

Visualising, analysing, and resynthesizing the Human Voice with PRAAT

1. Measuring F0 and Formants manually:

On the spectrogram, identify the fundamental frequency, the harmonics and the first four formants.

Click on the spectrogram. A cursor appears and the corresponding frequency is featured in red on the Y-axis at the left of the spectrogram. Using this cursor, measure the frequency of the fundamental frequency (F0) and of the first four formants (F1, F2, F3 & F4) at the level of the [e] of wh[e]re in your speech utterance and write these values below:

F0 = 253 Hz

F1 = 380 Hz F2 = 1922 Hz F3 = 2697 Hz F4 = 4624 Hz

In the “pitch” menu, select “pitch settings” and set the analysis parameters to *minimum pitch: 60Hz, maximum pitch: 400Hz*.

In the “Pitch” menu, select “show pitch”

You can now see the pitch contour as a blue line superimposed on the spectrogram. The pitch contour is detected by an algorithm called “cross-correlation” (CC), which identifies the periodicity of the vocal signal. Place the cursor on the blue line in the [e] of wh[e]re and read the fundamental frequency value, featured in blue on the Y-axis.

Report this value here:

F0 (Wh[e]re): 240 Hz

and compare it to the one you first measured with the cursor.

2. Measuring F0 and Formants automatically:

Another way to assess the overall pitch of the signal is to detect the pulses (each local amplitude peak, corresponding to an opening of the glottis). To achieve this, select “show pulses” in the “Pulses” menu. Each individual pulse appears as a vertical blue line on the waveform.

Using the cursor select the entire signal. Then go back to the “Pulses” menu and select “voice report”.

The voice report appears in the “Praat: Info” window. Under the “Pitch” heading, you can read the mean / median / standard deviation / minimum pitch.

Report your values in the table below:

Min F0	135.723 Hz
Mean F0	200.901 Hz
Max F0	278.342 Hz
Std dev F0	40.991 Hz

In the “formants” menu, select “formants settings” and set the values as below:

Go back to the formants menu and select “Show Formants”. You can now see the formants as red dots superimposed on the spectrogram. The formants are detected by an algorithm called “Linear Predictive Coding” – LPC, which identifies the peak frequencies of the spectral envelope.

Place the cursor in the [e] of wh[e]re and in the Formants menu select “formant listing”. Compare these values with the ones you originally measured with the cursor. If the values seem unreliable, you can increase the window length (0.1) and adjust the “number of formants” or “maximum formant” parameters.

Formants Listing f1=441 f2=1930 f3=2656 f4=4593

You can also measure formants using Cepstral smoothing as taught during the mammal vocal communication week.

3. Estimating your Vocal Tract Length.

Record your voice saying “HEAD”, “HEED” and “HOOD”.

Measure the values of the first four formants.

Open the excel file (“VTL calculator MoBi HVC”), and type the formant values in the corresponding rows. Select the points on the graph, and shift the blue frame on your data series. The slope of the linear regression on the graph is an estimate of ΔF , the spacing of the formants in your speech utterance, when your vocal tract is modelled as a linear tube. A formula that uses the equation above will automatically estimate the length of your vocal tract.

Compare it with the ones estimated by other students, and look out for gender differences.

F1 F2 F3 F4

	0.5	1.5	2.5	3.5	ΔF	eVTL
					1113.35714285	
Vowel hEAd	559	2039	2697	3800	714	15.72
					987.261904761	
Vowel hEEd	394	519	2580	3802	905	17.73
					1169.14285714	
Vowel hOOd	582	1765	2862	4131	286	14.97
					1344.30952380	
hEAd small	920	2100	3537	4508	952	13.02
					984.357142857	
hEAd big	640	1646	2323	3450	143	17.78

Repeat this for vowels in [heed] and [hood] – what do you observe and why?

Now record yourself saying “head” trying to sound as big as possible, and also trying to sound as small as possible - while sounding like a realistic same-sex adult (you must try to deceive listeners – so do not imitate a child, a mouse or an elephant ☺!). Measure F0 and ΔF (and eVTL using the excel file). Compare these values with your baseline values.

Baseline f0 = 253, Δf =1113 eVTL=15.72

head fo=253 df=1113

big fo=177 df=984

small fo=355 df=1344

4. Human voice resynthesis (PSOLA)

In the “Praat objects” window, select “Record mono sound” from the “New” menu. When you are ready, click on “record”, and say, speaking at a slow pace, “This is what I sound like when I modify the spectral components of my voice” (or anything else if you are more inspired) into the microphone. Click on “Stop”, give your name to the recording and “Save to list”. Your recording is automatically added to the “objects” list in the PRAAT objects window.

Select your recording, click on the “convert” button on the bottom right, and select the “change gender” option.

A new window appears in which you can specify the new parameters of your voice.

If you are female:

- type 0.8 in the “formant shift ratio” field – this will shift your formant frequencies down by 20%, giving you an apparent adult male vocal tract.
- type 110 in the “new pitch median” field – this will give your voice the pitch (glottal pulse rate) of an adult male.
- Click on OK, a new sound appears in the Objects list. Click on play.

If you are male:

- type 1.2 in the “formant shift ratio” field – this will shift your formant frequencies up by 20%, giving you an apparent adult female vocal tract.
- type 210 in the “new pitch median” field – this will give your voice the pitch (glottal pulse rate) of an adult female.
- Click on OK, a new sound appears in the Objects list. Click on play.

Redo this manipulation (selecting the original recording):
first changing the formants only,
then changing the pitch only.
How does this sound?

Selecting the original recording, perform a series of resyntheses combining gradual increases of the formant spacing (multiply the formant spacing by 1.2, 1.3, 1.4, and 1.5), and increases of the pitch (to 300, 400 and 500 Hz). What are the results?

Module Assessment:

Please enter your data from exercise 3 anonymously (for ID, just leave the current number) in the following google doc before 4 pm today:

<https://docs.google.com/spreadsheets/d/1rL1lbt5jH5DXK0XsgbruISGJ6LQ5NnLN/edit?gid=361208942#gid=361208942>

Do not use decimals ☺.

In no more than 4 pages (font Arial size 11) including figures, tables and references (citing 5 papers max), please write a short report (intro, methods, results and discussion, no abstract) investigating the relationship between sex, age, height (cm) and frequency variables in honest and size-deceptive speech.

Please submit this report by Friday 29th November @ 4pm.