Basic acoustics part 2
Spectrograms, resonance, vowels

See Rogers chapter 7 8

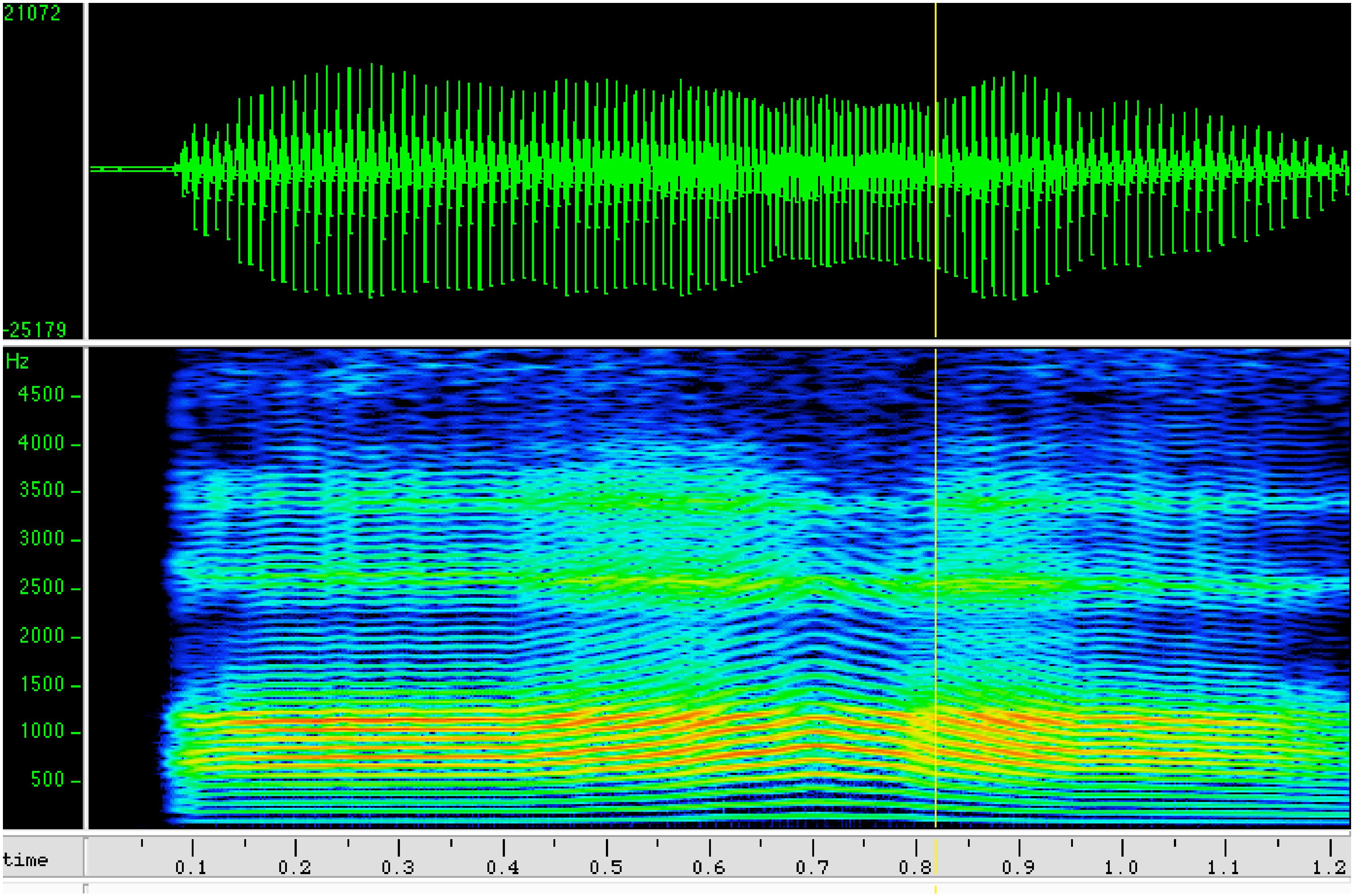
Spectrogram

- Spectrogram
- Represents ‘spectrum varying over time’
  - X-axis (horiz.) time (like waveform)
  - Y-axis Frequency (like spectrum)
  - Third dimension: pseudo-color or gray-scale representing amplitude

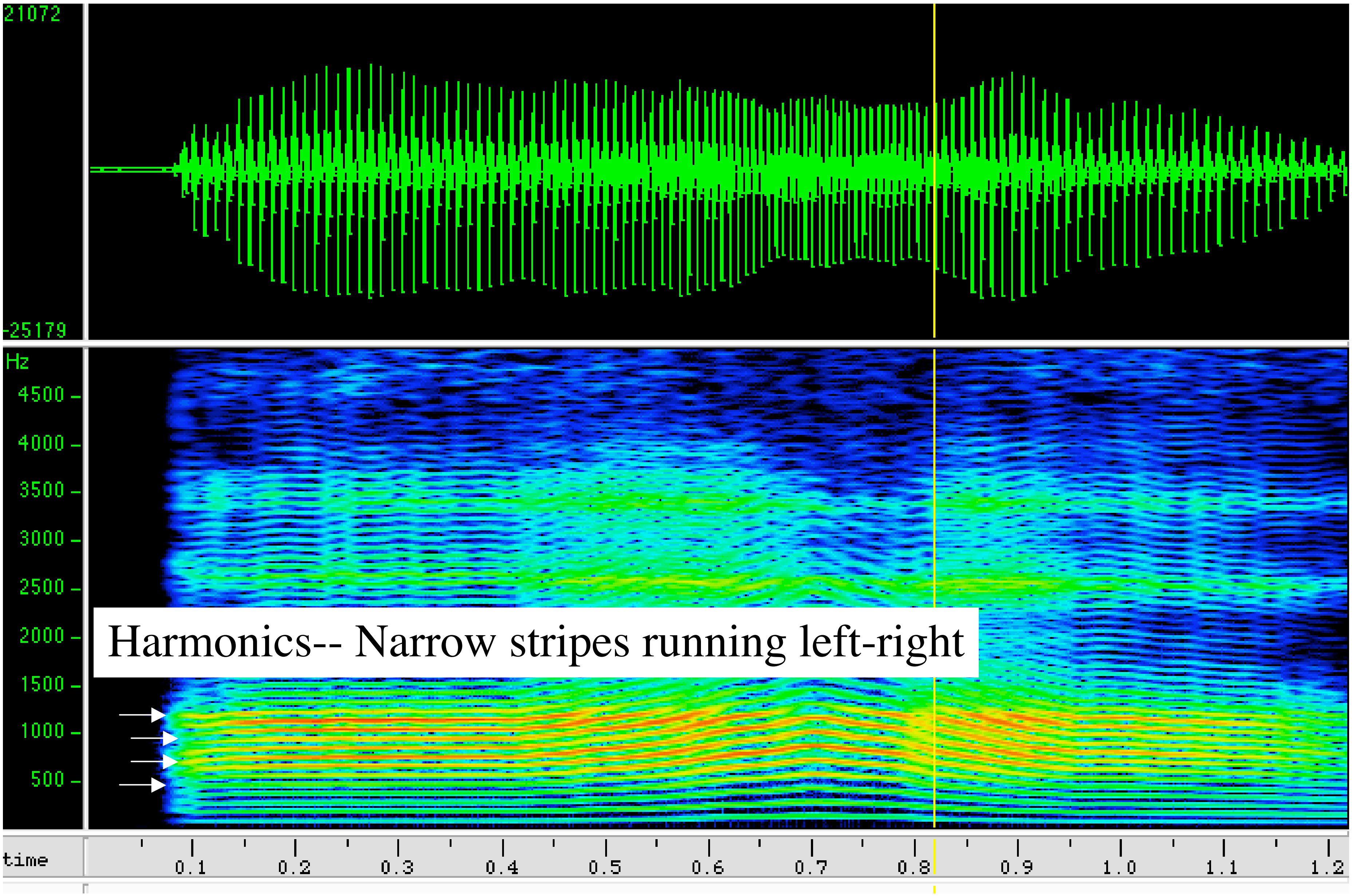
WaveSurfer….

- Allows us to see
  - Waveform
  - Spectrogram (color or gray)
  - Spectral section
    - ‘short-time spectrum’
    - = spectrum of a brief stretch of speech
- Demonstration spectrograms
  - of whistle
  - of speech

Narrow band spectrogram [aaa] pitch change



Harmonics [aaa] pitch change



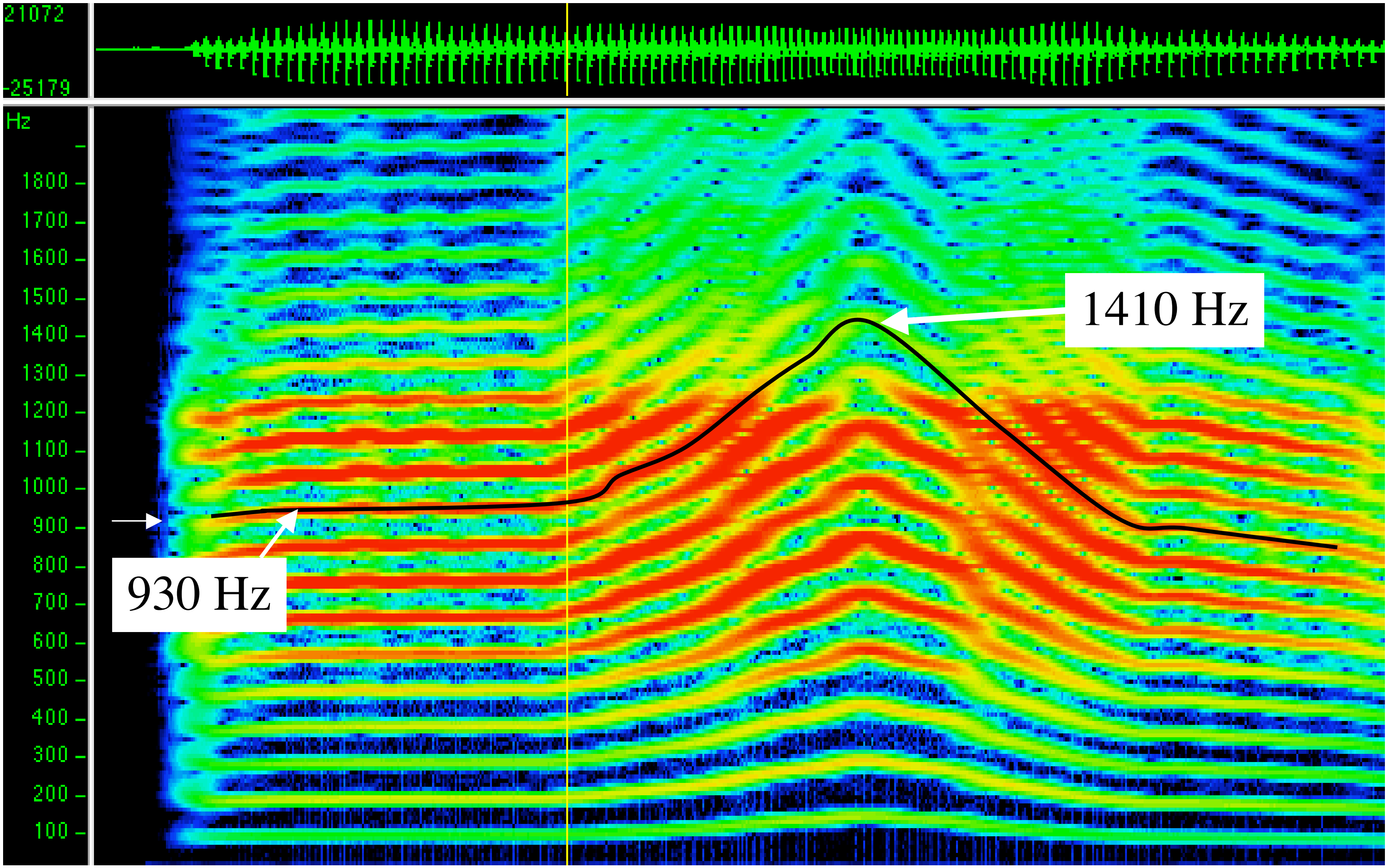
Narrow band spectrogram: changing pitch of [ɑɑɑ]

- Spectrogram of [ɑɑɑ] on varying pitches
- Narrow band spectrogram
  - Looks at fairly long stretch of time
    - 40 ms or so… sees several glottal pulses at once
      - Each glottal pulse about 10 ms long or less so several are blurred
    - Varying harmonic structure clear
  - Spectral sections at different times

Measuring F0 from narrow band spectrogram

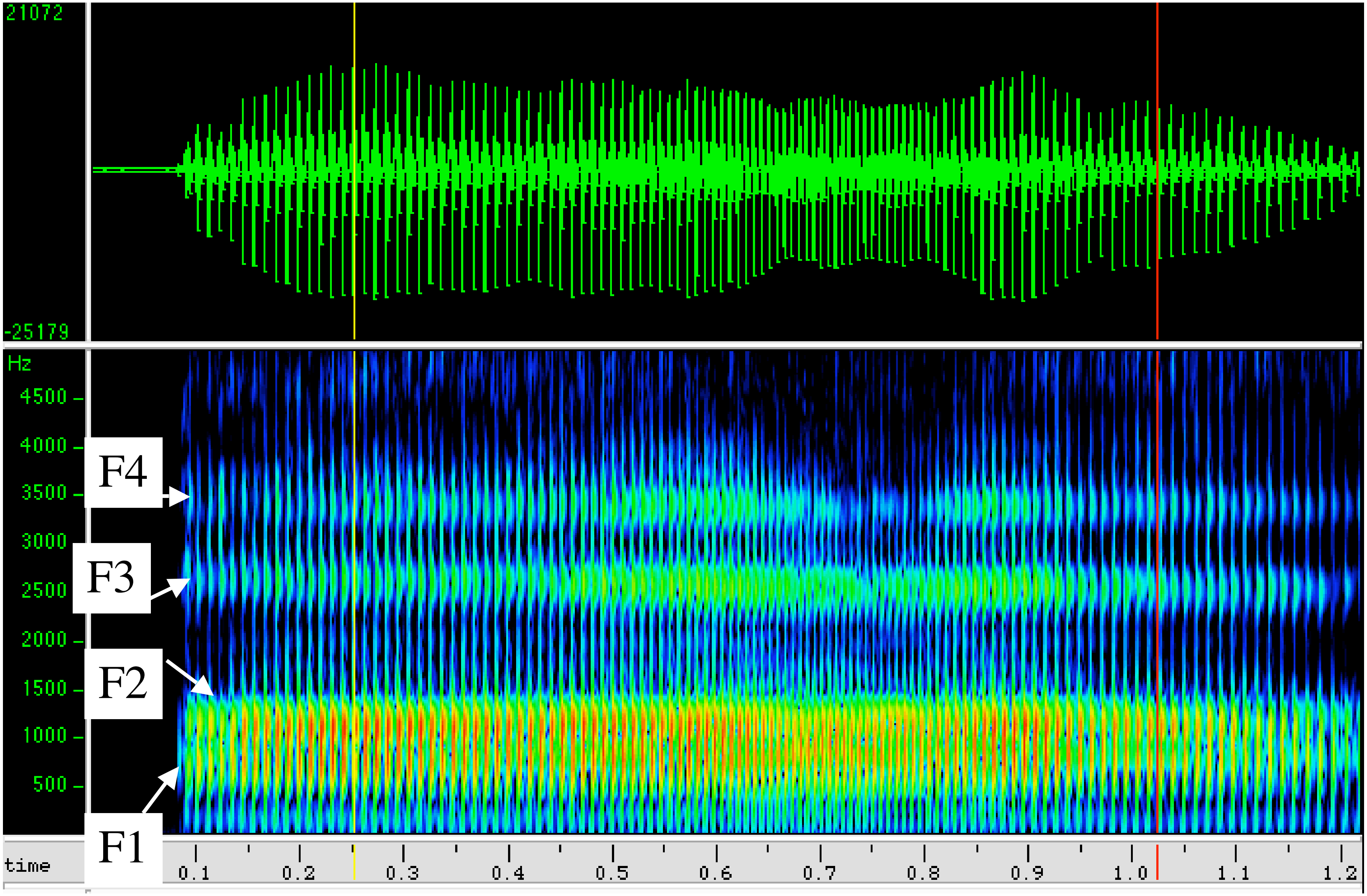
- Measure F0 from k-th harmonic
  - Hk = x Hz then F0= x/k Hz
    - 10th harmonic is convenient
- Expanding frequency scale makes this easier

Harmonics [aaa] pitch change: Freq. Expanded

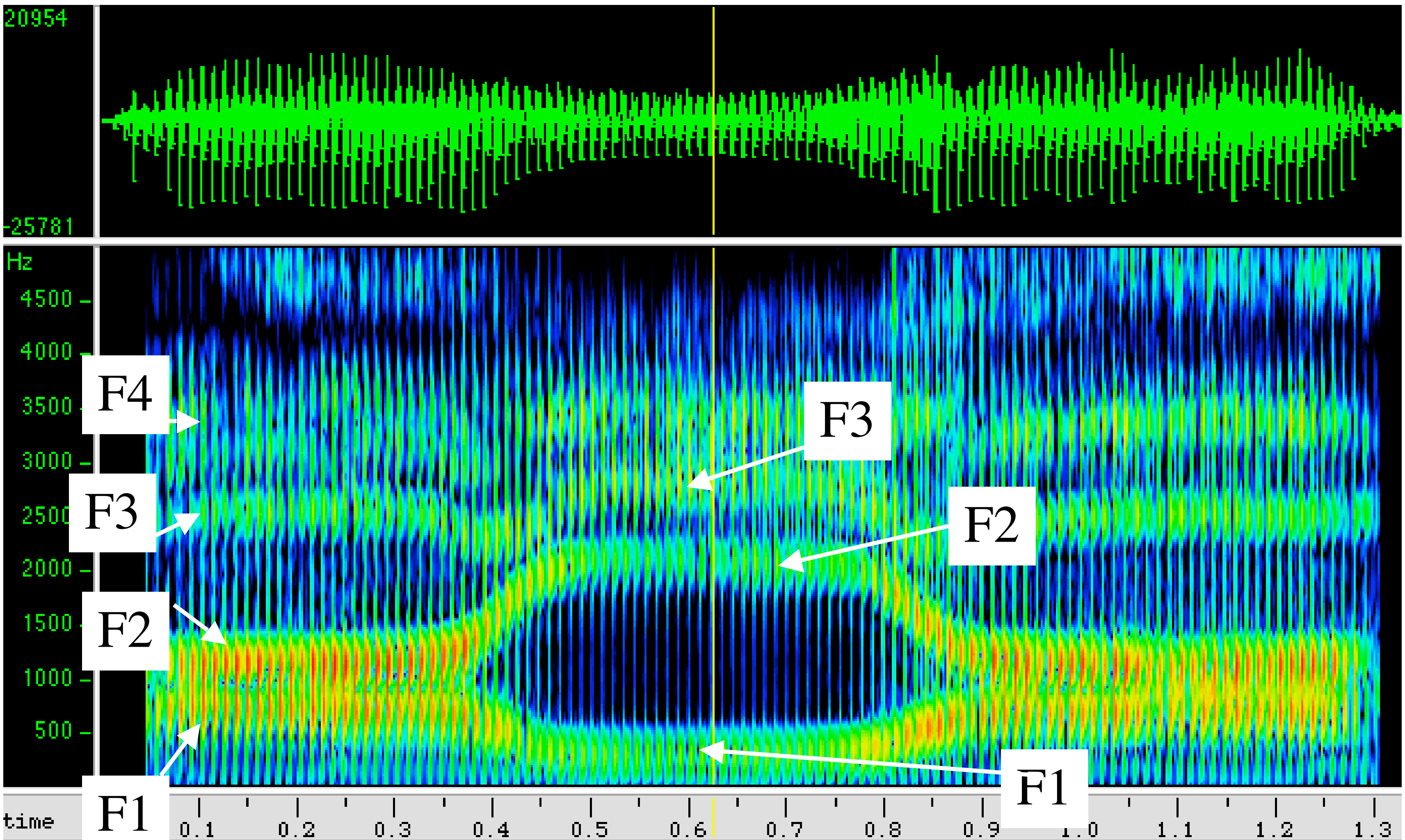


10-th Harmonic highlighted: F0 about 93 and 141 Hz at arrows

Wide band spectrogram [ɑɑɑ] pitch change



Wide band spectrogram [ɑɑiiɑɑ] No pitch change



Wide band spectrogram: changing pitch of [ɑɑɑ]

- Spectrogram of [ɑɑɑ] on varying pitches
- Wideband spectrogram
  - Looks at fairly long short of time
    - 2 to 3 ms only sees less than one full glottal period
    - Each glottal pulse about 10 ms long
  - Varying harmonic structure no longer clear
  - Dark bars show approximate location ‘formant peaks’
  - Formants don’t change much with pitch changes
  - They change lots with VOWEL changes

Wideband band spectrogram: [ɑɑiiɑɑ]

- Spectrogram of [ɑɑiiɑɑ] on SAME pitch
- Wideband spectrogram
  - Looks at fairly long short of time
    - 2 to 3 ms only sees less than one full glottal period
    - Each glottal pulse about 10 ms long
  - Harmonic structure no longer clear
  - Dark bars show approximate location ‘formant peaks’
  - Formants change lots with VOWEL changes

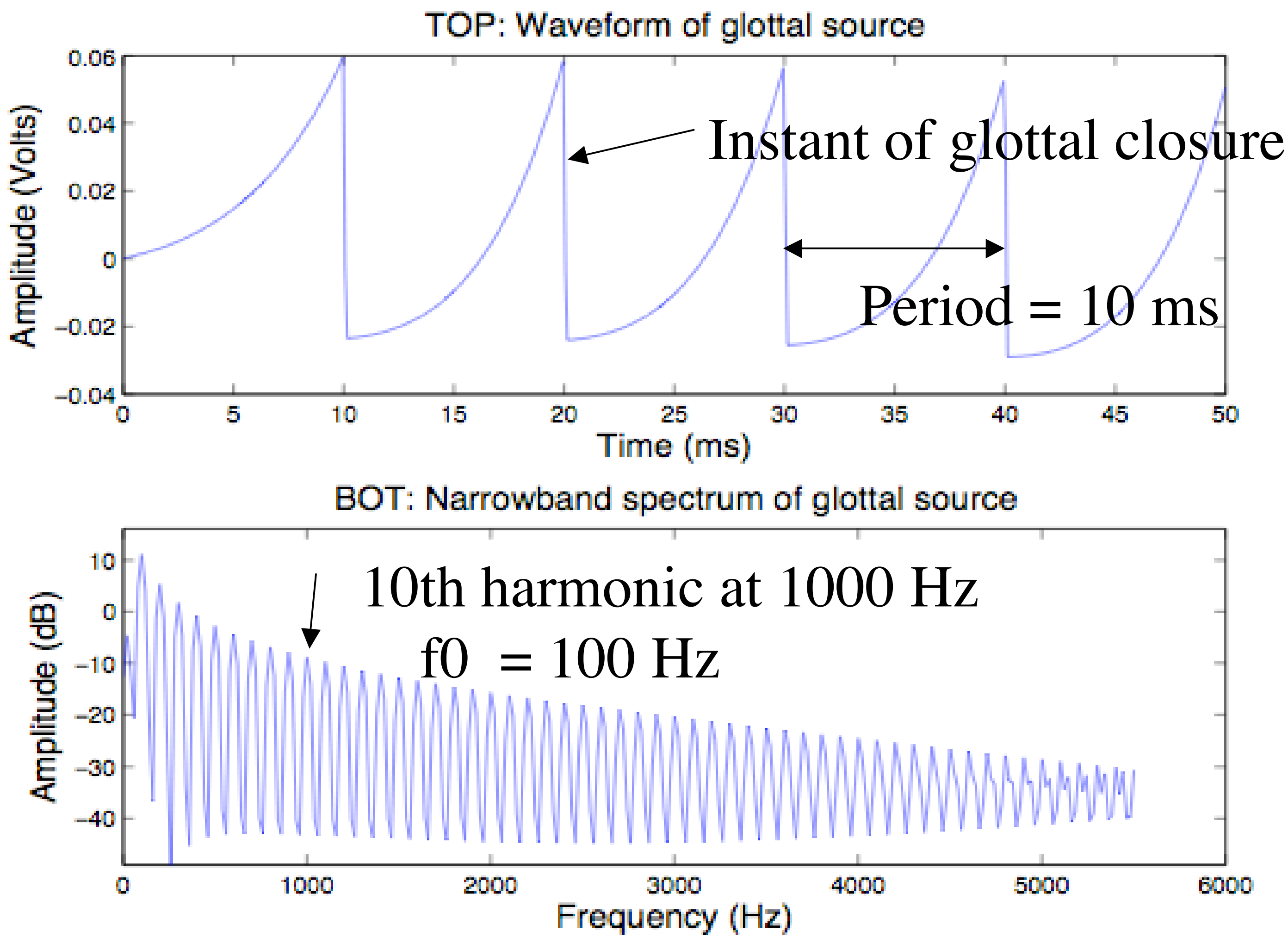
Wide and narrowband spectrograms

- Narrowband spectrogram makes **harmonic** structure clear
  - Associated with **glottal source**
- Wideband spectrogram makes **formant** structure clearer
  - Dark formant bands that change with vowel, not with pitch)
  - Formants associated 'filter properties' of vocal tract above the larynx

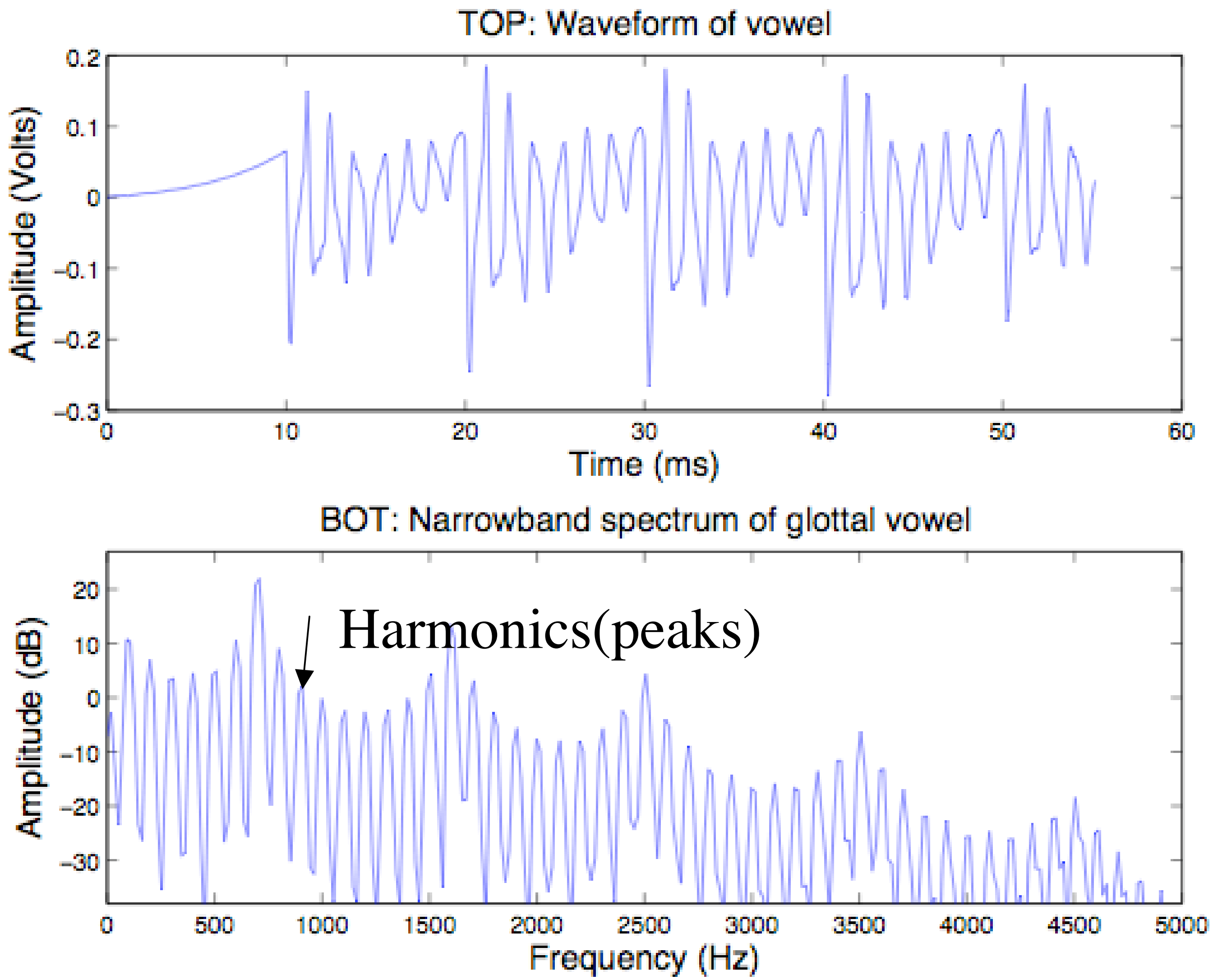
Source + Filter = Vowel

- Source + Filter theory of speech
- Consider vowel like sounds first
  - Source = voicing in glottis
  - Filter = tube-resonator system of SLVT
    - SLVT = supra-laryngeal vocal tract

Principle of source + filter : Glottal source



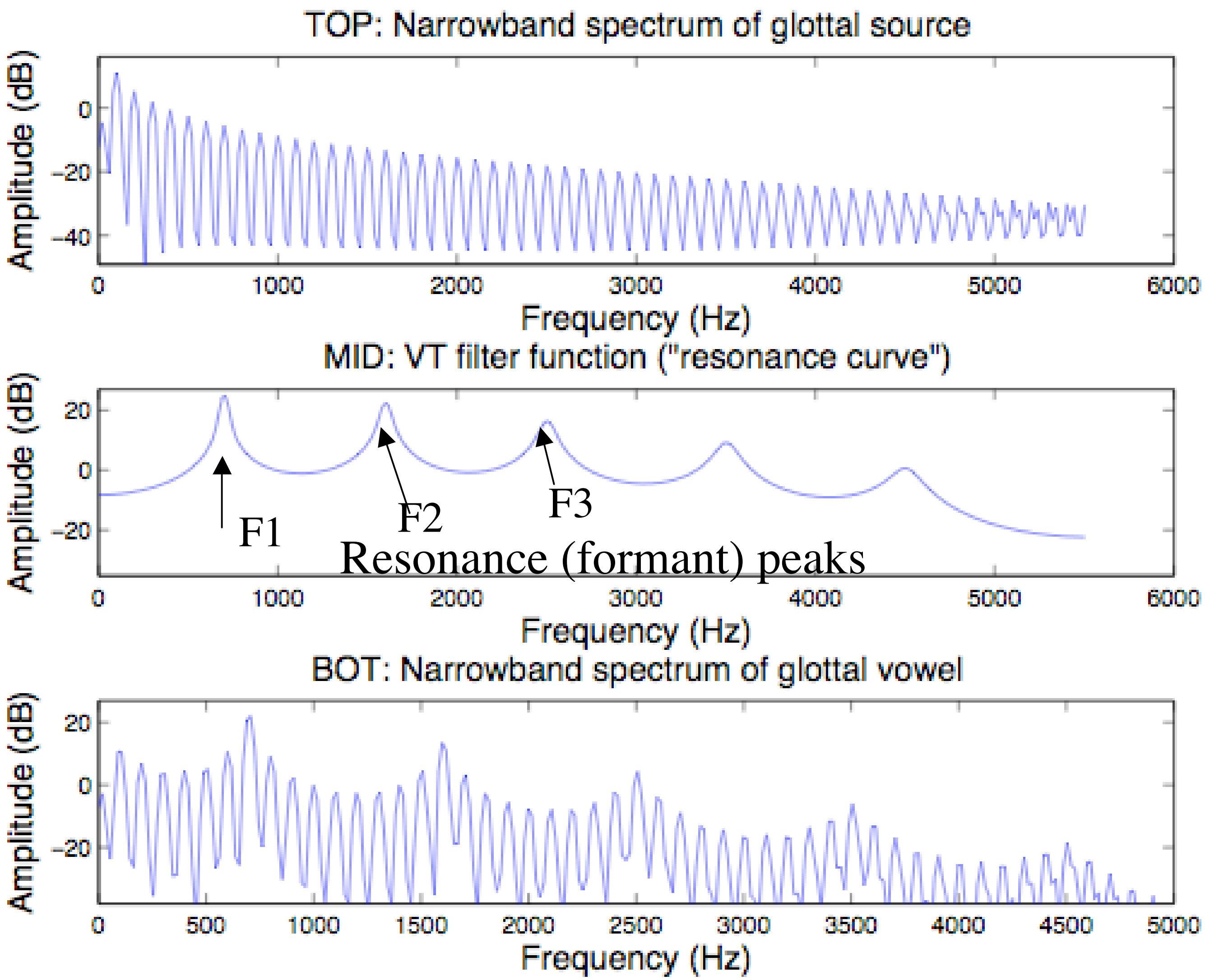
Principle of source + filter : Vowel



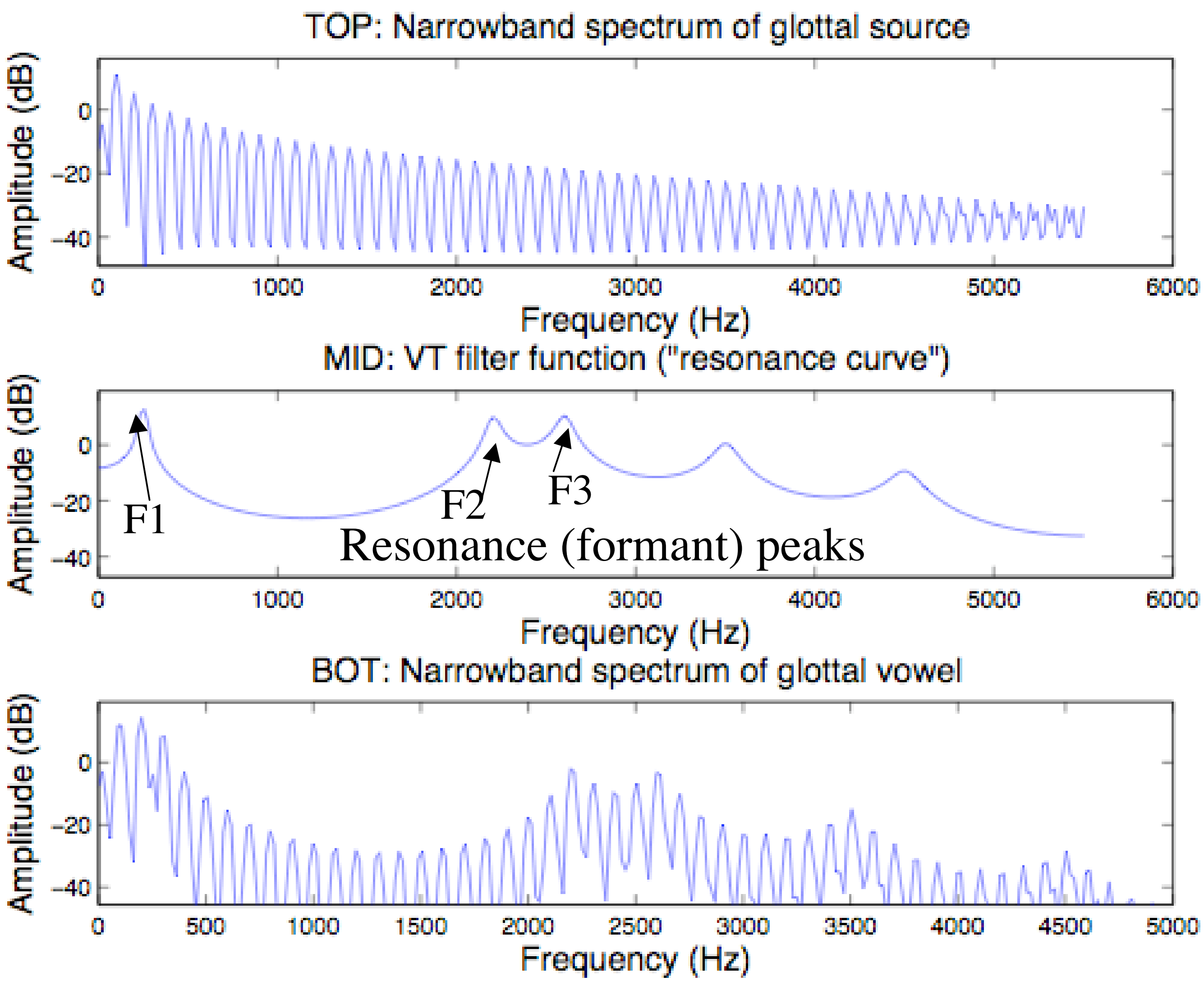
Compare last two slides

- Waveform and spectrum of glottal source are relatively simple compared to vowel
- SLVT filter imparts extra structure on vowel waveform
  - Oscillation between glottal pulses
  - Enhances (boosts) certain frequency regions

Source + Filter = Vowel [æ]



Source + Filter = Vowel [i]



Artificial glottal source

- Transformer ® 'robot voice'
- Replace glottal source with a simple buzz
- Use my SLVT as the filter

Spectrum of the Robot source

- Robot source:
  - Lots of harmonics across the frequencies
  - Ideally each harmonic would be near same amplitude
  - Note we see little ‘pointed pickets’ in spectral section
  - Not narrow lines
  - Real time-limited spectra look like this
    - As we increase time for a ‘steady’ signal we get more line-like harmonic peaks

What about filter?

- We’ve seen the robot source that can be filtered by real vocal tract
- Can we make a ‘robot filter’
  - Yes: Plastic tubes
- Slap them with palm of hand and get an ‘impulse response’ of filter

Robot vowels stage 1

- WaveSurfer analysis slapped tubes of different lengths
  - (ThreeTappedTubes)
- WaveSurfer … rapidly tapped mid-size tube
  - (TappedTubeEmpty.wav)
- WaveSurfer tapped tube with partial block
  - (TappedTubeBlock.wav)

Robot vowels stage 2

- WaveSurfer analysis slapped tubes of different lengths
  - (ThreeTappedTubes)
- WaveSurfer … rapidly tapped mid-size tube
  - (TappedTubeEmpty.wav)
- WaveSurfer tapped tube with partial block
  - (TappedTubeBlock.wav)

Robot vowels stage 3

- Add Robot source to tube
- Move ‘robot tongue’ to change shape
- WaveSurfer: robot /aaaiiiaaa/

Review: Displays

- Waveform
  - Time x amplitude
- Spectrum or spectral section
  - Frequency by amplitude (dB)
- Spectrogram
  - Time by frequency by amplitude
  - (horiz.) (vert.) (color or darkness)

Waveforms:

- Waveform
  - Time x amplitude
- Good for measuring durations of some events (especially when displayed with spectrogram).
  - Period of a repetitive waveform (e.g. glottal pulse duration of voiced speech)
  - VOT

Spectral section

- Spectral section (spectrum)
  - Frequency by amplitude in a brief interval of time (a section of a longer signal)
- Narrow band spectra look at moderately long chunks of speech (30-40 ms)
  - Show harmonics for voiced speech
- Broad band spectra look at shorter chunks of speech (less than glottal period)
  - can show formant structure

Narrow band spectrogram

- Narrow band spectrogram is a way to display **many narrow-band spectral sections at once**
- At each point in time, look at moderately long chunks of speech (30-40 ms) centered on that time point ('windowed' sections)
  - Represent amplitude at each frequency for that center time by darkness or color coding
- Shows harmonics as horizontal bands that bend as fundamental frequency changes
- Formant patterns visible only indirectly by which harmonics are strong

Measuring F0 from waveforms

- Find duration of one period
- Convert period duration to frequency
    - 1 period in .005 seconds (= 5 ms) that means
    - = X periods in 1 second?
    - Answer 1/ .005 = 200 Hz
- Alternate method: count several periods (k)
    - x periods in x sec means frequency of k/x Hz
    - That is k/x periods occur in one second

Measuring F0 from wide band spectrogram

- Find duration of one period
  - Distance between vertical striations (stripes)
- Proceed as with waveform
- "Ballpark" method for average F0:
  - Count number of striations in 100 ms and multiply by 10

Measuring F0 from narrow band spectrogram or spectral section

- Count up to the 10th harmonic
- Measure its frequency against the frequency scale
- Divide by 10
  - Can be very accurate
  - Can use harmonic number k (instead of 10) if that's easier to find
  - Then divide by k

Measuring Formants

- Use wide band spectrogram
- Try to identify wide bars that move a bit up and down
- Measure the center frequency of ‘darkest’ or ‘redest’ part.

Note: I will provide ‘formant tracks’ from WaveSurfer which will put thin lines through the formants