Ling 110 - Introduction to Phonetics and Phonology

Acoustic speech analysis in WaveSurfer

This is a set of exercises for us to do together in a course section meeting. The aim of these exercises is to acquaint you with how to make spectrograms and spectra and some of the analysis parameters that can be adjusted to improve your acoustic speech analysis.

Exercise 1. Complex periodic waves in waveform and spectrogram displays.

a. Open both “sine1k.wav” and “sine2k.wav” in WaveSurfer

> WaveSurfer - “File” “open”
> choose configuration - “speech analysis”

b. Mix them together

> in sine1k - select all, copy
> in sine2k - select all, mix paste (try different mixture amplitudes,
> 50/50 is a good starting point)

c. Look at the spectrogram and wave form. Measure the frequencies of the two components by positioning the cursor on the dark bands in the spectrogram. Narrow the analysis bandwidth using the “spectrogram properties” dialog box (right click, or <control> click on the spectrogram to get the “properties” dialog). The screen capture below (figure 1) shows the dialog box with the number of points set to 256 and the top frequency of the spectrogram set to 3000 Hz.

Changing the number of points in the analysis window changes the analysis window bandwidth and thus the frequency resolution of the spectrogram. Figure 2 shows the result of changing the analysis bandwidth.

Q1: Can you figure out what trick was necessary to get the number of points up to 1024?

Q2: How is the spectrogram different for Hamming versus Hanning windowing?

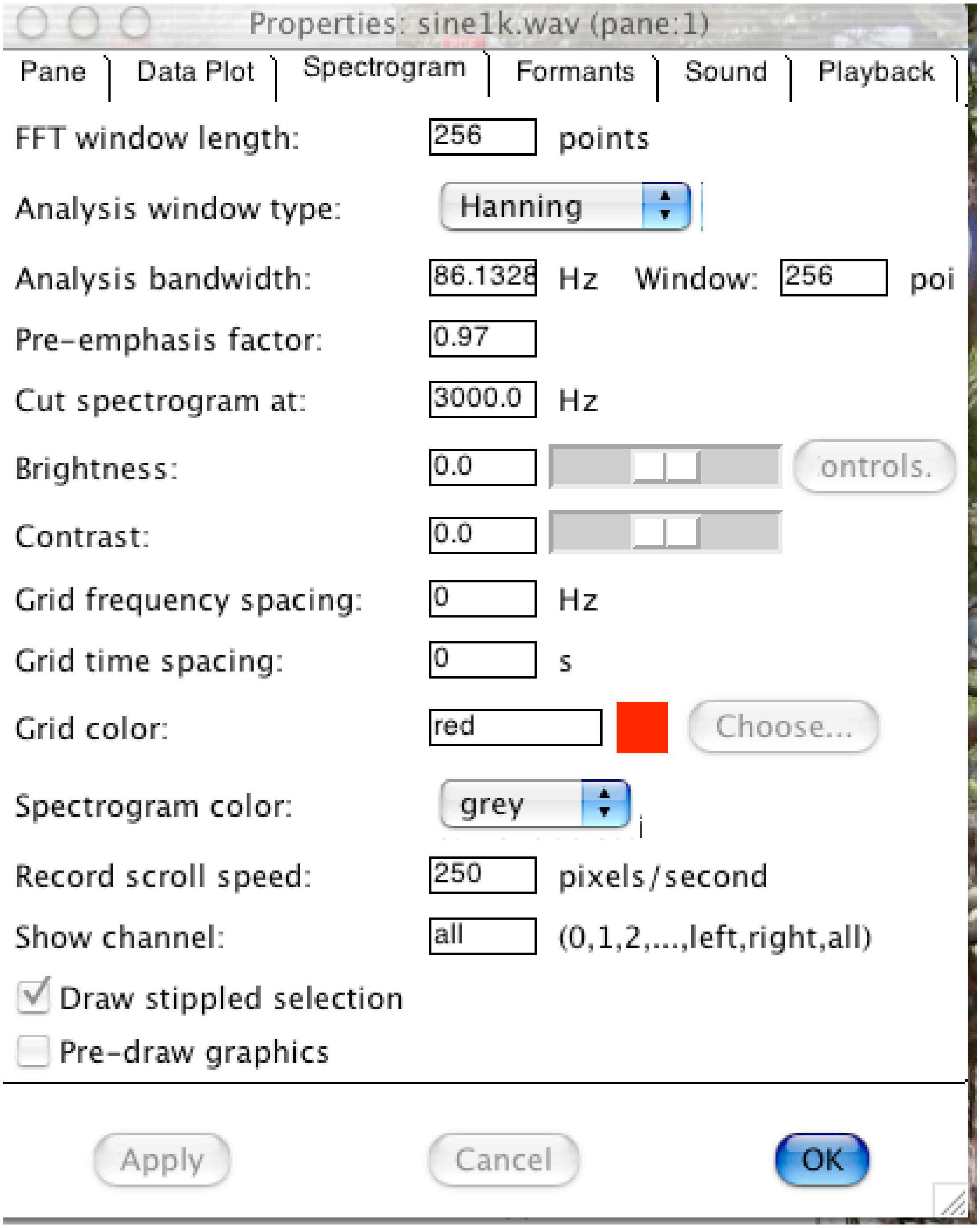

Figure 1. The "spectrogram properties" dialog box in WaveSurfer.

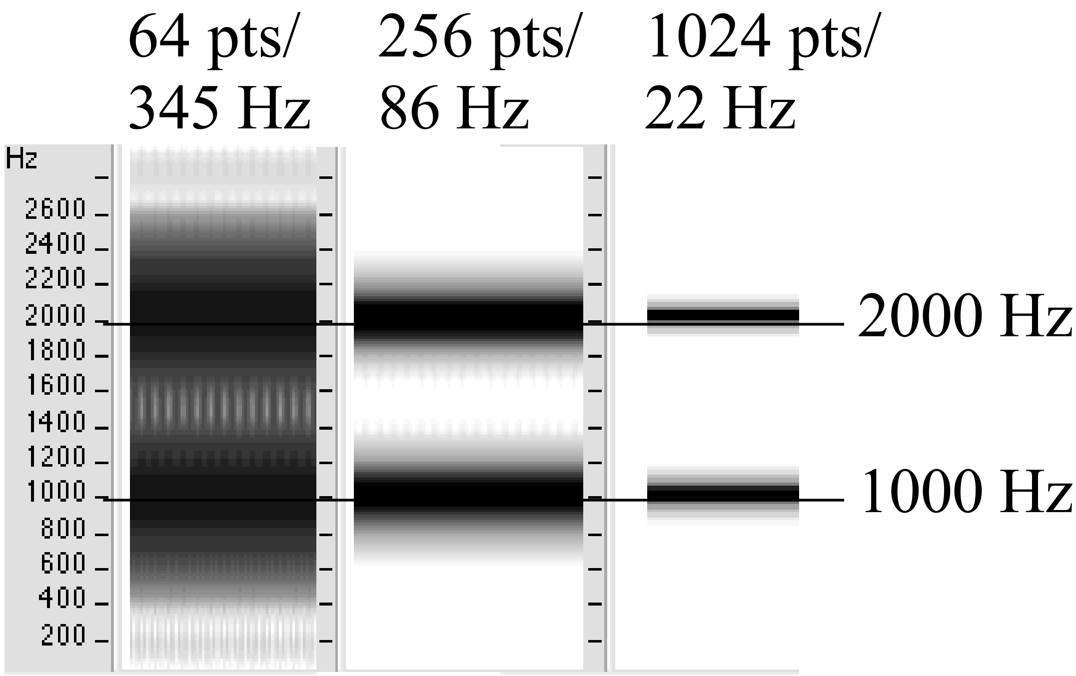

Figure 2. Spectrograms of a two component complex periodic wave with different analysis window bandwidth settings.

d. Open a "spectrum section" window. Set the FFT points to 1024. Now click in the spectrogram and move the mouse around. There are vertical and horizontal line cursors in the spectrum section, and notice that these cursors move in the spectrum section as you move the mouse over the spectrogram.

> Q3: Describe the relationship between the mouse position on the spectrogram and the line cursors in the spectrum section.
>
> Q4: The sound that we are looking at only has true components at 1000 and 2000 Hz, all other peaks in the spectrum are artifacts of the analysis algorithm. Keeping this in mind, do you prefer a Hanning window or a Hamming window?

Exercise 2. Speech spectrograms and pitch contours.

a. Open two copies of "AH_ganeli_c.wav". Change the sampling rate of one of them to 8,000 samples per second ("transform" "convert"). Change parameters in the spectrogram properties (both under the "data plot" and "spectrogram" tabs) until you get spectrograms that look like the ones in figure 3 for your two sound files.

Q5: If you set the Analysis bandwidth to be 300 Hz WaveSurfer will automatically adjust the number of points in the analysis window so that the bandwidth is as close as possible to 300. What is the number of points for an analysis bandwidth of about 300 Hz when the sampling rate is 8000 Hz? What about for 22050 Hz? How does the number of points in the analysis window relate to the sampling rate?

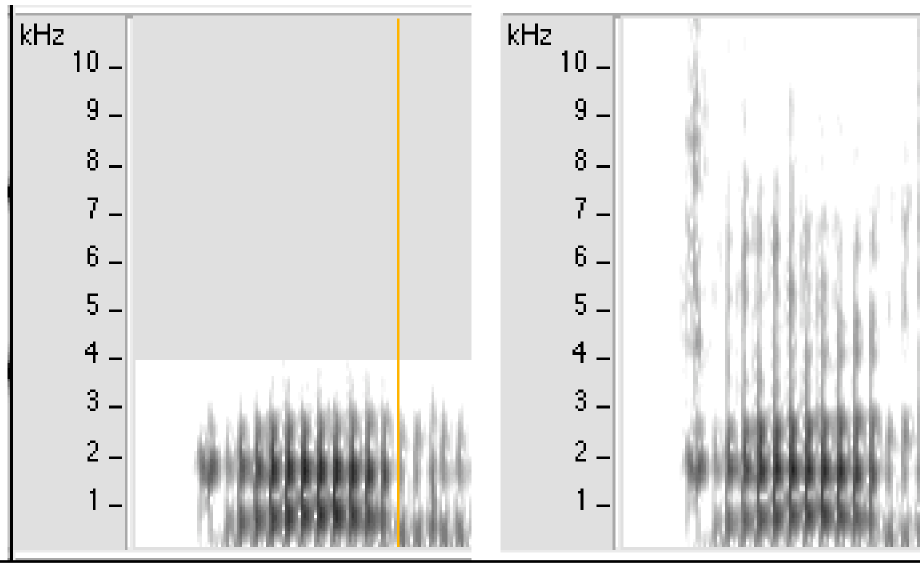

Figure 3. Spectrogram of the first syllable [gan] in “ganeli” with sampling rate at 8000 Hz and 22050 Hz.

b. Add a pitch contour to the analysis (right click or <control> click on wave - “create pane” - “pitch contour”). Figure 4 shows a pitch contour with an analysis window size of 0.009 seconds (nine milliseconds) and a new pitch estimate taken every millisecond (frame interval 0.001 seconds). The data vertical limits of the display were also adjusted to fit this particular pitch contour closely. The spectrogram in figure 4 was computed with an analysis window that is 50 Hz wide, and the spectrogram display was cut so that only 2000 Hz is displayed. The narrow bandwidth of the spectrogram analysis window results in a display of the individual vowel harmonics and these reflect the pitch of the voice - the same information shown in the pitch curve.

Q6: Notice that there is a “glitch” in the pitch contour at the location of the vertical cursor in figure 4 at the boundary between the [n] and the [e] of “ganeli” and another glitch at the end of the [e] vowel. Can you make these discontinuities in the pitch trace go away by changing the analysis parameters of the pitch tracking algorithm (in pitch contour “properties”)? Does the narrowband spectrogram add any insight into whether it is desirable to find settings that eliminate the pitch contour glitches?

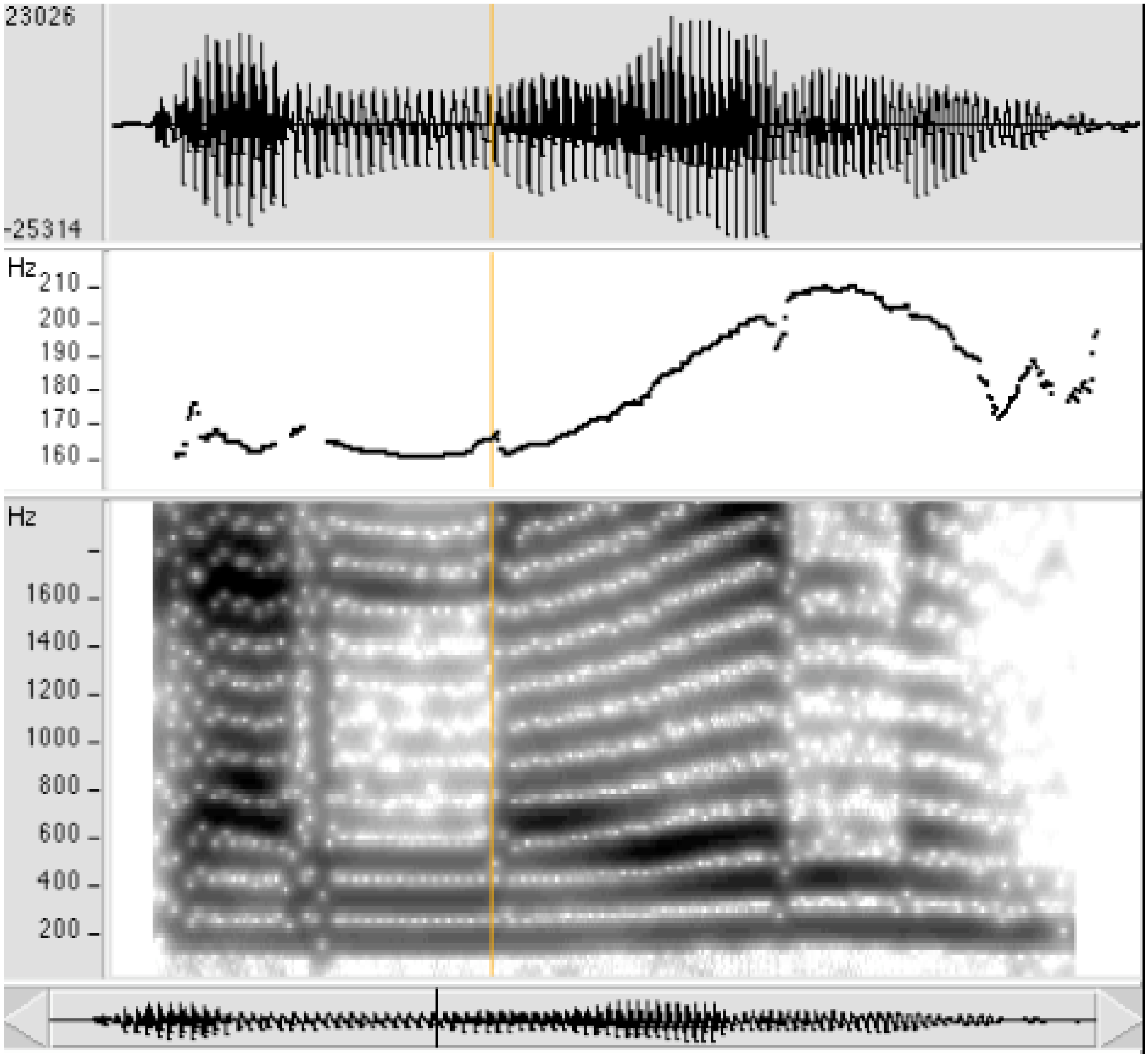

Figure 4. From the top of the figure: an acoustic waveform, pitch contour, and narrowband spectrogram of the word "ganeli".