Phonetics 3.2: Formants and Spectrograms

April 17, 2020

Fork in the pedagogical road...

Need to know:

. How to identify formants on spectrogram — VIDEO 3

. What a spectrogram represents — SLIDE 9+ THIS VIDEO, VIDEO 3

. Be able to use formants to identify speech sounds in a spectrogram — VIDEO 3

. A formant is a resonant (prominent) component frequency — SLIDES 2–8 IN ONE SENTENCE

Don't need to know

. Where formants come from — THIS VIDEO
. How a spectrogram is computed — OPTIONAL VIDEO?
. Source-filter theory — THIS VIDEO

Recap — speech sounds modeled as complex waves

Suppose you have two **periodic** speech sounds S_1 and S_2. We know we can decompose each into a sum of simple waves of varying amplitudes and frequencies.

Let ω_i be the frequency of the i^{th} component wave (simple) and A_i, B_i be the amplitude of that wave for signals S_1, S_2 respectively. Write a simple wave as $A_i \cdot simple(\omega_i)$ for S_1 and as $B_i \cdot simple(\omega_i)$ for S_2.

Then, $S_1 = \sum_{i=1}^{n} A_i \cdot simple(\omega_i)$ $\qquad S_2 = \sum_{i=1}^{n} B_i \cdot simple(\omega_i)$

Recap — speech sounds modeled as complex waves

A_i = amplitude of the simple wave for S_1's decomposition

B_i = amplitude of the simple wave for S_2's decomposition

frequency	A_i	B_i
ω_1	$A_1 = 10$	$B_1 = 2$
ω_2	3	14
ω_3	8	9
ω_2	0.001	0.02
$\vdots$	$\vdots$	$\vdots$
ω_n	20	16

Typically only few ω_i have amplitude which is large; we characterize S_1, S_2 by the frequencies ω_i which have large amplitude (so large A_i, B_i in this case).

Recap — speech sounds modeled as complex waves

Characterize S_1 by its prominent (large A_i) frequencies ω_1, ω_3 and ω_n

Characterize S_2 by its prominent (large B_i) frequencies ω_2, ω_3 and ω_n

frequency	A_i	B_i
ω_1	$A_1 = 10$	$B_1 = 2$
ω_2	3	14
ω_3	8	9
ω_2	0.001	0.02
⋮	⋮	⋮
ω_n	20	16

We characterize S_1, S_2 with reduced information, but, importantly, we use most important information (analogous to what we did with articulatory phonetics)

Questions from last time

Question: Where do these waves come from in speech?

Where do S_1 (or S_2) come from?

Follow-up question: What determines the relative amplitudes of the components which make up the complex wave?

How do the A_i (or B_i) get determined for frequency ω_i?

A great first answer: Source-filter theory

Where do these waves come from in speech? — Source

Recall how we produce voicing: forcibly passing air through the closed glottis causes the vocal folds to vibrate.

This vibration produces the lowest frequency of the speech signal, which generally has the highest amplitude.

This frequency is called the **fundamental frequency** or $\mathbf{f}_0$.

What determines the relative amplitudes of the components which make up the complex wave? — Filter

The vocal tract is 'filters' the signal from the source (vocal folds) — i.e. when the vocal tract is in different configurations (e.g., for [i] versus [ɛ]), it transforms the source signal in a different way.

The transformed signal can be decomposed into components, and different filters lead to different transformed signals.

Only component waves of certain frequencies have amplitudes that aren't substantially dampened — we call these **resonant frequencies**.

So, the different configurations of the vocal tract will preserve different frequencies (the ones that match its resonant frequencies).

Formants

The fundamental frequency is always prominent (typically called ‘pitch’); the other resonant frequencies are called **formants**.

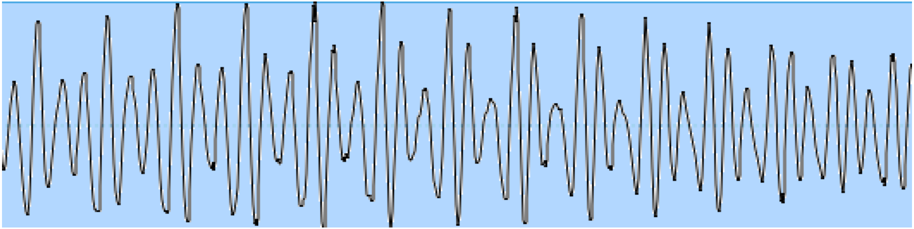
Note: not all sounds have prominent formants — only speech sounds which have a highly periodic waveform. Vowels are highly periodic, approximants are somewhat periodic; other consonants don’t produce waveforms with much periodicity. Next video will elaborate on this.

Formants are numbered in order starting from the first formant or F1, the lowest resonant frequency above fundamental frequency.

Generally, the first three formants are most important in speech signals, but in principle they go upward indefinitely.

Recovering formants (resonant frequencies) from waveform

How can you see a formant on a complex speech waveform?



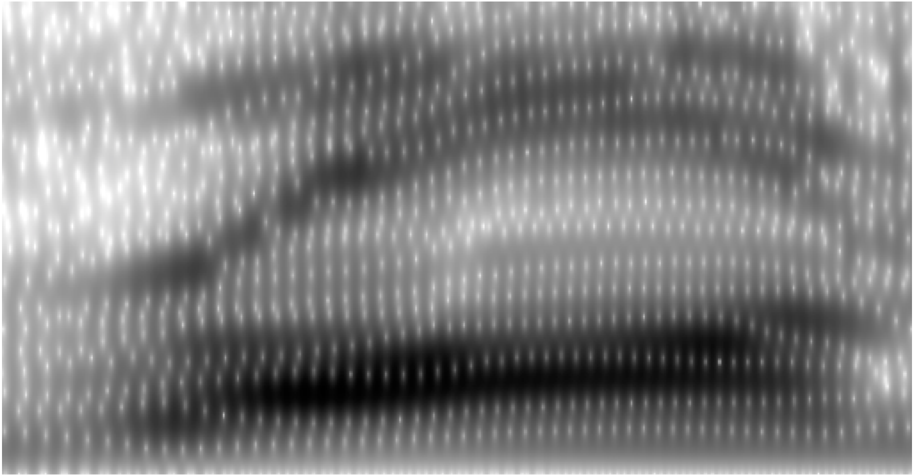
Recovering formants (resonant frequencies) from waveform

Not really.

That's where another tool comes in: the **spectrogram**.

A spectrogram is another way of representing the speech signal. It is a three-dimensional plot.

Example spectrogram



Spectrogram

A spectrogram shows the complex wave decomposed into its various component frequencies. Unlike a waveform, it shows three variables.

Time on the x-axis

Frequency on the y-axis

Amplitude as the darkness at a particular frequency

End of video's lecture material. Look out for optional video?