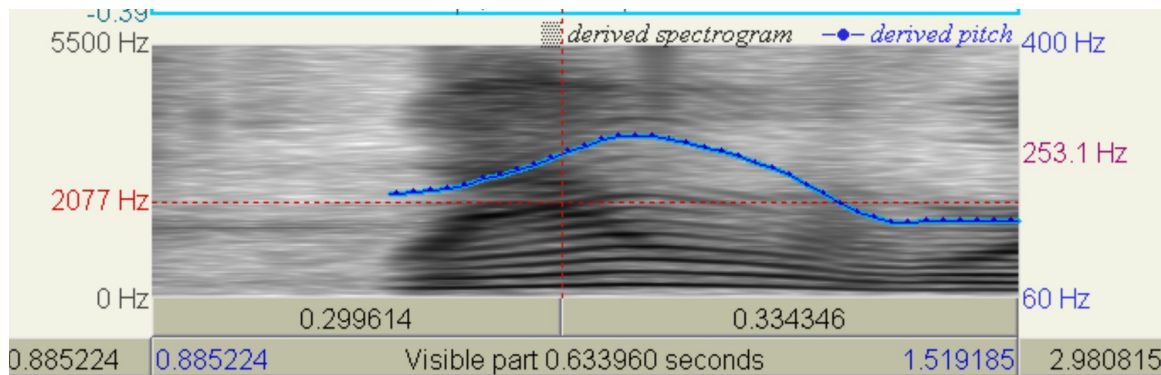


## Day 1 - Visualizing, analyzing, and resynthesizing the Human Voice with PRAAT

### 1. Measuring F0 and Formants manually:

Spectrogram of WHERE WERE YOU with pitch curve:



Fundamental and First Four Harmonics, Manual Measurements:

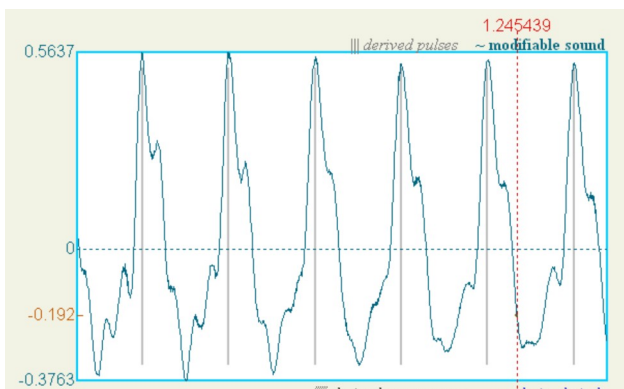
F0 = 258 Hz  
F1 = 380 Hz  
F2 = 1922 Hz  
F3 = 2697 Hz  
F4 = 4624 Hz

Pitch contour value of F0 wh[e]re: 253 Hz

My measured value of f0 was very close to the pitch contour value, with a difference of only 5Hz.

### 2. Measuring F0 and Formants automatically:

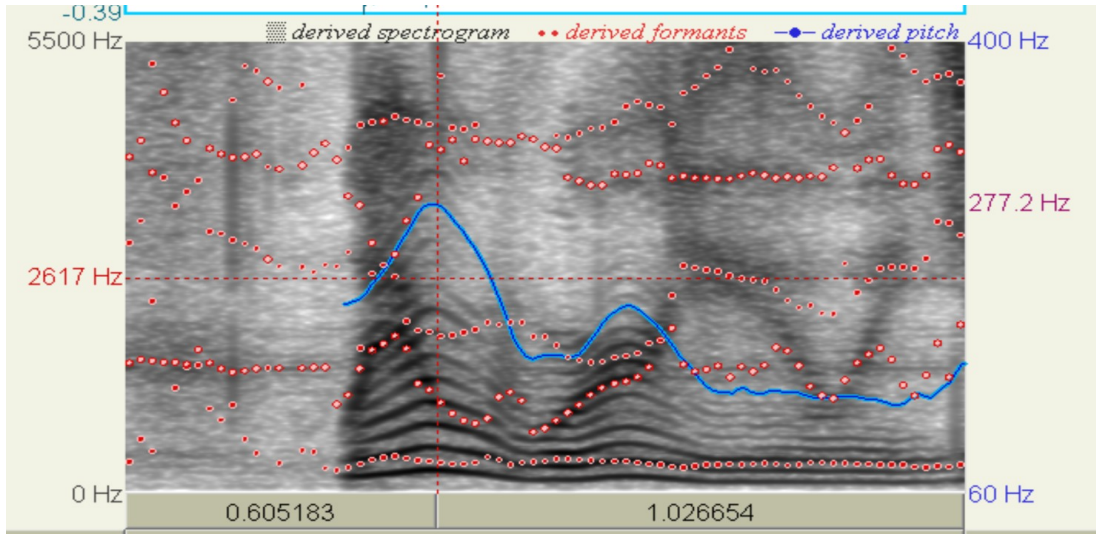
Pulse Spacing on Waveform:



Pulse-derived Voice Report Values:

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

Spectrogram with formants:

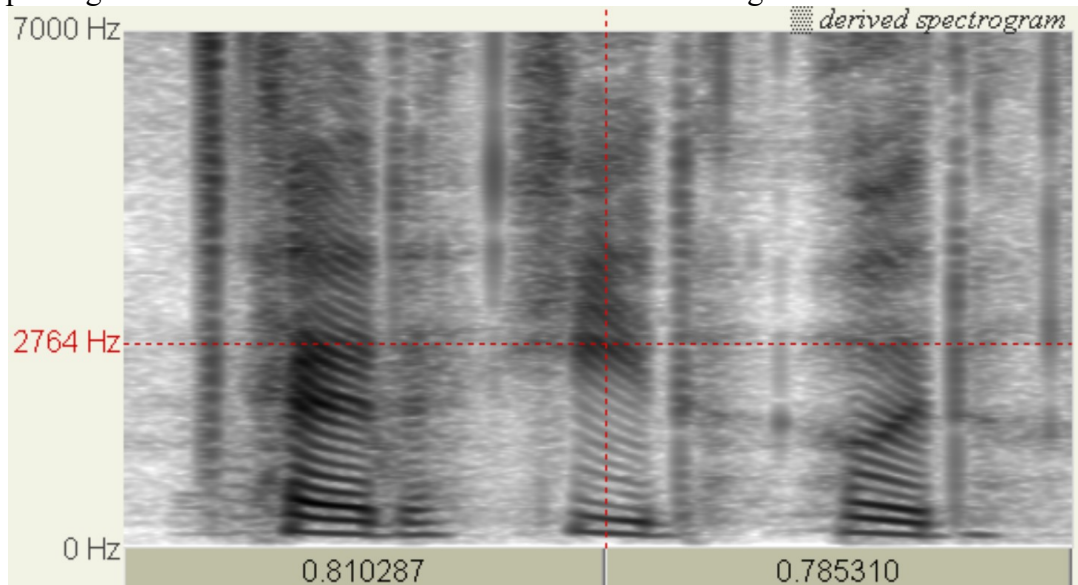


Formant Report-derived Formants for whEre:

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

### 3. Estimating your Vocal Tract Length: HEAD HEED HOOD

Spectrogram of HEAD HEED HOOD with formant tracking:



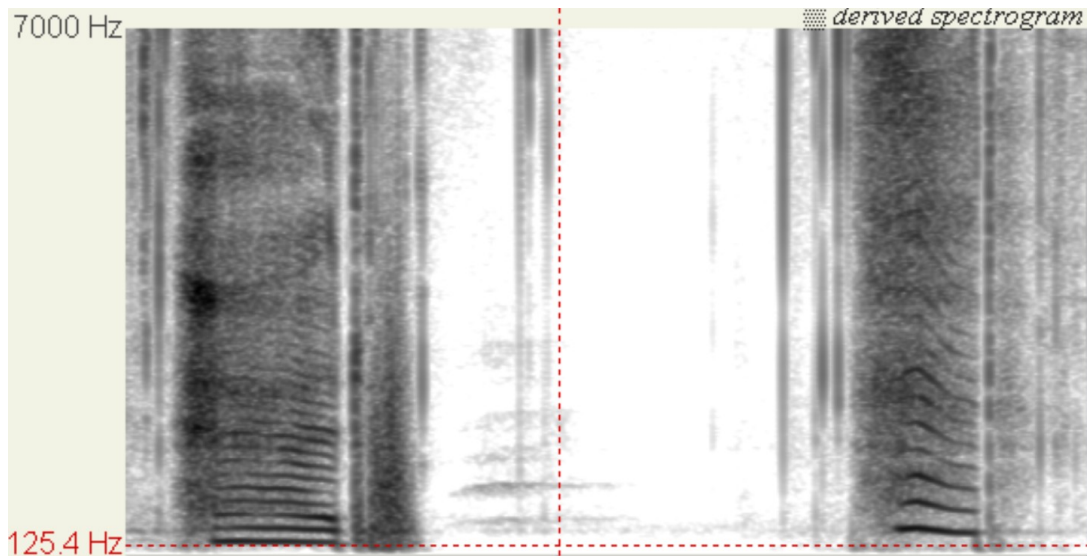
### First Four Formants and VTL Calculator:

Audio	F1	F2	F3	F4	Calculated $\Delta F$	Calculated eVTL
Vowel hEAd	559	2039	2697	3800	1113.35714285714	15.72
Vowel hEEd	394	519	2580	3802	987.261904761905	17.73
Vowel hOOd	582	1765	2862	4131	1169.14285714286	14.97
hEAd small	920	2100	3537	4508	1344.30952380952	13.02
hEAd big	640	1646	2323	3450	984.357142857143	17.78

### Comparison of vowel production HEAD HEED HOOD:

The mouth is open the widest during the production of the vowel for head, it narrows and widens while producing hood, and it slightly narrows as the lower lip drops into an 'o' shape during hood. This is reflected in the formants: f1 starts around 550, drop to 400, and again to near 600. F2 starts around 2000, drops to 500, and raises to 1800. F3 starts around 2700, drops to 2600, and raises to 2900. F4 starts around 3800, slightly raises to 3802, then moves to 4100. The second formant fluctuates the most, and this is seen from the curvy dark grey line on the spectrogram above. The higher order formants don't vary as wildly, but do increase a couple hundred Hz with a slight positive slope over the 3 sounds. It can be inferred that vowel production and formant spacing, therefore, is closely linked with the internal structure of the oral cavity.

Spectrogram of HEAD with big and small vocalizations:



(from above) First Four Formants and VTL Calculator:

Comparative f0 and spacing for HEAD across normal, big, and small vocalizations:

head	fo=253	df=1113
big	fo=177	df=984
small	fo=355	df=1344

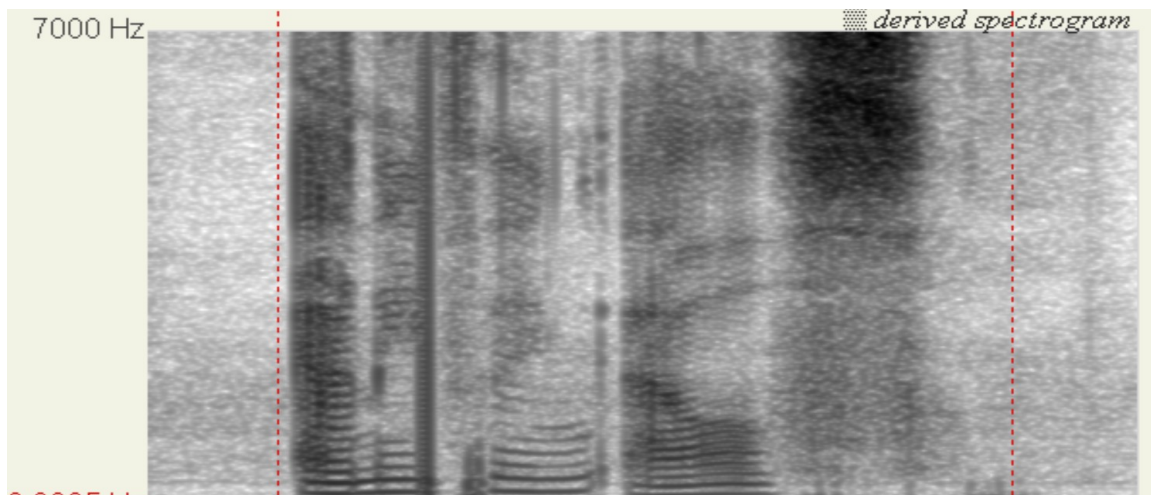
We can see that, compared with the normal vocalizations of HEAD, the big vocalization fundamental frequency decreases from 253 to 177 Hz ( $df=-76$ ), and the formant spacing decreases from 1113 to 984 ( $df=-129$ ). We also see that, compared with normal vocalizations, the small vocalization fundamental frequency increases from 253 to 355 Hz ( $df=102$ ) and the formant spacing increases from 1133 to 1344 Hz ( $df=211$ ). This shows that the  $f_0$  N2B difference is about 70% smaller than the N2S difference, and the formant spacing for N2B is a little more than half that of the N2S vocalizations. Thus, the impact of N2S vocalizations varies more significantly than N2B vocalizations.

N2B (normal-to-big):	$df_0=-76$	$df_n=-129$
N2S (normal-to-small):	$df_0=102$	$df_n=211$

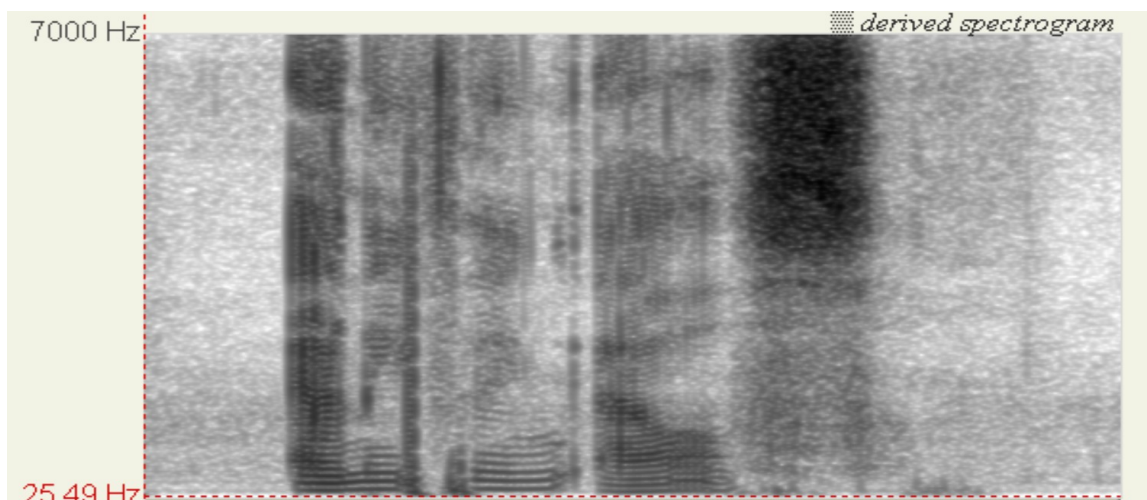
For the dataset: baseline  $f_0 = 253$ ,  $\Delta f = 1113$  eVTL=15.72

#### 4. Human voice resynthesis (PSOLA)

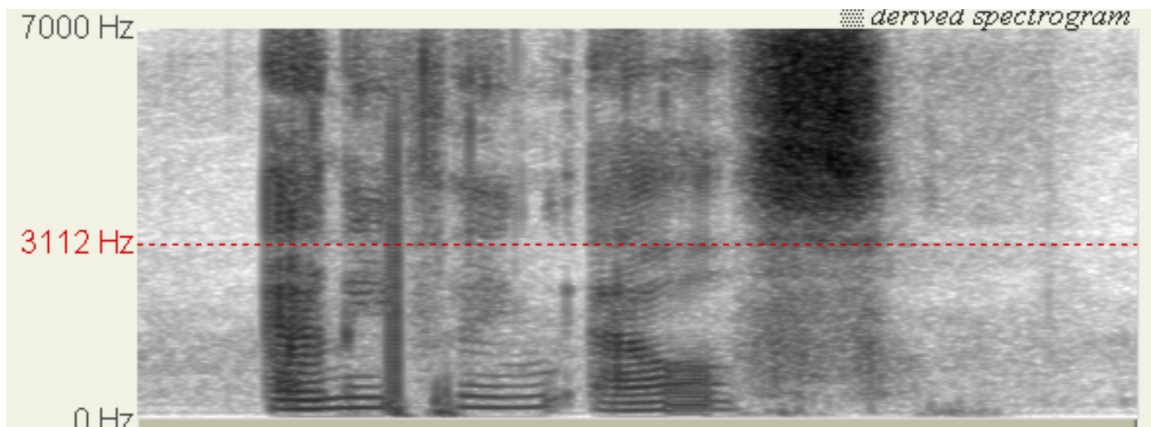
Spectrogram of ANYTHING ELSE:



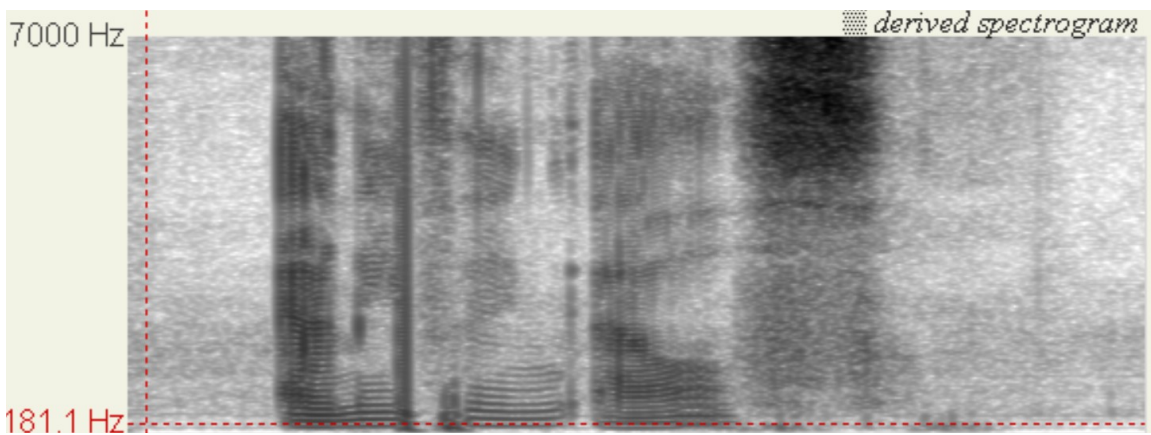
Spectrogram of ANYTHING ELSE, formants=0.8, pitchcenter=110; sounds like a speaker of the binary opposite sex



Spectrogram of ANYTHING ELSE, only formants=0.8; Sounds muffled and like the mouth is closed off, sounds like the speaker is not fully dropping the jaw



Spectrogram of ANYTHING ELSE, pitchcenter=110; voice sounds gravelly or raspy, sounds like the palate is widely stretched



Comparative Analysis:

It's interesting to perceive these same changes over both a gradual as well as wider spectrum. A variety of the sounds remind me of vocal manipulations heard in cartoons (think: Chipmunks) or other older forms of media. I'm also reminded of when transgendered individuals begin to take hormone therapy, and as their bodies adjust to the hormones, their voices also experience a phase of adjustment as they learn to speak in a way that corresponds with their gender and not with their sex assigned at birth.