SHS homework solutions 7: formants

Solution 7.3.

a. In formant measurement, we use pre-emphasis to make the spectrum flatter. Without pre-emphasis, the spectrum of a vowel typically drops off by 6 dB per octave, and pre-emphasis takes this down to approximately 0 dB per octave, so that all formants look equally strong in the spectrum (on average; remember that they are spaced 1000 or 1100 Hz apart for schwa). The Linear Predictive Coding formant analysis algorithm will find the most important peaks in the spectrum, and if the formants are all equally strong in the spectrum, the algorithm is likely to find all of them. Without pre-emphasis, the algorithm tends to find too many low-frequency peaks that we don't like to associate with formants.

Note that we had pre-emphasis in spectrogram drawing as well. For vowels, switching on pre-emphasis makes sure that all formants look equally strong in the spectrogram (on average), as they do to the human ear (as a result of the integration that takes place on the basilar membrane because of its logarithmic frequency representation). Without pre-emphasis in the spectrogram, lower formants would look stronger than higher formants.

b. The **Maximum formant** setting refers to vocal tract length. If the vocal tract length is 17 centimetres and can be regarded as a straight tube that is closed at the glottis and open at the lips, than the wavelengths of standing waves in this tube are 4 tube lengths, 4/3 tube lengths, 4/5 tubelengths, 4/7 tube lengths, 4/9 tube lengths, and so on (as shown in class). These wavelengths are 68, 68/3, 68/5, 68/7, and 68/9 centimetres (and so on). With a velocity of sound of 340 metres per second, the frequencies are 500, 1500, 2500, 3500, and 4500 Hz (and so on), so these are the resonance frequencies, i.e. the formants, of a straight-tube vowel, which is approximately schwa ([ə], the sound in Dutch *de* or *ehhh..*). This means that for a 17-cm long vocal tract (typical of males, whose larynx drops by 1.5 cm during puberty), we expect that there will be five formants below 5000 Hz. For a vocal tract 15.5 cm long (typical of females), the wavelengths of the resonances will be 10 percent shorter, and their frequencies will be 10 percent higher, so that we expect to find five formants below 5500 Hz.

Solution 7.4.

The following is your teacher's attempt:

```
sound = Read from file: "vowels.wav"
iamMale = 1   ; true (otherwise 0)
maximumFormant = if iamMale then 5000 else 5500 fi   ; Hz
timeStep = 0.001   ; seconds: use many measurements
formant = To Formant (burg): timeStep, 5, maximumFormant, 0.025, 50.0
textgrid = Read from file: "vowels.TextGrid"
numberOfIntervals = Get number of intervals: 1
writeInfoLine: "vowel", tab$, "F1", tab$, "F2"
for i to numberOfIntervals
     # When we arrive here, either the TextGrid or the Formant stands selected!
     # Therefore, make sure that the TextGrid gets selected always:
     selectObject: textgrid
     vowel$ = Get label of interval: 1, i
     if vowel$ <> ""
          t1 = Get start time of interval: 1, i
          t2 = Get end time of interval: 1, i
          selectObject: formant
          # The median is the 50 percent quantile:
          f1 = Get quantile: 1, t1, t2, "hertz", 50%
          f2 = Get quantile: 2, t1, t2, "hertz", 50%
          appendInfoLine: vowel$, tab$, round (f1), tab$, round (f2)
     endif
endfor
```

(The median is more robust against outliers than the mean.) My output is:

```
vowel     F1   F2
u    267  1593
oː   376  990
ɔ    530  674
ɑ    558  1023
aː   364  1334
ɛ    415  1616
eː   328  2055
i    238  2188
ɪ    351  1867
ʏ    366  1505
y    245  1825
øː   354  1516
```

Notes:

• I had expected a much lower F2 for /u/. Perhaps the maximum formant should have bene lower than 5000 Hz, especially for this vowel.

• In class we showed that the second visible resonance in the waveform of /i/ lies above 3000 Hz. The F2 found here is a weak formant that is usually not visible in the waveform, but it is visible in the spectrogram. The one that is visible in the waveform comes from a strong cooperation between F3 and F4, which together cause a high peak in the spectrum as a result of being close together (e.g. 3000 and 3700 Hz).

• The measurements of /aː/ (very low F1) and /ɔ/ (high F1) are probably wrong.

Solution 7.5.

If you don't like to type the values of F1 and F2 in your drawing script, you could insert drawing commands directly into the script of **7.4**:

```
#
# Praat script analyseAndDrawVowels.praat
# Paul Boersma, 2019-09-25
#
sound = Read from file: "vowels.wav"
iamMale = 1   ; true (otherwise 0)
maximumFormant = if iamMale then 5000 else 5500 fi   ; Hz
timeStep = 0.001   ; seconds: use many measurements
formant = To Formant (burg): timeStep, 5, maximumFormant, 0.025, 50.0
textgrid = Read from file: "vowels.TextGrid"
numberOfIntervals = Get number of intervals: 1
Erase all
Times
Font size: 12
Select outer viewport: 0, 6, 0, 5.0
Axes: 3000, 500, 1000, 200
Draw inner box
Marks left every: 1.0, 200, "yes", "yes", "yes"
Marks bottom every: 1.0, 500, "yes", "yes", "yes"
for i to numberOfIntervals
     # When we arrive here, either the TextGrid or the Formant stands selected!
     # Therefore, make sure that the TextGrid gets selected always:
     selectObject: textgrid
     vowel$ = Get label of interval: 1, i
     if vowel$ <> ""
          t1 = Get start time of interval: 1, i
          t2 = Get end time of interval: 1, i
          selectObject: formant
          f1 = Get quantile: 1, t1, t2, "hertz", 50%
          f2 = Get quantile: 2, t1, t2, "hertz", 50%
          Text special: f2, "centre", f1, "half", "Times", 18, "0", vowel$
     endif
endfor
```

Now I take an alternative approach, namely with two scripts: one that analyses the sound and TextGrid and saves the formant values in a table file, and one that reads the table file and draws its contents. This method will be useful is you have much more data than just this mini-set:

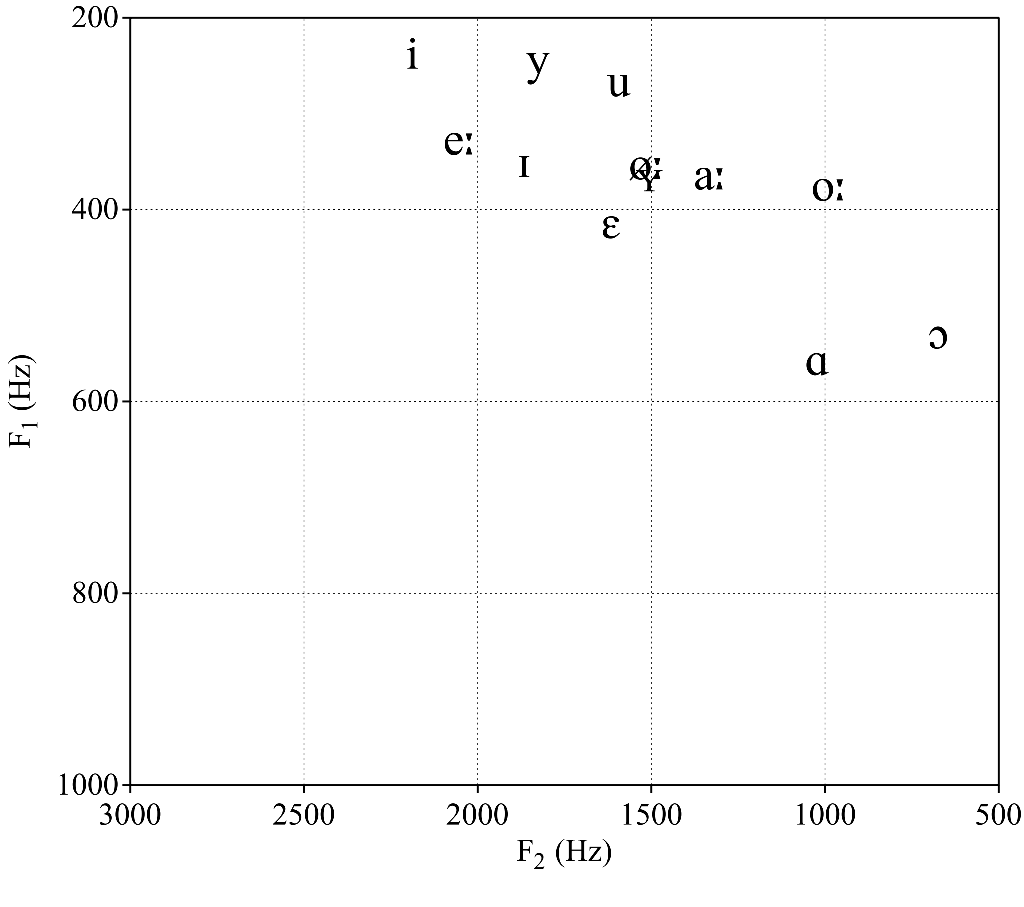
```
#
# Praat script analyseVowels.praat
# Paul Boersma, 2019-09-25
#
sound = Read from file: "vowels.wav"
iamMale = 1   ; true (otherwise 0)
maximumFormant = if iamMale then 5000 else 5500 fi   ; Hz
timeStep = 0.001   ; seconds: use many measurements
formant = To Formant (burg): timeStep, 5, maximumFormant, 0.025, 50.0
textgrid = Read from file: "vowels.TextGrid"
numberOfIntervals = Get number of intervals: 1
table = Create Table with column names: "table", 0, "vowel F1 F2"
numberOfVowelsFound = 0
for i to numberOfIntervals
     selectObject: textgrid
     vowel$ = Get label of interval: 1, i
     if vowel$ <> ""
          t1 = Get start time of interval: 1, i
          t2 = Get end time of interval: 1, i
          selectObject: formant
          # The median is the 50 percent quantile:
          f1 = Get quantile: 1, t1, t2, "hertz", 50%
          f2 = Get quantile: 2, t1, t2, "hertz", 50%
          selectObject: table
          Append row
          numberOfVowelsFound += 1
          Set string value: numberOfVowelsFound, "vowel", vowel$
          Set string value: numberOfVowelsFound, "F1", fixed$ (f1, 3)
          Set string value: numberOfVowelsFound, "F2", fixed$ (f2, 3)
     endif
endfor
selectObject: table
Save as tab-separated file: "vowels.Table"
```

You can inspect this table in Praat by selecting it and doing **View & Edit**. The Table file can be read not only by Praat, but also by Excel and R (for external software, perhaps make sure that the table file is saved as UTF-8 and not UTF-16, via **Praat** > **Preferences** > **Text writing preferences**).

Here is the drawing script:

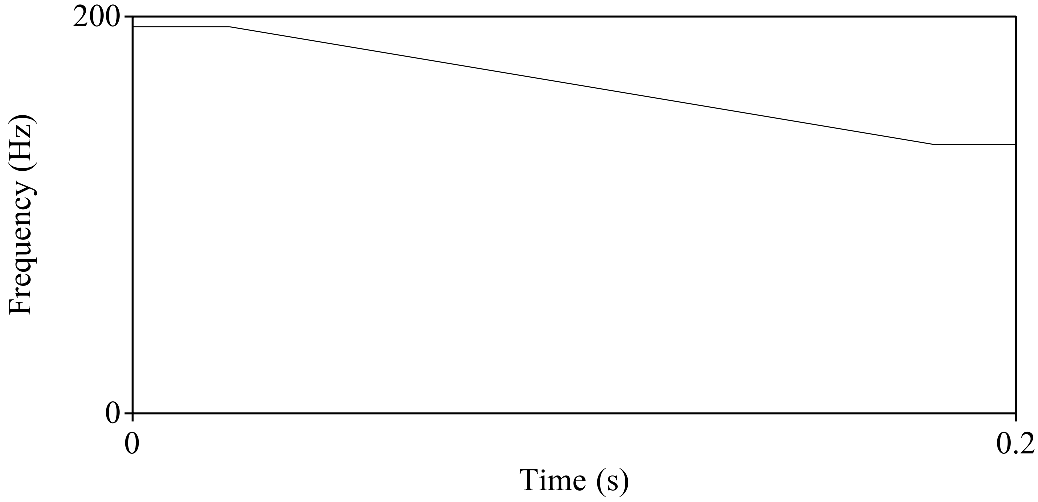
```
#
# Praat script drawVowels.praat
# Paul Boersma, 2019-09-25
#
Read Table from tab-separated file: "vowels.Table"
Erase all
Times
Font size: 12
Select outer viewport: 0, 6, 0, 5
Axes: 3000, 500, 1000, 200
Draw inner box
Marks left every: 1.0, 200, "yes", "yes", "yes"
Marks bottom every: 1.0, 500, "yes", "yes", "yes"
Text bottom: "yes", "F_2 (Hz)"
Text left: "yes", "F_1 (Hz)"
numberOfVowels = Get number of rows
for ivowel to numberOfVowels
     vowel$ = Get value: ivowel, "vowel"
     f1 = Get value: ivowel, "F1"
     f2 = Get value: ivowel, "F2"
     Text special: f2, "centre", f1, "half", "Times", 18, "0", vowel$
endfor
```

The resulting picture shows problems with /u/, /a:/ and /ɔ/:

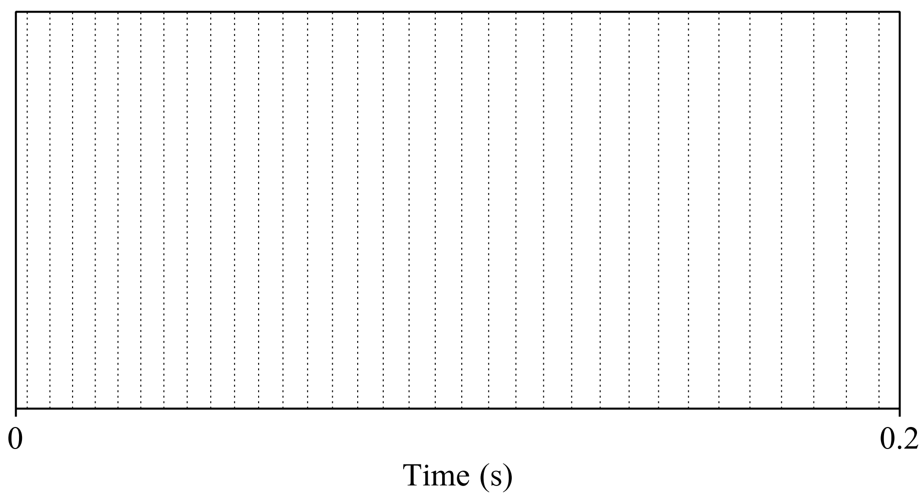


Solution 7.6.

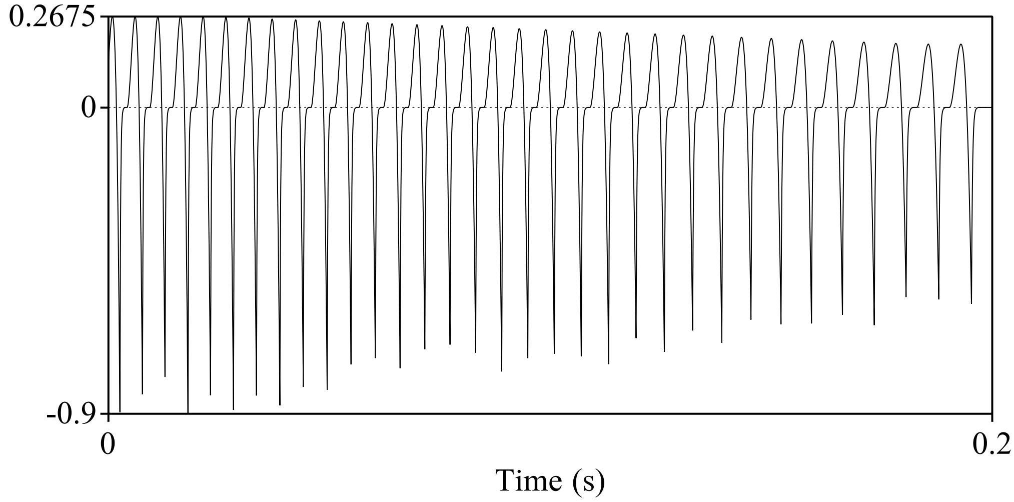
This was my PitchTier:



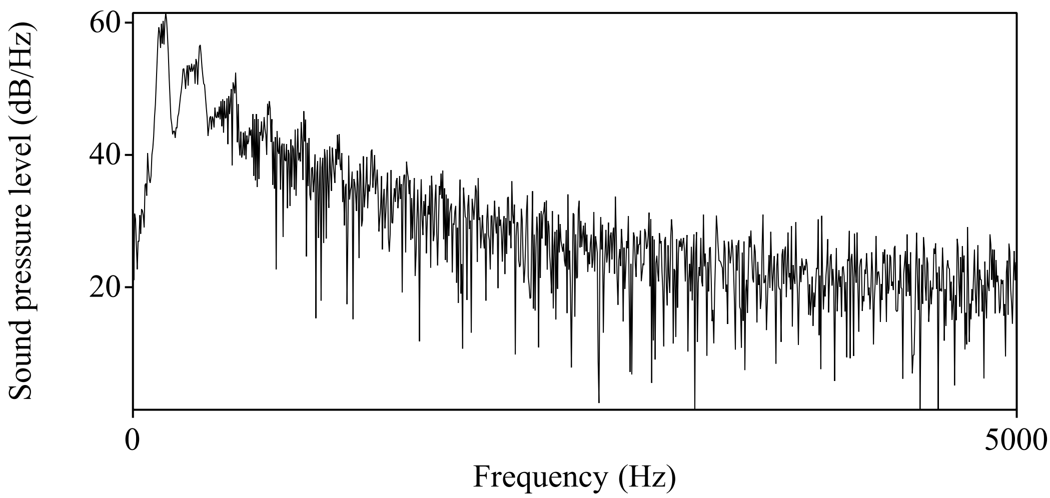
That’s a falling pitch, so the periods should grow, as is confirmed by the PointProcess:



The resulting phonation:

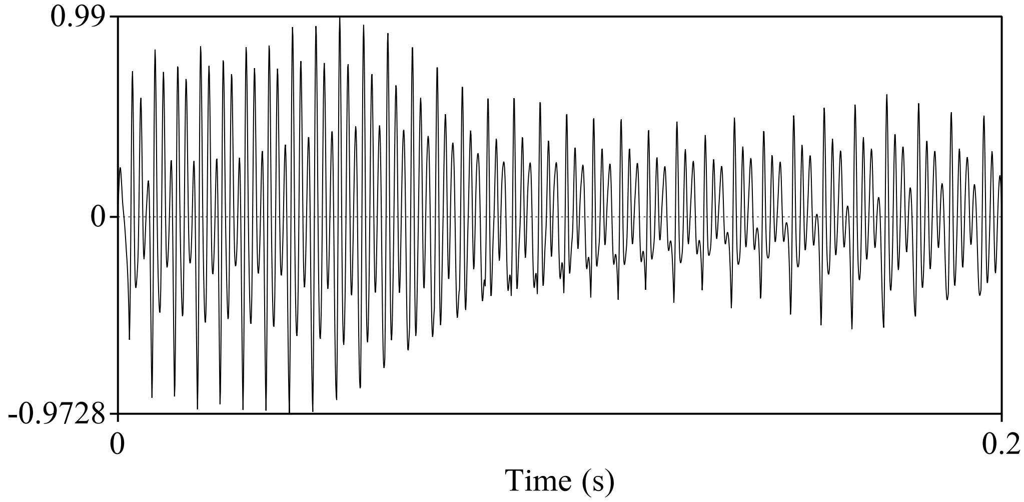


That’s the time-derivative of the airflow through the glottis. The spectrum of the airflow drops off by approximately 12 dB/octave, and taking the derivative reduces this to a drop-off of 6 dB/octave (acoustically, taking the derivative corresponds to taking the radiation at the lips into account). Here is the spectrum of the source sound above:

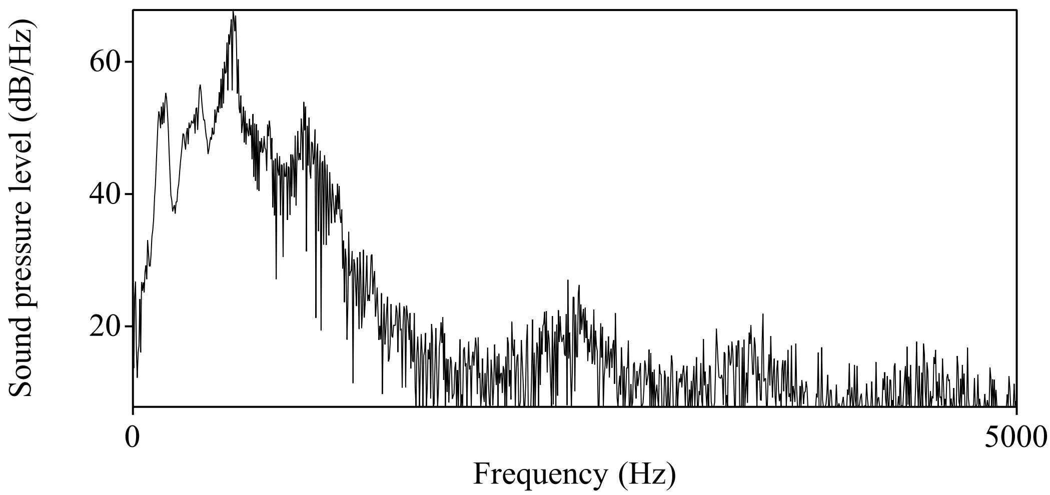


This seems to drop off by a bit more than 6 dB/octave, namely approximately 8 or 9 dB/octave.

Filtering with F1 = 558 Hz and F2 = 1023 Hz, with 55.8-Hz and 102.3-Hz bandwidths, plus the three higher formants mentioned, yields the following sound:

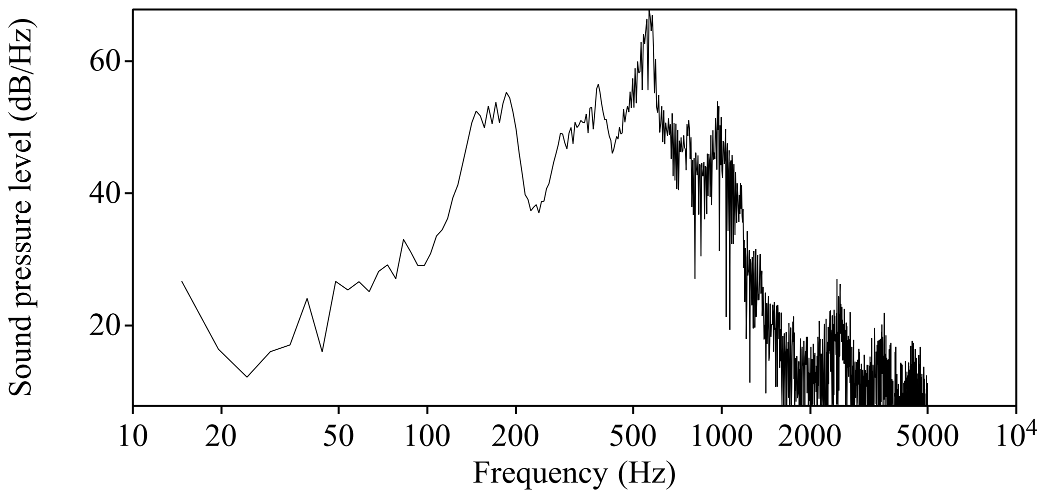


Its spectrum is



This drops of much faster than 6 dB/octave in the low frequencies, and this is because F1 and F2 are close together (they collaborate to yield an especially strong peak). but between 2500 and 5000 Hz the drop is indeed 6 dB/octave.

The assignment talked about drawing the spectrum along a logarithmic frequency scale:



A constant 6 dB/octave drop-off would be a straight line in this figure, because both the horizontal axis and the vertical axis are logarithmic (and both dB and octave are logarithmic measures).