LING 450/550
5 – Spectrograms 1

Power Spectra and Spectrograms

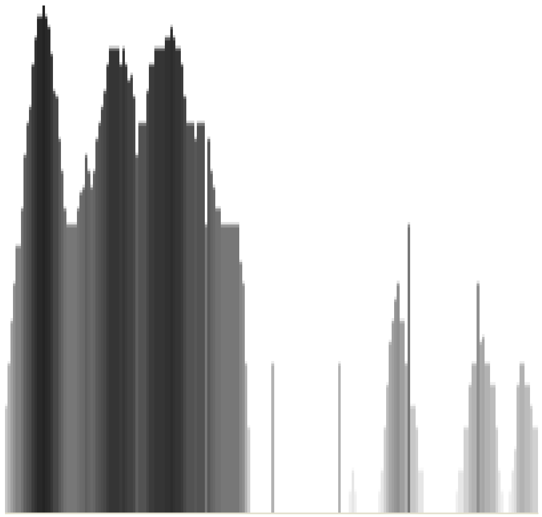
Power spectra allow us to see the component frequencies and amplitudes that make up a complex wave.

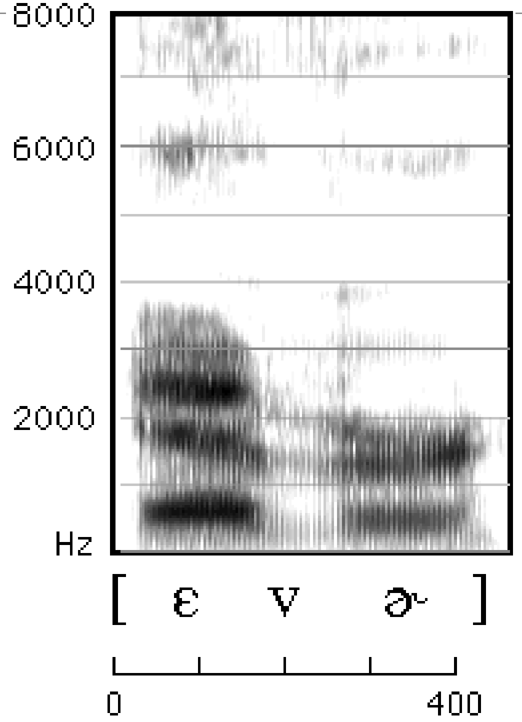
Spectrograms are a way of viewing hundreds of sequential power spectra, so we can see how the spectrum of a speech signal changes over time.

- This is done by turning the spectrum sideways, reducing each bar on the spectrum to a dot, and using the color of the dot to symbolize the height of the bar (i.e., the intensity of the component frequency).

Building a Spectrogram

1. Power spectrum amplitudes are represented as different shades of grey (taller = darker)
2. Spectrum is rotated and collapsed to 1-dimension.
3. Hundreds of spectra are arrayed side by side, allowing us to see changes in formant frequencies over time.





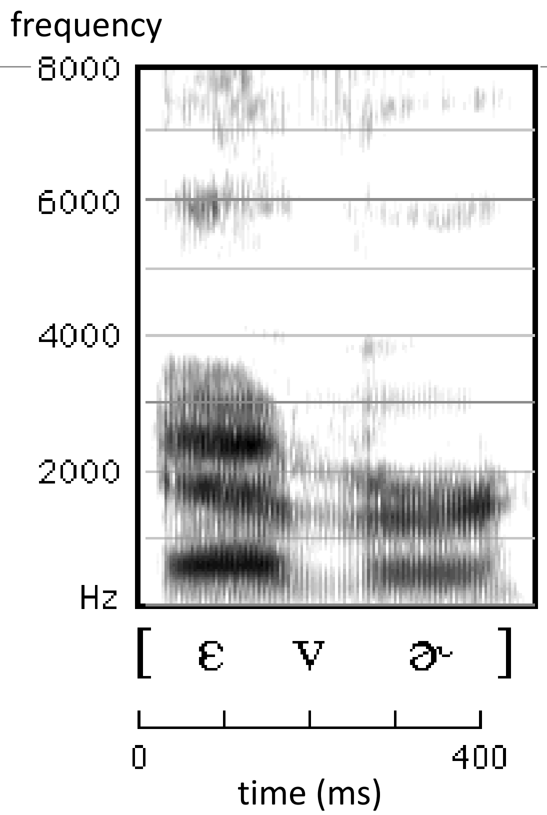
Reading a Spectrogram

A spectrogram plots three variables:

- The horizontal axis is *time*, usually measured in milliseconds.
- The vertical axis is *frequency* (Hz).
- The third variable, *intensity*, is represented by the color of each point that is plotted.

The dark bands show the *formants* of the voiced sounds.

- The first three formants vary depending on the shape of the vocal tract and are important for understanding the speech signal.



Narrowband Spectrograms

When calculating the frequency components that make up the speech signal, it is necessary to take a chunk of time to figure out the frequency of the wave.

Narrowband spectrograms show calculations based on longer chunks of time, which allows more accurate calculation of the *frequencies* making up the signal and can reveal individual *harmonics* of the signal.

- Because the chunks of time are longer, narrowband spectrograms are less precise about when a given speech event happened (e.g., a release burst or the onset of voicing). This lack of precision in the temporal dimension is called *time smear*.
- Because of their detailed frequency information, narrowband spectrograms are useful for investigating changes in pitch, tone, and intonation, all aspects of the *source*.

Narrowband Spectrograms

[ɑbɑ]

- Frequency range: 0 - 5000 Hz; duration: ≈0.75 s
- Window size: 0.03 s

Wideband Spectrograms

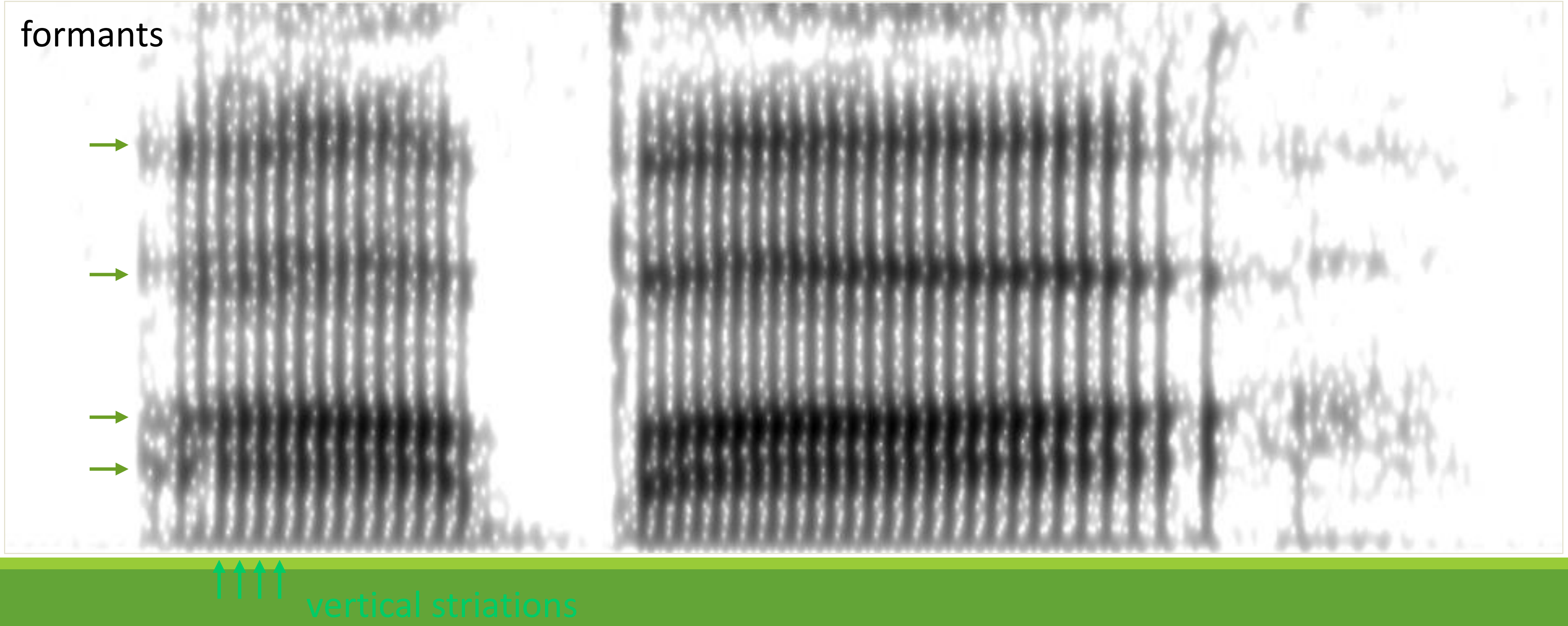
Wideband spectrograms show calculations based on short chunks of time, which allows accurate measurement of event timing at the expense of detailed frequency information. Wideband spectrograms are vertically *striated*; each vertical striation shows an individual vibration of the vocal folds.

- Remember that formants are *broad peaks* of intensity that range over several neighboring frequencies. Since the wideband spectrograms "smear together" neighboring frequencies, this actually helps us to see the formants more clearly.
- Vowel quality, place of articulation, and other segmental features are all aspects of the *filter* and thus cause broad-frequency changes to the signal. Thus they are best viewed using wideband spectrograms.
- Because of their accurate timing calculations, wideband spectrograms are also useful for measuring VOT, vowel duration, and other time-based characteristics of speech.

Wideband Spectrograms

[ɑbɑ]

- Frequency range: 0 - 5000 Hz; duration: ≈0.75 s
- Window size: 0.005 s



Reading Spectrograms

READ LADEFOGED & JOHNSON, CHAPTER 8 (PP. 208-225)

adapted from slides by Richard Wright, Dan McCloy, and Valerie Freeman

Spectrogram Review

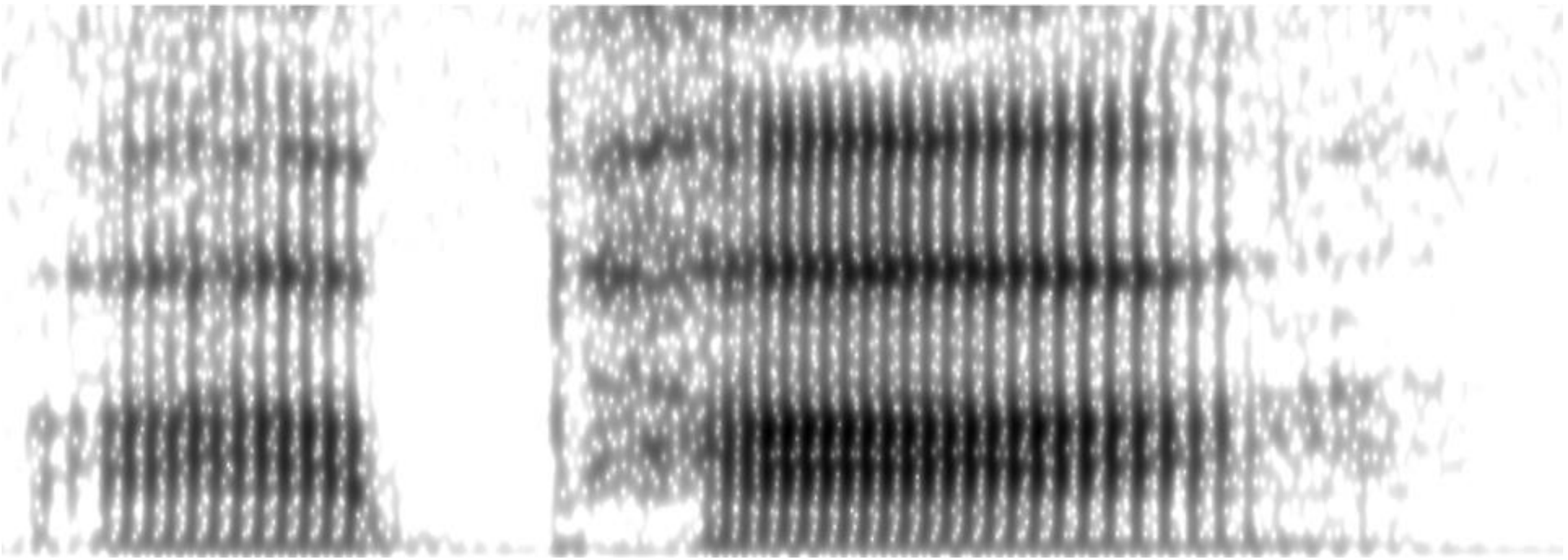
Remember that spectrograms tell us about the intensity of different frequencies over time.

When we change the vocal tract filter, it changes the way air resonates in the vocal tract. We can see the acoustic effects of these changes by examining spectrograms.

Reading Spectrograms: Oral Stops

[ɑpɑ]

- Frequency range: 0 - 5000 Hz; duration: ≈0.75 s



Reading Spectrograms: Oral Stops

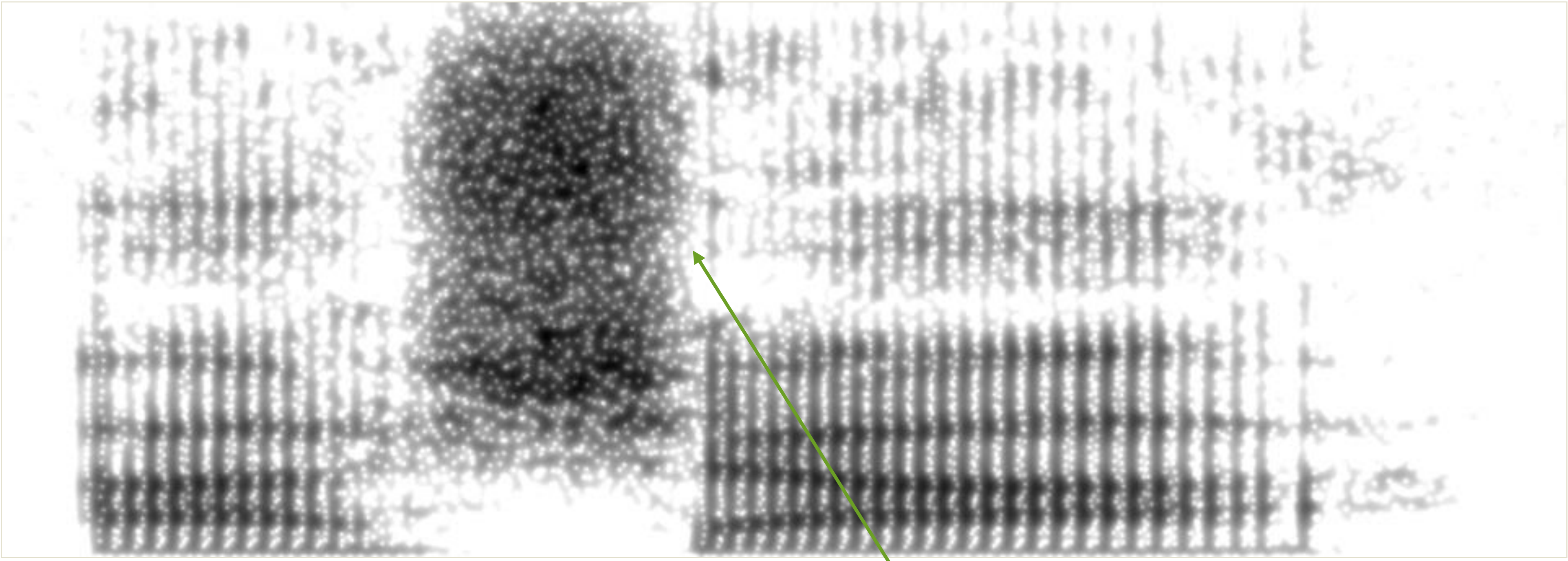
Oral stops involve complete stoppage of airflow and thus show up as a “gap” in the spectrogram.

- The formant transitions of the voiced sounds before or after a stop are the best way to determine the stop’s place of articulation.
  - More on this later.
- The release burst of a stop (if present) may be continuous with a following voiced sound, or may stand alone at the end of a word as a brief vertical band of high energy. The distribution of that energy can also provide clues to the place of articulation of the stop.

Reading Spectrograms: Fricatives

[ɑʃɑ]

- Frequency range: 0 - 10,000 Hz; duration: ≈0.75 s



high energy in the higher frequencies

Reading Spectrograms: Fricatives

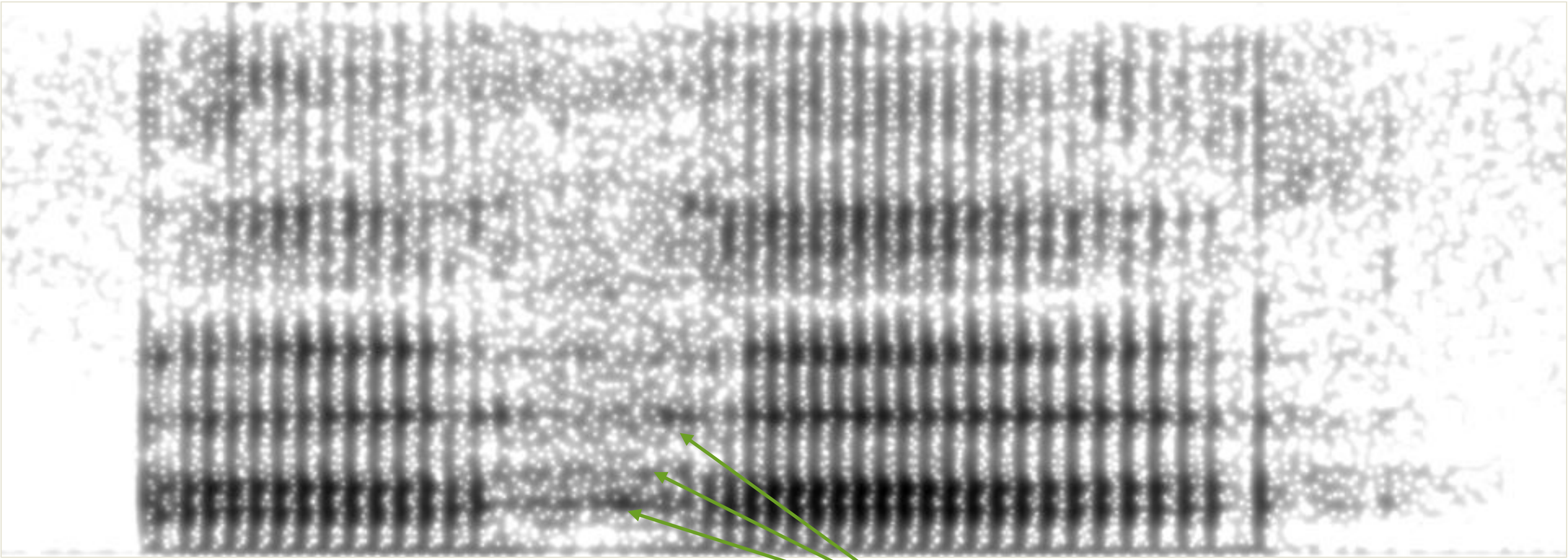
Fricatives are characterized by fairly high energy (darker color) especially in the higher frequencies, and (usually) a lack of formant bands.

- Voiced fricatives have less energy (are lighter in color) than voiceless fricatives because air is only flowing through the constriction when the vocal folds are open.
- [h], a voiceless glottal fricative, often does show formant bands. This is because [h] generates turbulence in the same place that voicing originates (the glottis), so the turbulence of [h] is subject to the same filter that vowels are (namely, the entire vocal tract). The spectrum of [h] can vary widely depending on the following vowel.

Reading Spectrograms: Fricatives

[ɑhɑ]

- Frequency range: 0 - 10,000 Hz; duration: ≈0.75 s

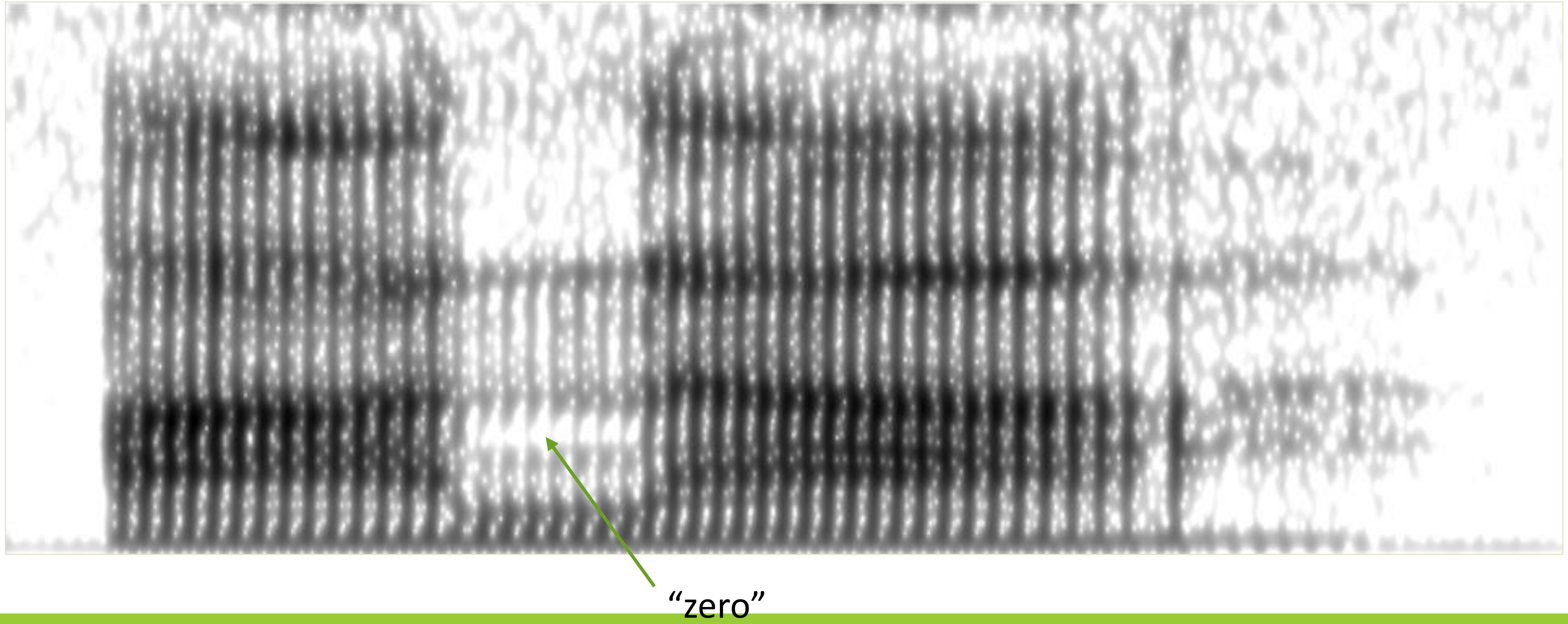


formant bands visible through the [h]

Reading Spectrograms: Nasal Stops

[ɑnɑ]

- Frequency range: 0 - 5000 Hz; duration: ≈0.75 s



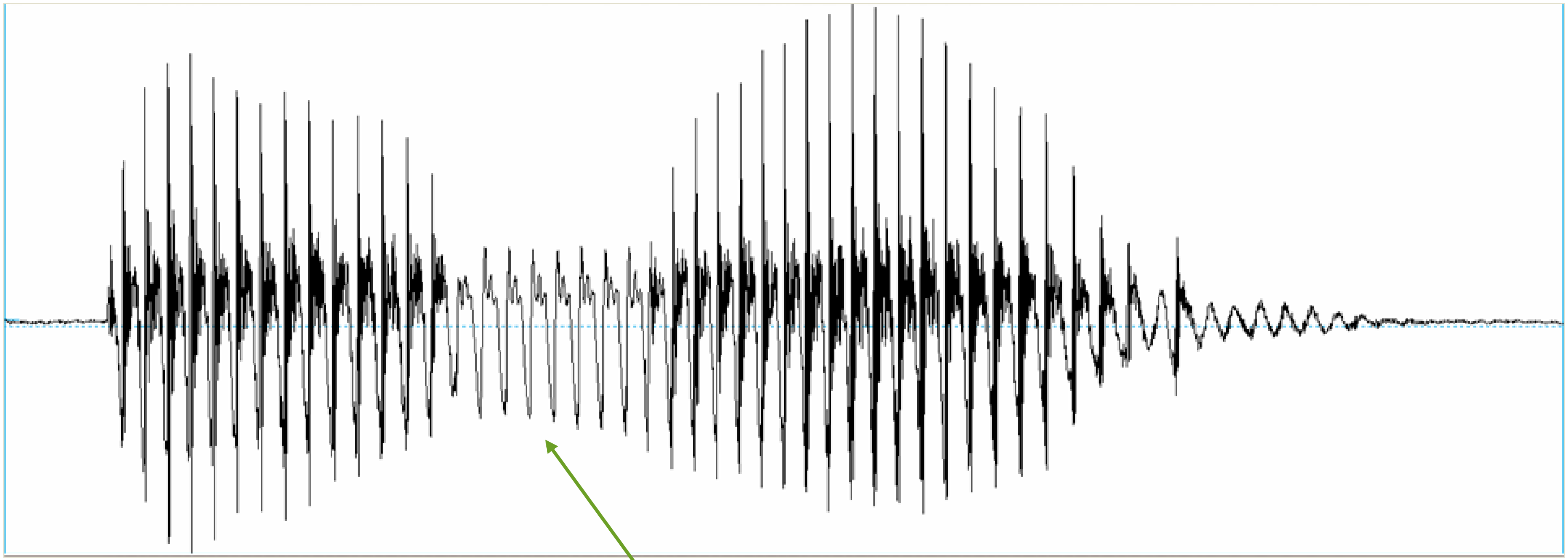
Reading Spectrograms: Nasal Stops

Nasal stops allow continuation of airflow and show formant structure much like vowels.

- Like oral stops, nasal place of articulation is most easily determined by the formant transitions of the surrounding vowels.
  - Again, more on this later.
- Unlike vowels, nasals typically show “zeroes” (white areas) between the lower formants.
  - These are sometimes called *antiformants*.
- Nasals can be distinguished from vowels by a much lower amplitude in the waveform, and little high-frequency energy.
  - The walls of the nasal passage attenuate (absorb) much of the energy of these sounds.

Reading Spectrograms: Nasal Stops

[ɑnɑ]

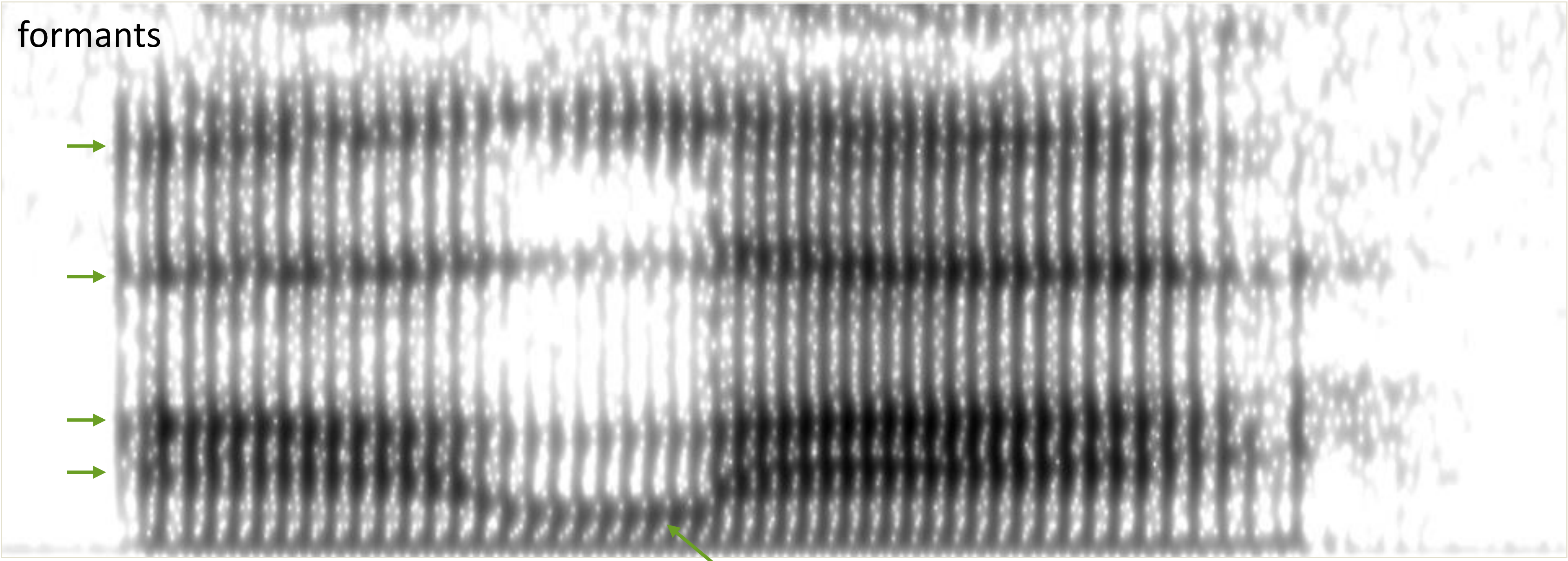


lower amplitude

Reading Spectrograms: Approximants

[ɑlɑ]

- Frequency range: 0 - 5000 Hz; duration: ≈0.75 s



low F_1 value for the approximant [l]

Reading Spectrograms: Approximants

Vowels and *approximants* show vertical striations in wideband spectrograms, and clear dark formant bands.

- Vowels are identified on a spectrogram primarily by the values of their F_1 and F_2 formants.
- Approximants involve constriction comparable to that of a high vowel, and so tend to have fairly low F_1 values. Approximants can often be distinguished by the movement of their F_2 and F_3 formants.

IPA Conventions for American English

Voiceless stops at the beginning of a syllable (and not preceded by [s]) are generally aspirated.

- [kʰəˈtʰæstɹɨfi] (but [skəˈtʰæstɹɨfi])

[u] is generally fronted to [ʉ] in some dialects of American English.

- [ʃʉz] vs. [ʃuz]

https://www.youtube.com/watch?v=3HjIljJd-o0

IPA Conventions for American English

Vowels are often nasalized before a nasal consonant.

- [ˈdæ̃mp]
- A nasalized vowel may even replace a nasal consonant entirely: [ˈdæ̃p]

Secondary articulations contribute an approximant “coloring” to a consonant.

- Consonants that sound like they’re followed by [j] may be transcribed as *palatalized*.
  - [ˈbʲuri] (or [ˈbjuri])
- “Dark l,” at the end of a syllable, is *velarized*.
  - [ˈlʌɫ]

Exploring Spectrograms with Praat

Try Recording Some of the following:

- Oral Stops-
  - [pa] –[apa], [ta]-[ata], [ka]-[aka]
- Nasals-
  - [ma]-[ama], [na],[ana]
- Fricatives-
  - [fa]-[afa], [sa]-[asa], [ʃa]-[aʃa], [ha]-[aha]
- Approximants-
  - [la]-[ala], [ɹa]-[aɹa]

Use the pre-recorded sounds on Canvas to investigate manner in spectrograms

You have the following types of sounds:

Oral stops [pa] –[apa], [ta]-[ata], [ka]-[aka], **Nasals** [ma]-[ama], [na],[ana], **Fricatives**-[fa]-[afa], [sa]-[asa], [ʃa]-[aʃa], [ha]-[aha] & **Approximants** [la]-[ala], [ɹa]-[aɹa]

Look at identifying factors for manner-when comparing, compare like sounds (not differing in additional features) when possible.

Refer to the slide pdf

Reminders

Do HW 4

Read Ladefoged & Johnson chapter 8