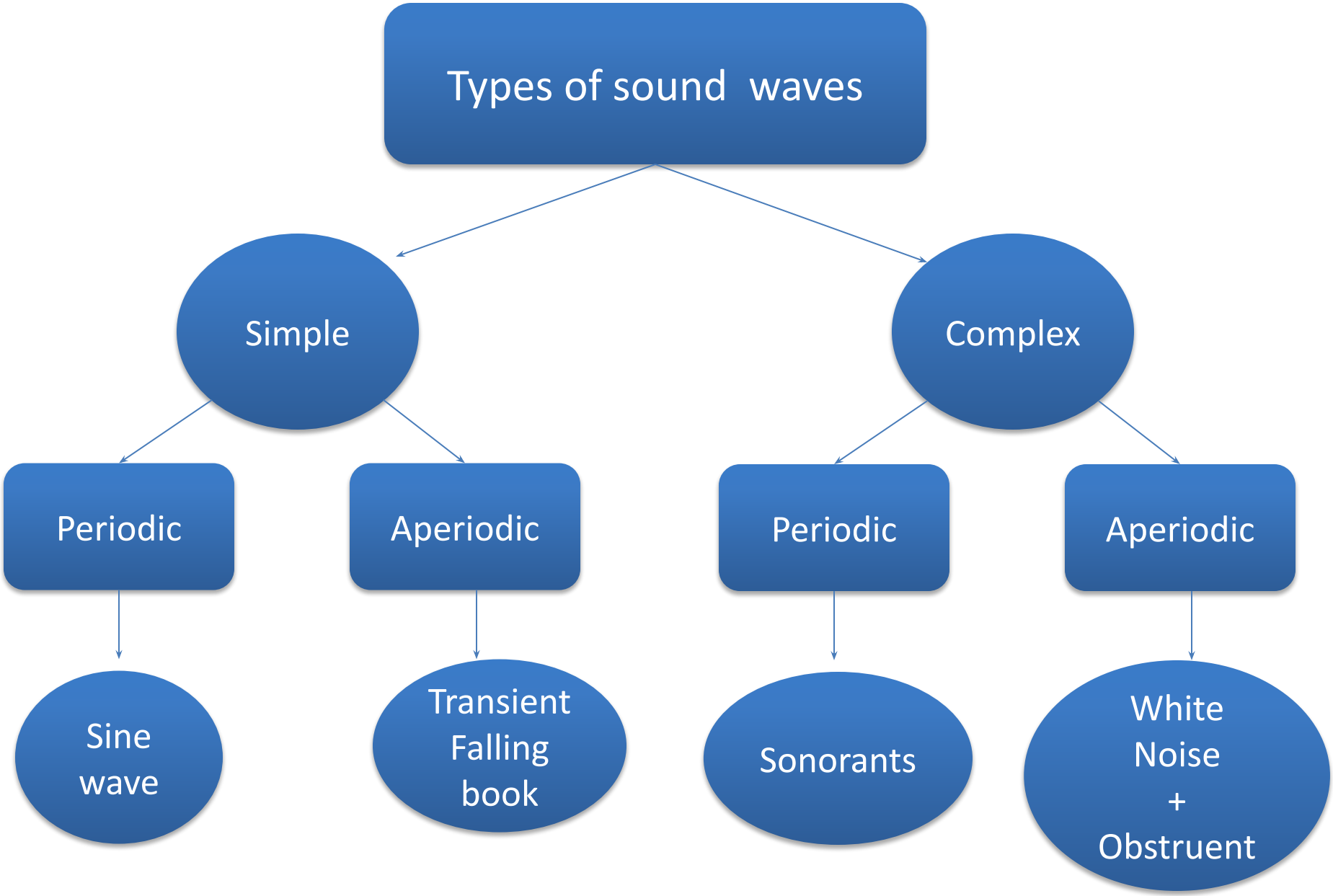
UNIVERSITÄT DES SAARLANDES

How to describe a sound?

1. Loudness → Amplitude → Power/energy.
2. Pitch → Frequency → Length/Tension/Mass.
3. Quality→ way of source vibration

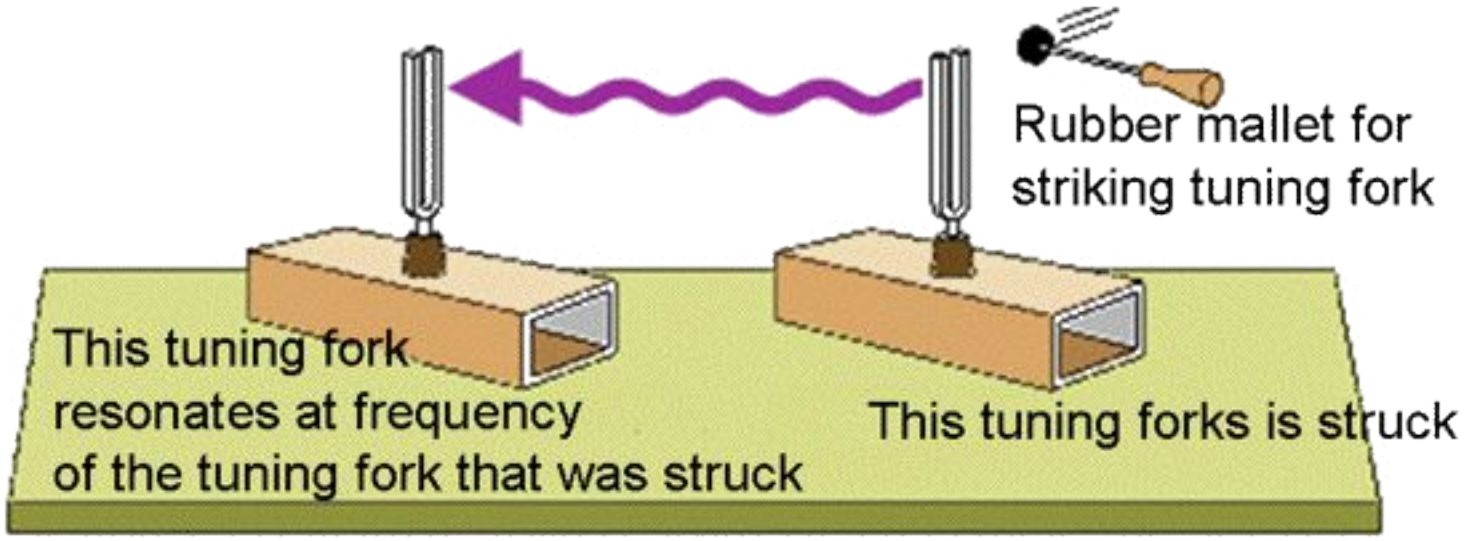


Basics


Types of sound waves
Simple
Complex
Periodic
Aperiodic
Periodic
Aperiodic
Sine wave
Transient Falling book
Sonorants
White Noise + Obstruent

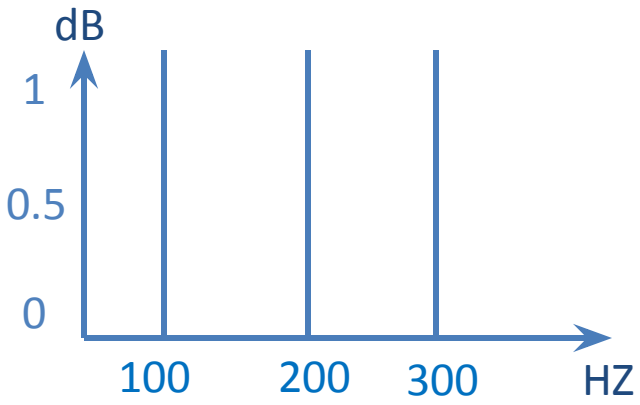
Resonance

- **Resonance:** one body can be set in motion by the vibration of another body.
- If the object is made to vibrate at its natural frequency by an external source of vibration



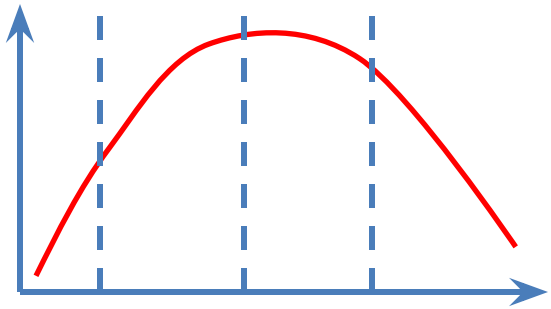
Resonance

Source Input



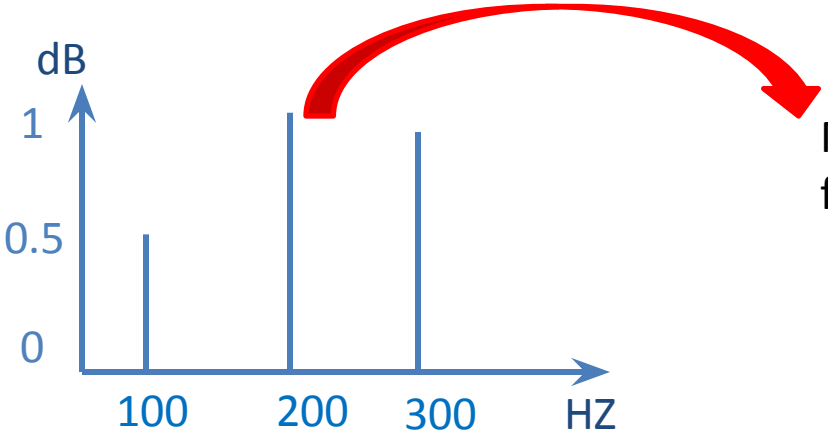
3 Pure tones with
Same amplitude

Transfer function



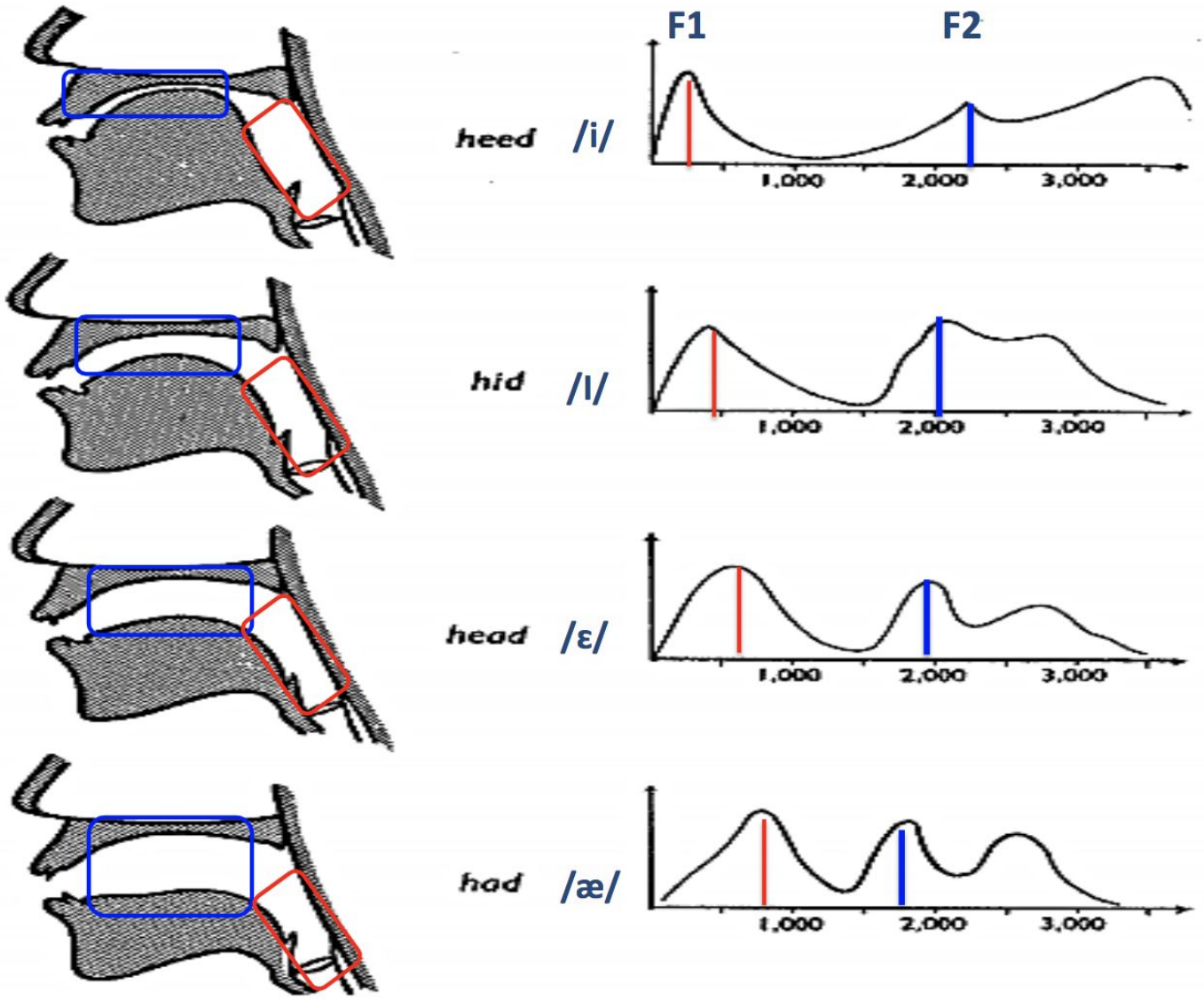
Resonator or filter
Prefer to vibrate at
200 cps

Output



Most effective
frequency

Vowel formants



Vowel formants

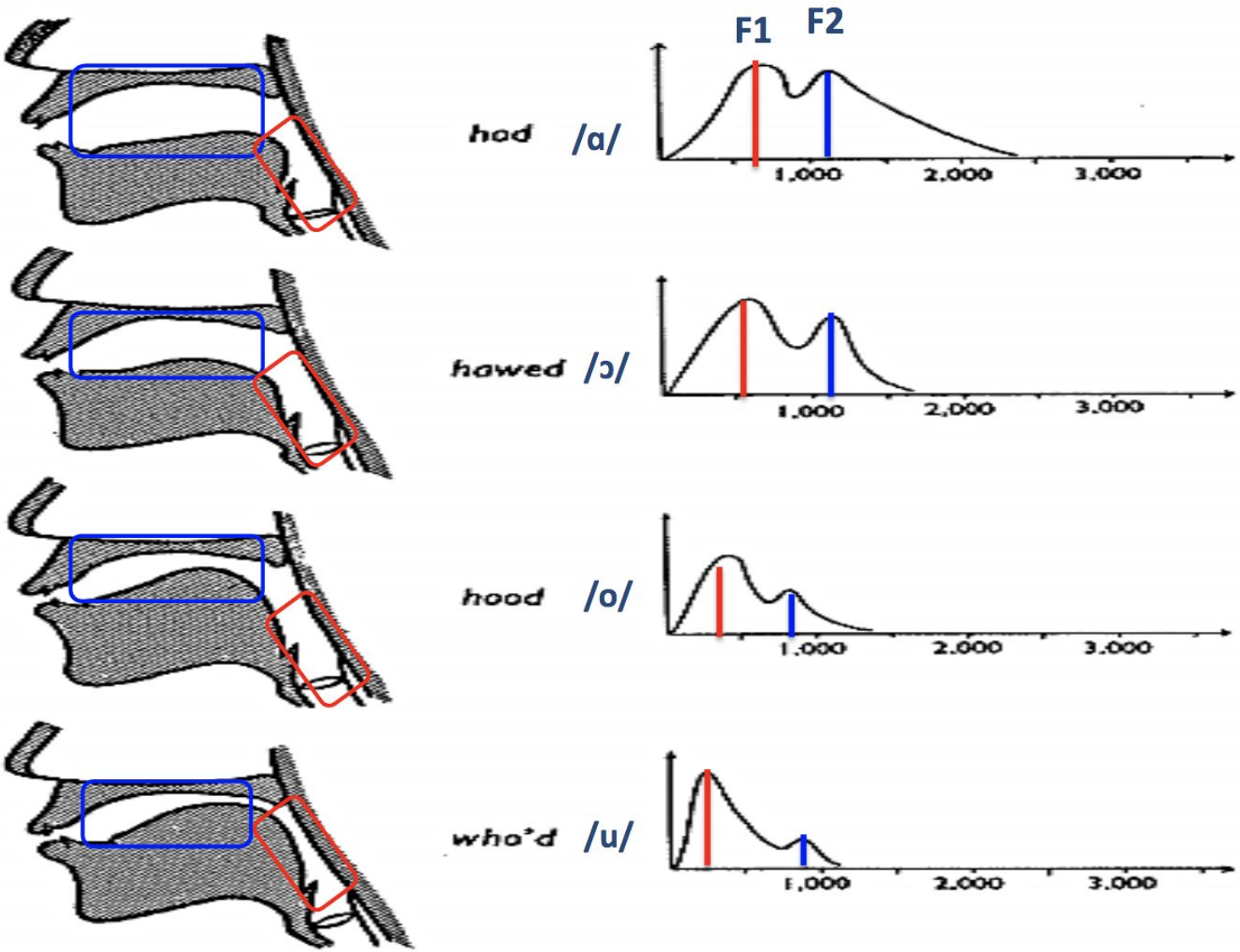


Figure 8.4 A spectrogram of the words *heed, hid, head, had, hod, hawed, hood, who'd* as spoken by a male speaker of American English. The locations of the first three formants are shown by arrows.

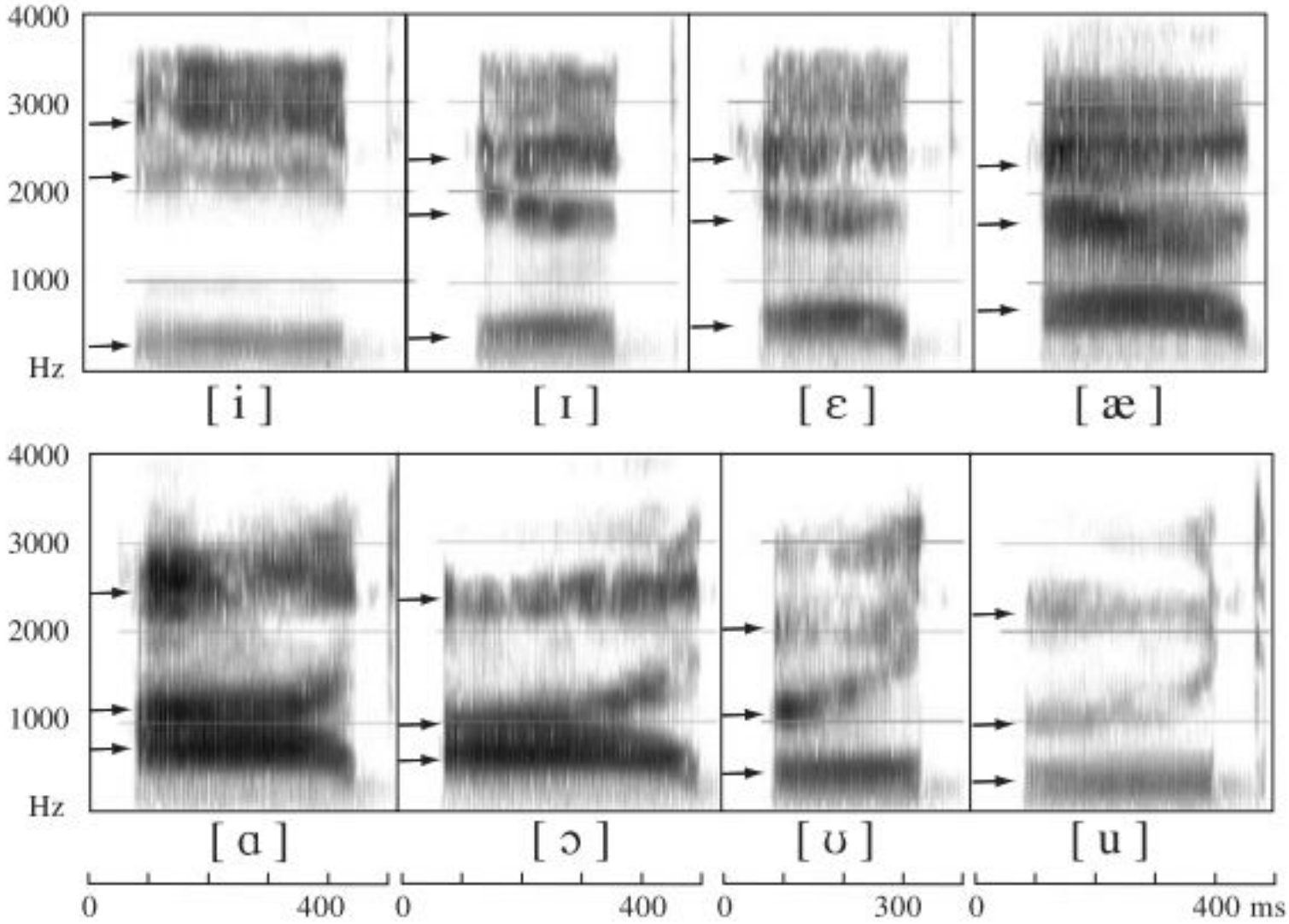
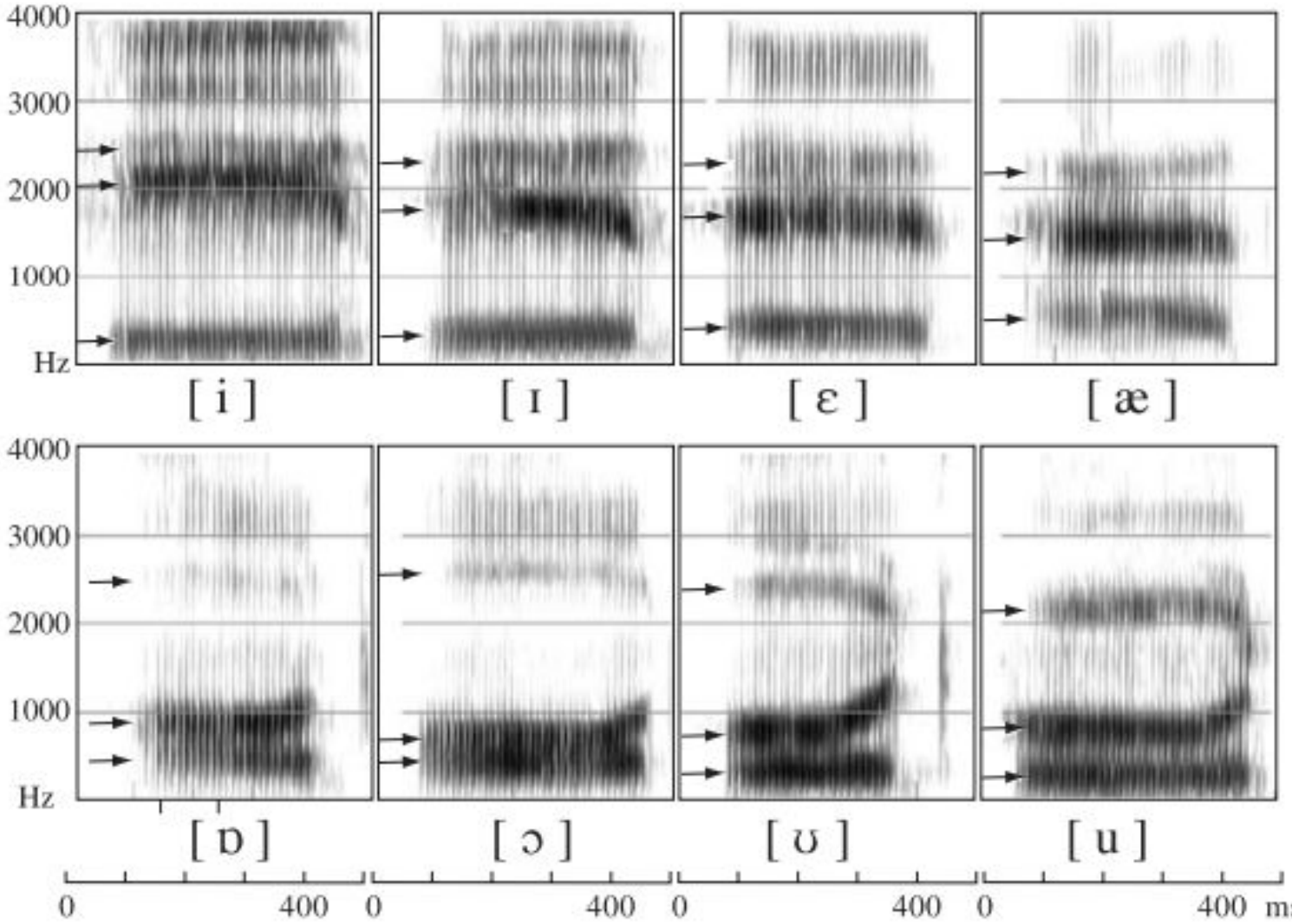


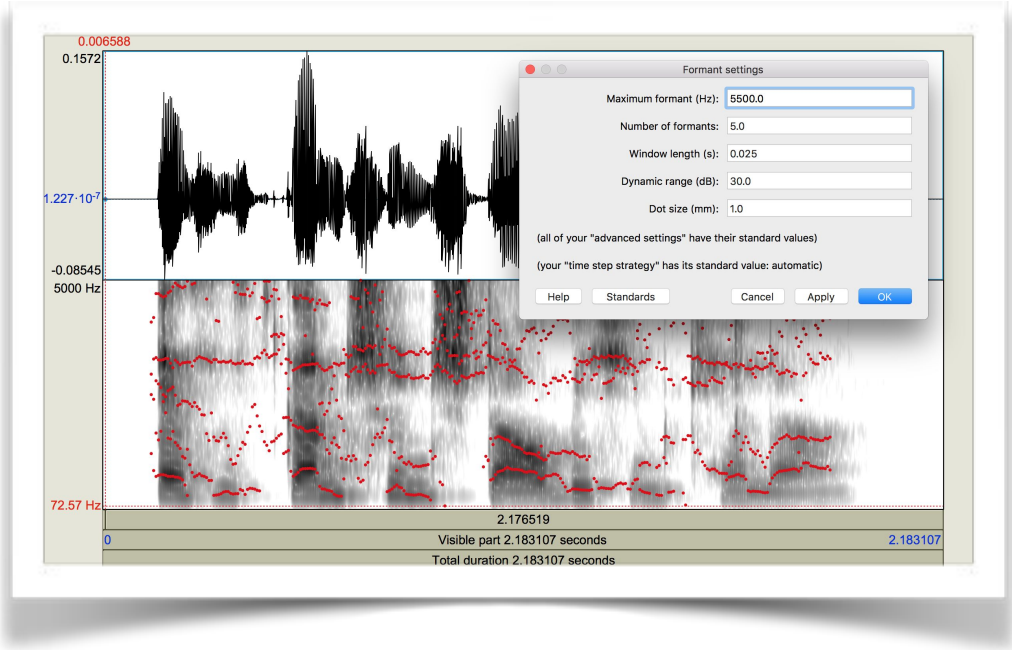
Figure 8.5 A spectrogram of the words *heed*, *hid*, *head*, *had*, *hod*, *hawed*, *hood*, *who'd* as spoken in a British accent. The locations of the first three formants are shown by arrows.



Acoustic analyses: formants

- *maximum formant* depends on speaker:
  - ♂ 5.000 Hz
  - ♀ 5.500 Hz
- *number of formants*: 5
- `formant listing`: F1-F4 for all time stamps in selection
- `get first/second/third/fourth formant`

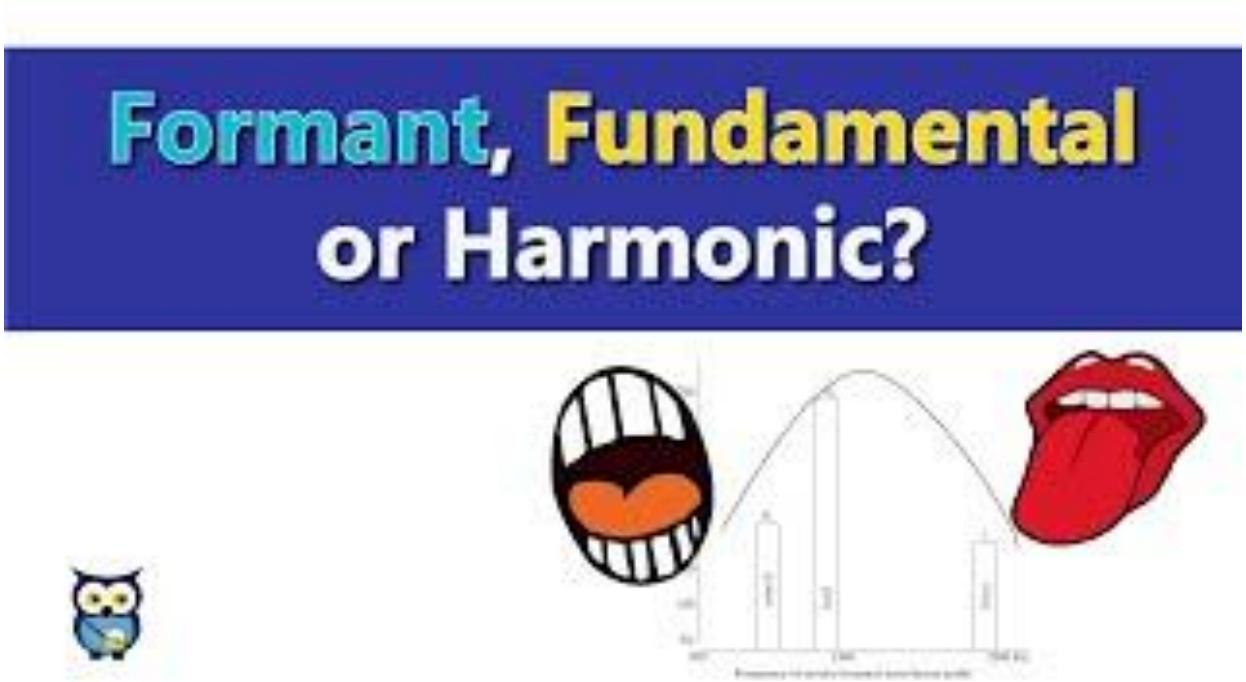
measured at temporal midpoint



- `draw visible formant contour`: draws to *Praat Picture* window
- `extract visible formant contour`: creates *formant object* in *Praat Object* window

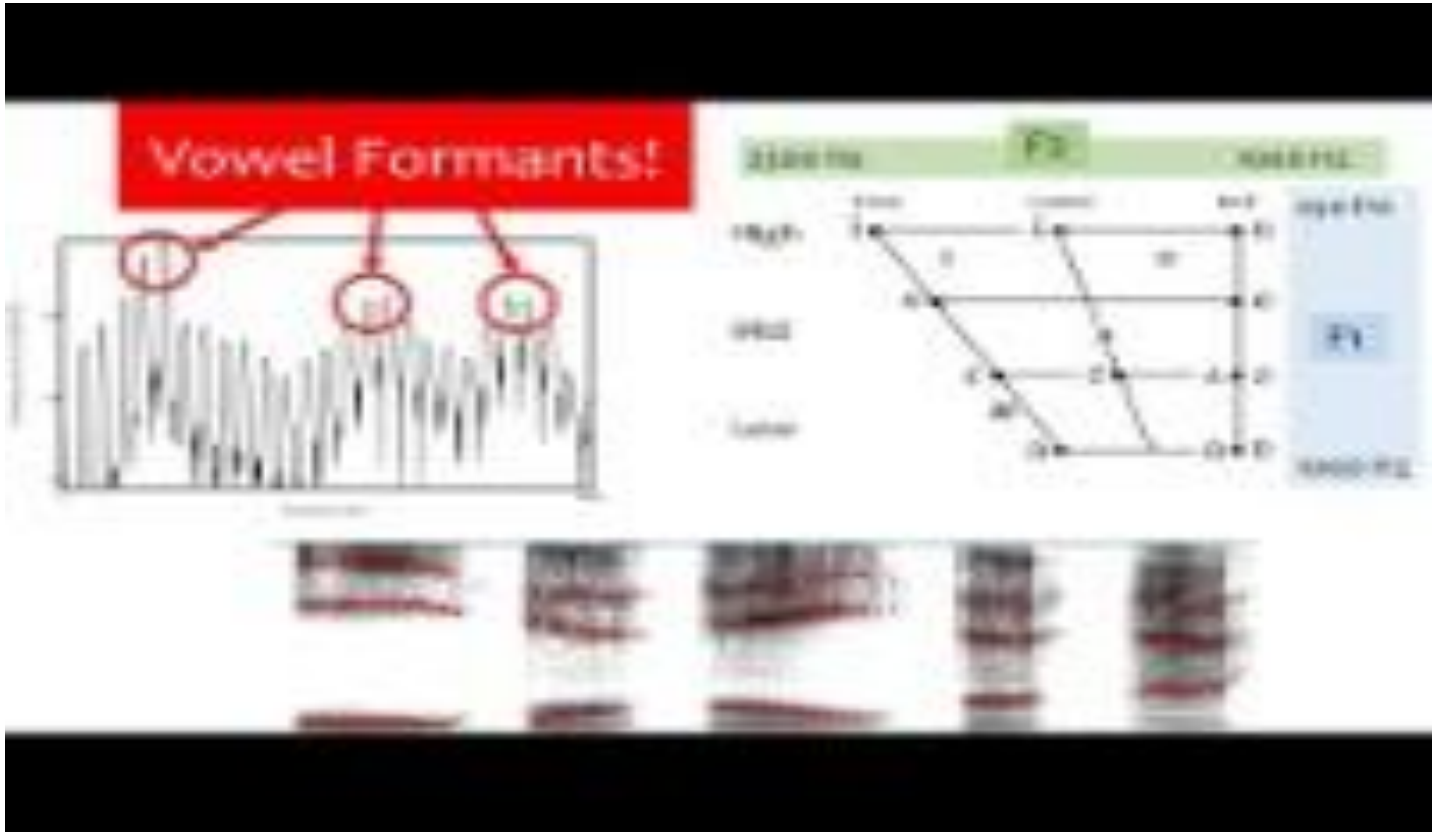
UNIVERSITÄT DES SAARLANDES

Resources: Formants, F0, Harmonics



https://youtu.be/PLzhqqNaEP0?t=43

Resources: Vowel Formant Measurements

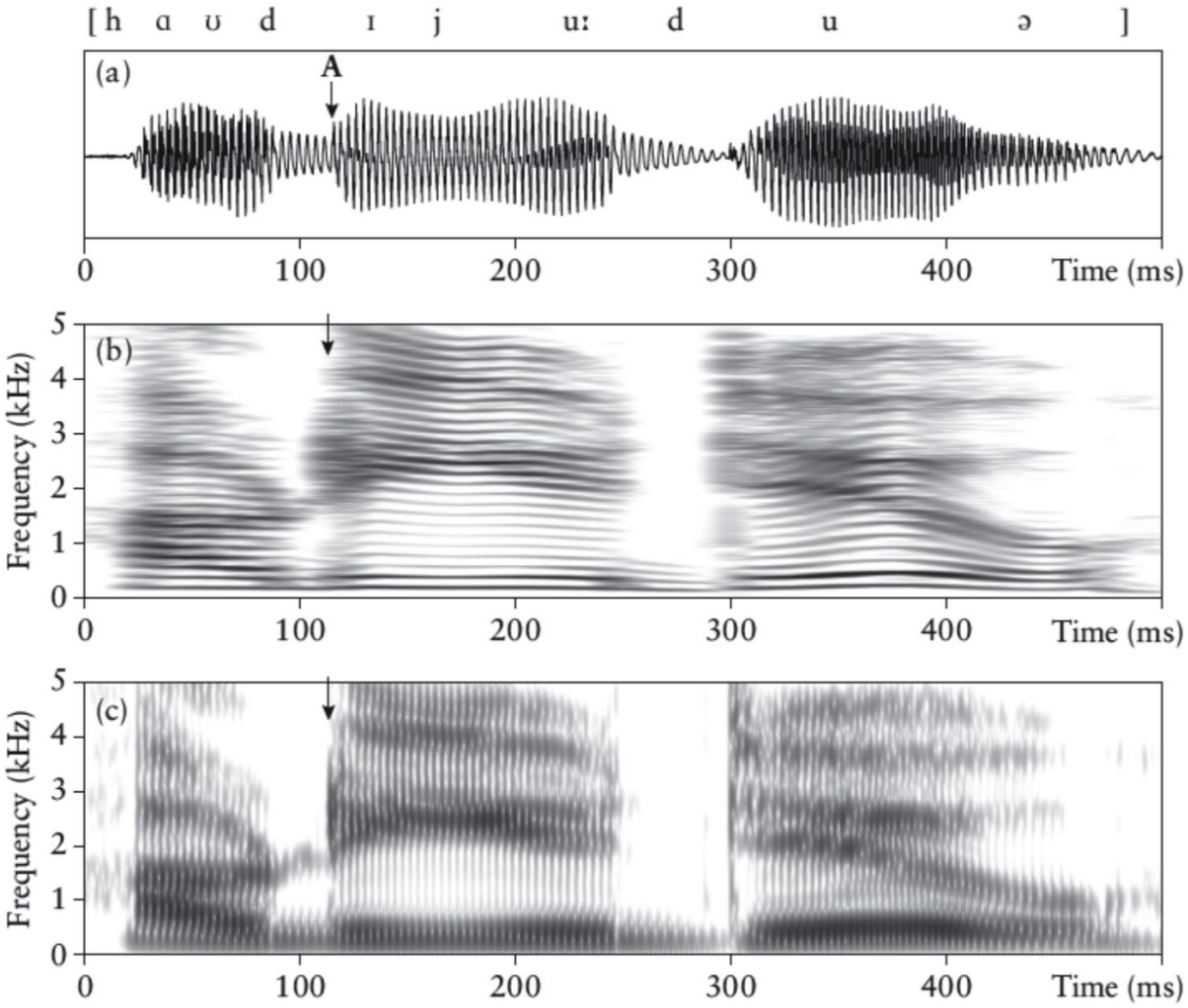


https://www.youtube.com/watch?v=gInUFa2fLyE&ab_channel=ListenLab

Resources: Vowel Formant Measurements

https://www.youtube.com/watch?v=5gVraRxer6o&ab_channel=LinguistiklaborAlbert-Ludwigs-Universit%C3%A4tFreiburg

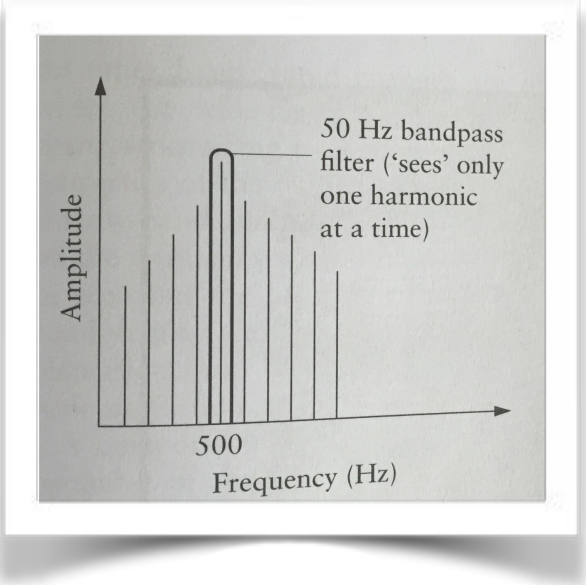
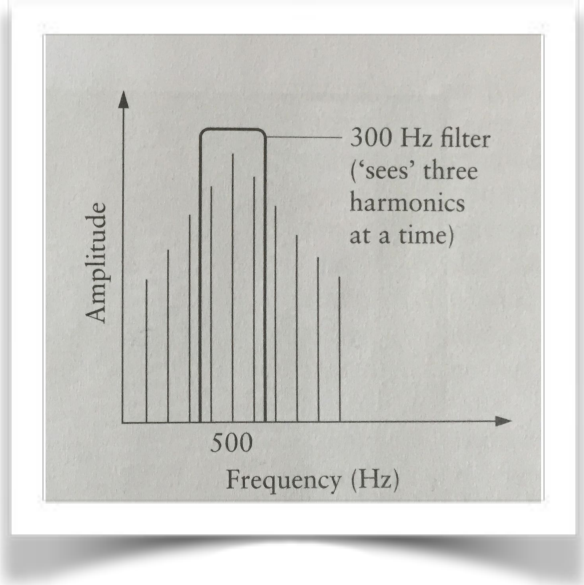
Wideband and Narrowband Spectrogram



narrow band spectrogram

wide band spectrogram

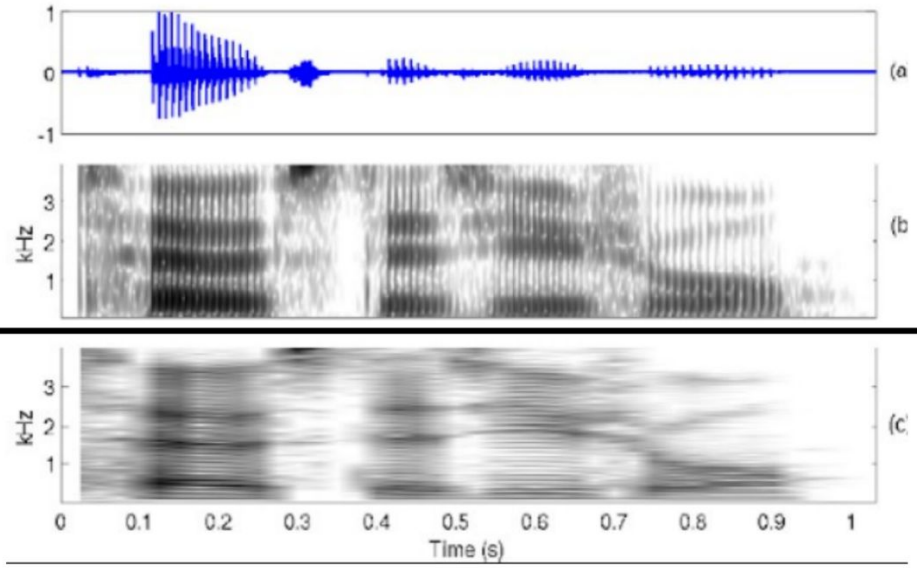
Wideband and Narrowband Spectrogram

spectrogram	**narrow band**	**wide band**
		
time	low resolution	high resolution
frequency	high resolution	low resolution
analysis	✓ harmonics ✓ *f*0 / intonation	✓ individual glottal pulses ✓ formant structure

Wideband and Narrowband Spectrogram

wide band spectrogram

Good temporal resolution



Poor frequency resolution

Poor temporal resolution

Good frequency resolution

narrow band spectrogram

Wideband and Narrowband Spectrogram Praat Settings

Intro 3.2. Configuring the spectrogram

With Spectrogram settings... from the **Spectrogram** menu, you can determine how the spectrogram is computed and how it is displayed. These settings will be remembered across Praat sessions. All these settings have standard values ("factory settings"), which appear when you click **Standards**.

View range (Hz)

the range of frequencies to display. The standard is 0 Hz at the bottom and 5000 Hz at the top. If this maximum frequency is higher than the Nyquist frequency of the Sound (which is half its sampling frequency), some values in the spectrogram will be zero, and the higher frequencies will be drawn in white. You can see this if you record a Sound at 44100 Hz and set the view range from 0 Hz to 25000 Hz.

Window length

the duration of the analysis window. If this is 0.005 seconds (the standard), Praat uses for each frame the part of the sound that lies between 0.0025 seconds before and 0.0025 seconds after the centre of that frame (for Gaussian windows, Praat actually uses a bit more than that). The window length determines the *bandwidth* of the spectral analysis, i.e. the width of the horizontal line in the spectrogram of a pure sine wave (see below). For a Gaussian window, the -3 dB bandwidth is 2*sqrt(6*ln(2))/(π**Window length*), or 1.2982804 / *Window length*. To get a `broad-band' spectrogram (bandwidth 260 Hz), keep the standard window length of 5 ms; to get a `narrow-band' spectrogram (bandwidth 43 Hz), set it to 30 ms (0.03 seconds). The other window shapes give slightly different values.

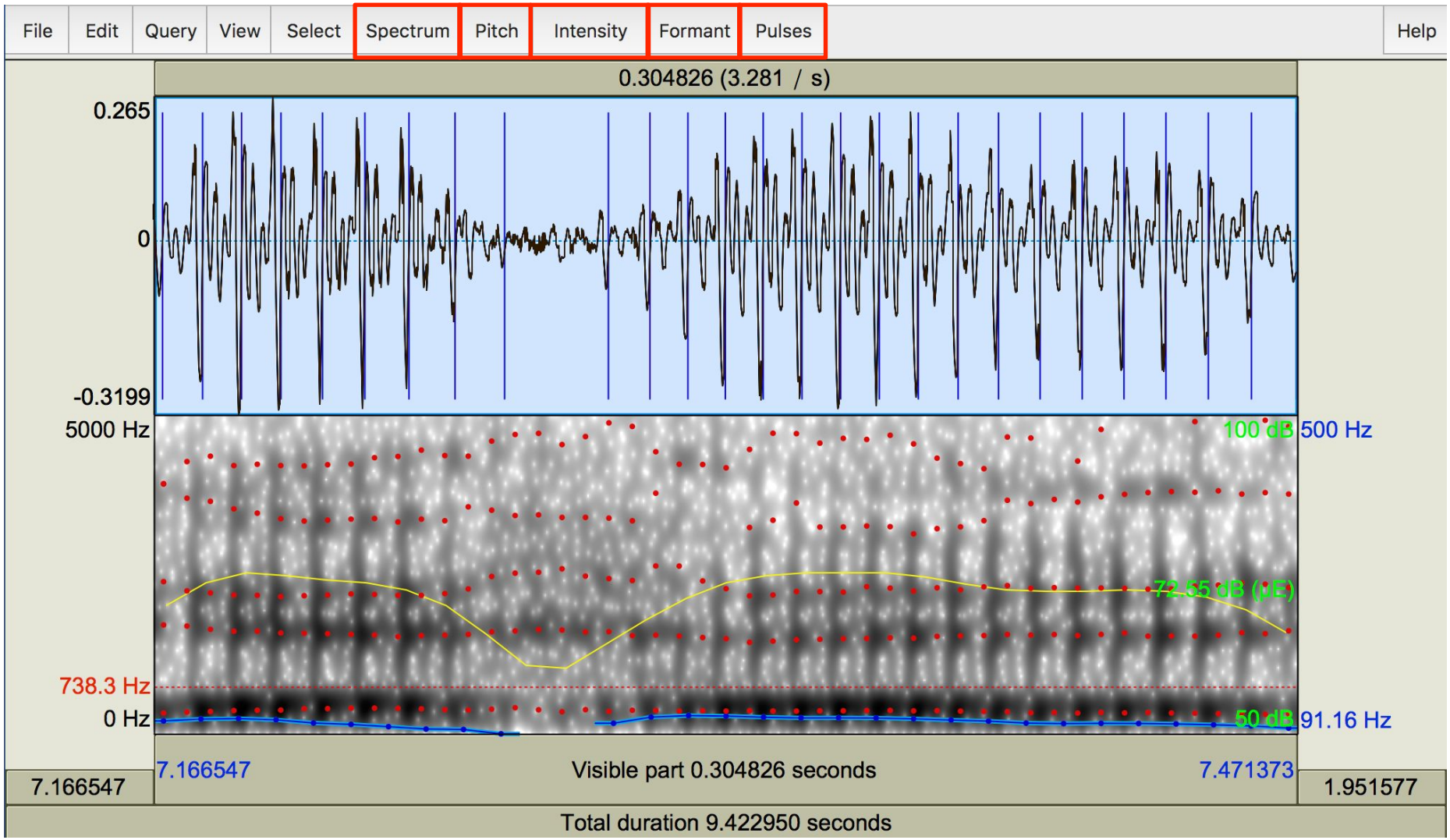
Dynamic range (dB)

All values that are more than *Dynamic range* dB below the maximum (perhaps after dynamic compression, see Advanced spectrogram settings...) will be drawn in white. Values in-between have appropriate shades of grey. Thus, if the highest peak in the spectrogram has a height of 30 dB/Hz, and the dynamic range is 50 dB (which is the standard value), then values below -20 dB/Hz will be drawn in white, and values between -20 dB/Hz and 30 dB/Hz will be drawn in various shades of grey.

https://www.fon.hum.uva.nl/praat/manual/Intro_3_2__Configuring_the_spectrogram.html

UNIVERSITÄT DES SAARLANDES

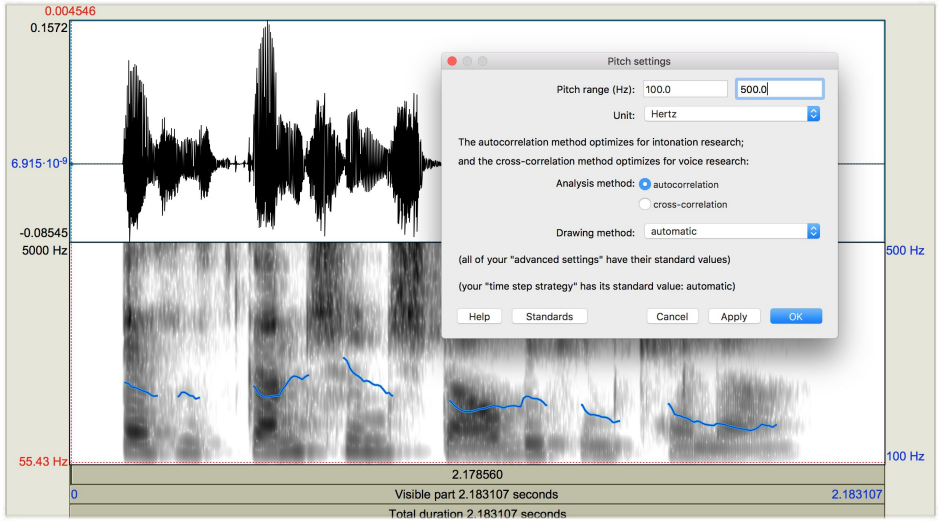
Acoustic analyses



UNIVERSITÄT DES SAARLANDES

Acoustic analyses: pitch

→ fundamental frequency (*f*0)



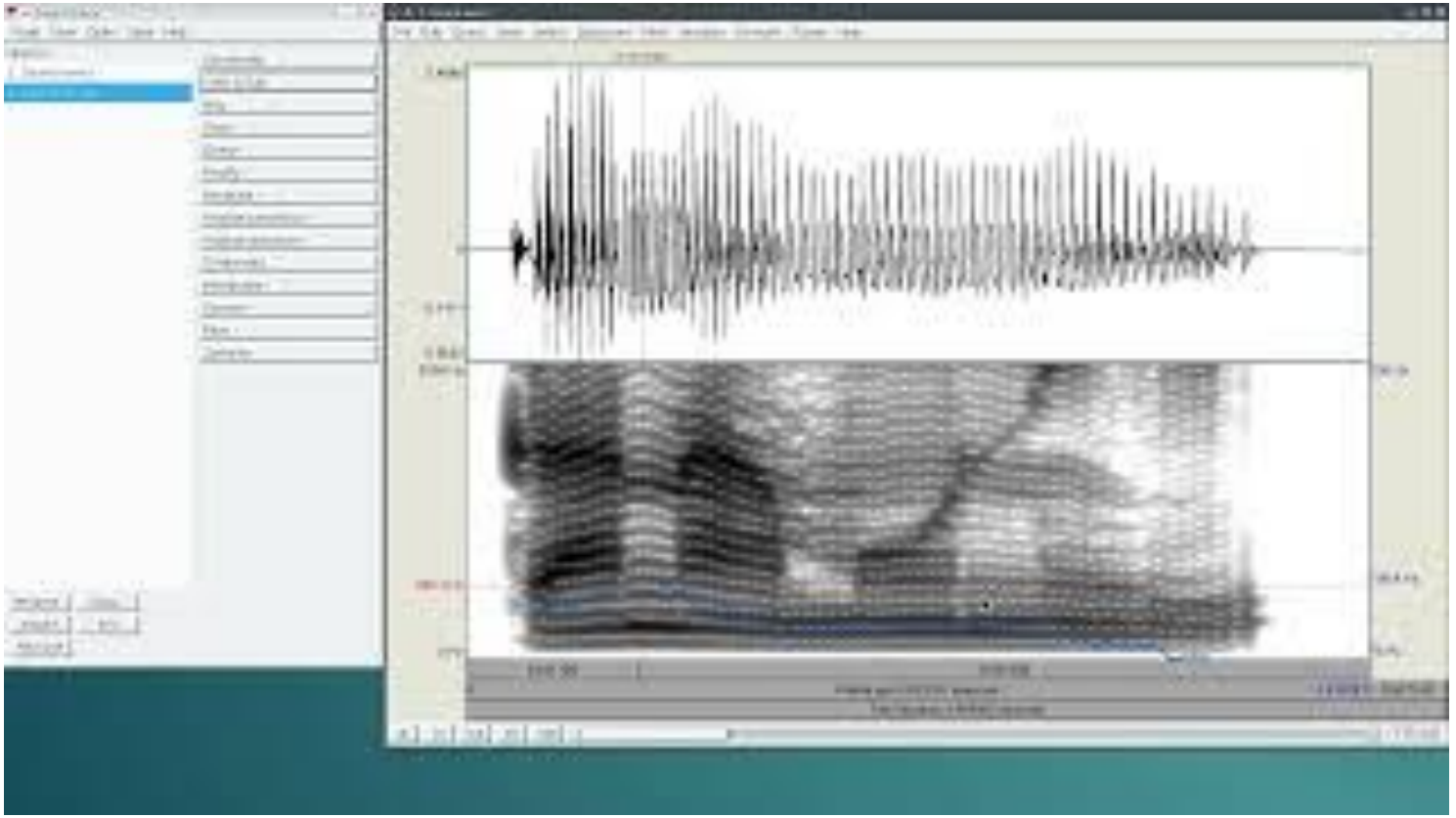
- calculated for voiced sounds only
- *pitch range* depends on speaker:
  - ♂ 75 - 300 Hz
  - ♀ 100 - 500 Hz

- `pitch listing`: *f*0 for all time stamps in selection
- `get pitch`: mean *f*0 in selection

→ mean *f*0 and *f*0 range in utterance

- `get minimum / maximum pitch`
- `draw visible pitch contour`: draws to *Praat Picture* window
- `extract visible pitch contour`: creates *pitch object* in *Praat Object* window

Resources: Measuring F0



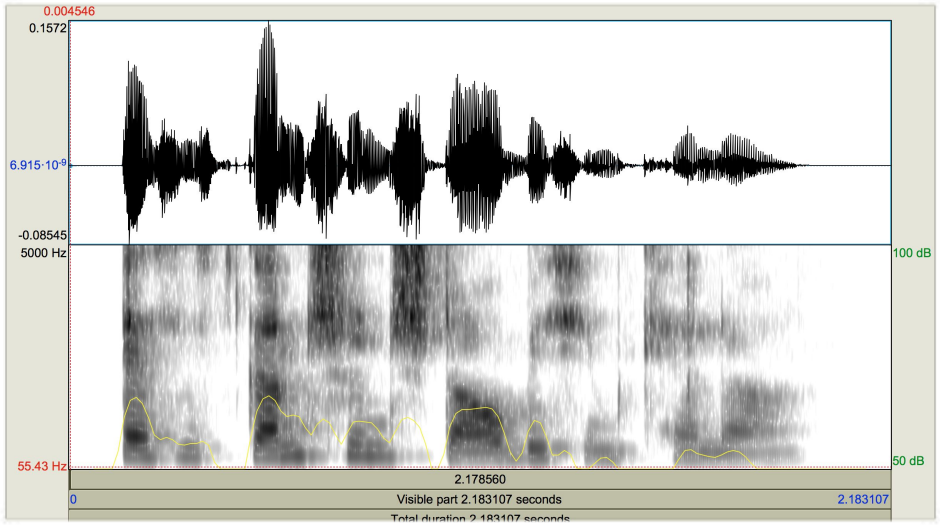
https://www.youtube.com/watch?v=qaMH5GeXdO0&list=PL9yHkZ12XpJYU1o3uySJ7-3R_4Z8Euqmn&index=10&ab_channel=StefanWerner

Acoustic analyses: intensity

- `intensity listing`: intensity for all time stamps in selection
- `get intensity`: mean intensity in selection

Some Common Decibel Levels	
140 dB	airplane takeoff (30m)
130 dB	hearing threshold of pain
120 dB	jackhammer, car horn (1m)
115 dB	iPod at peak volume, crying baby
110 dB	chainsaw (1m), airport
100 dB	rock concert, helicopter
90 dB	hair dryer, lawnmower
70 dB	vacuum cleaner (1m)
60 dB	normal conversation (1m)
40 dB	quiet room

→ mean intensity in utterance



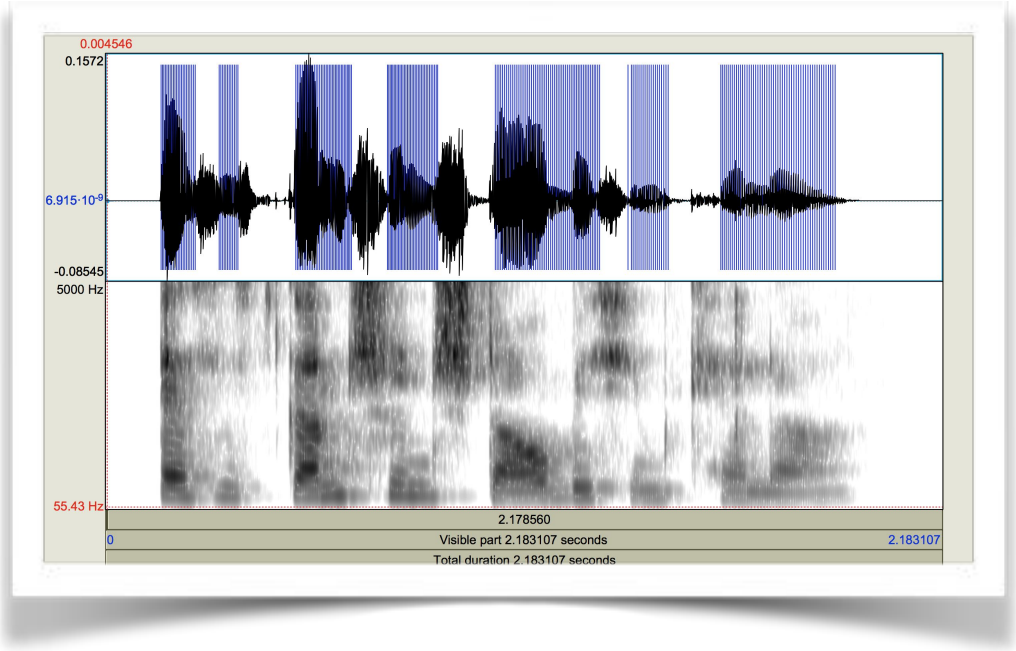
- `get minimum / maximum intensity`
- `draw visible intensity contour`: draws to *Praat Picture* window
- `extract visible intensity contour`: creates *intensity object* in *Praat Object* window

UNIVERSITÄT DES SAARLANDES

Acoustic analyses: glottal pulses

- `pulse listing`: time stamps of glottal pulses in selection

- `voice report`: information about voice quality (*pitch settings → analysis method → cross-correlation*)



- `draw visible pulses`: draws to *Praat Picture* window

- `extract visible pulses`: creates *point process object* in *Praat Object* window

UNIVERSITÄT DES SAARLANDES

Resources: Plots

https://www.youtube.com/watch?v=R0opWFiKelU&list=PL9yHkZ12XpJYU1o3uySJ7-3R_4Z8Euqmn&index=12&ab_channel=StefanWerner

Resources: Praat Experiments

https://www.youtube.com/watch?v=4XdVn_HTBAU&list=PL9yHkZ12XpJYU1o3uySJ7-3R_4Z8Euqmn&index=15&ab_channel=StefanWerner

Assignment : Extract vowel formants

From your recording of the north wind and sun, extract from the first 10 vowels:

- f0
- F1,F2,F3 (midpoint)