Voicing and phonation

In this exercise you will examine electroglottographic (EGG), acoustic, and oral airflow signals to discover some key points in the acoustic theory of speech production and in the physical process of voicing in speech. In short the exercise questions what we mean when we say that a segment is voiced, but we will also examine some conditional effects on phonation type.

A speaker said the odd sentence "He who has the only applicable answer will keep it to himself" and we recorded a microphone signal (you'll calculate how far the microphone was from the mouth), an EGG signal, and an oral airflow signal.

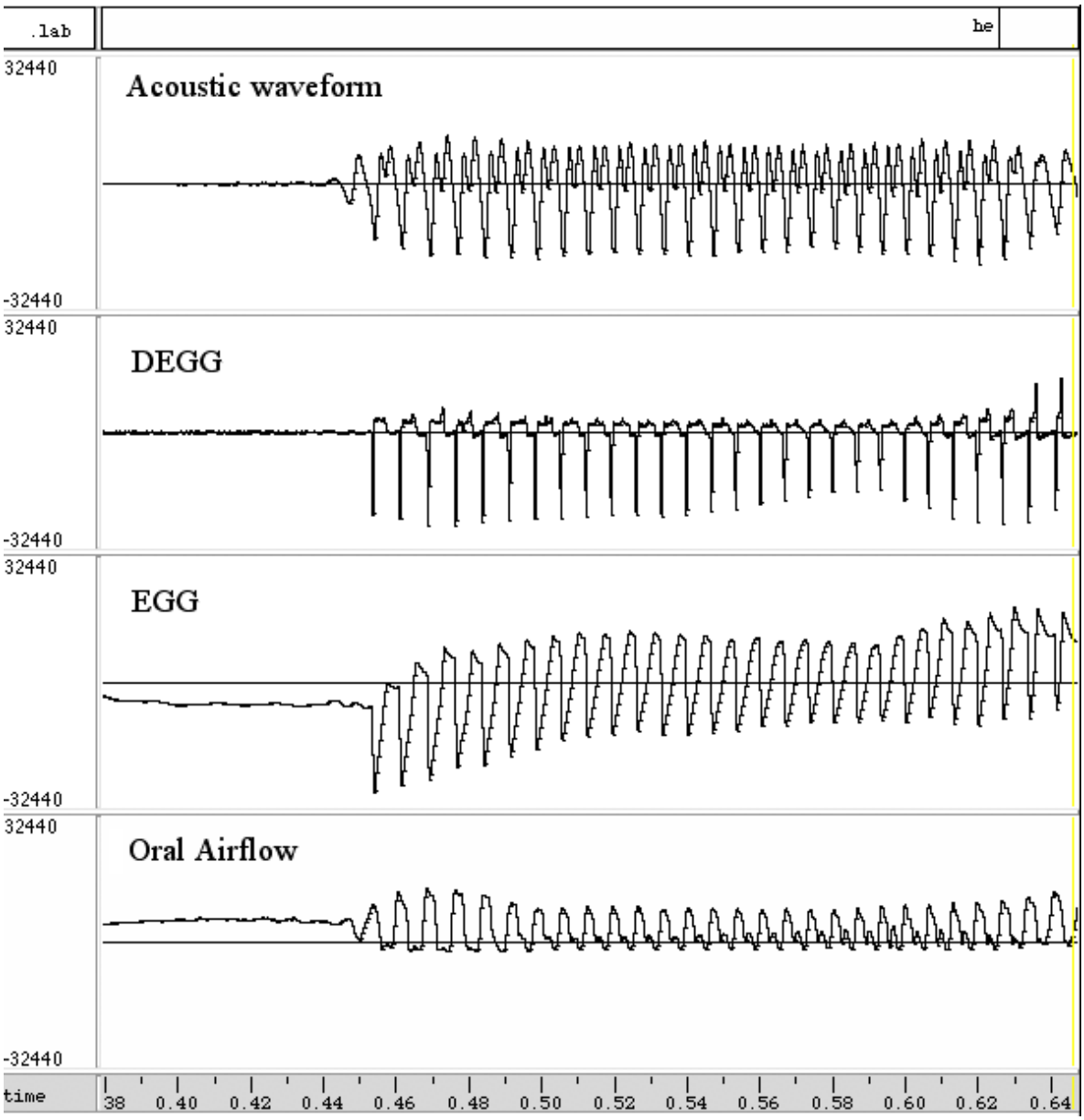

Figure 1. A four-channel recording of "he". The signals are sampled with 16 bit precision and at 44.05 kHz, but are not calibrated.

Figure 1 shows the basic display that we will be using in WaveSurfer. The sound file answer1_4chan.wav, as the name implies, is a 4 channel recording, with the speech waveform in channel 0, the EGG in channel 1, the DEGG in channel 2 and the oral airflow record in channel 3. These channels are all sampled at 44.05 kHz. To set up WaveSurfer with the display you should follow these steps (the steps with * are unnecessary if you use the Phonation configuration file):

(1) open the sound file "answer1_4chan.wav" using the "Phonation" configuration.
(2) in the "properties" pop up menu for the waveform (right click or control click the pane), specify that the "sound" has 4 channels.
*(3) create 3 additional waveform panes.
*(4) specify in the properties pop up menu that the "waveform" to be shown in the windows are (from top to bottom - 0, 2, 1, 3).
*(5) optional - you may want to reduce the pane height to 140 pixels to see all the panes at once.
(6) create a transcription pane and specify (properties load transcription) that the transcription can be read from "answer1_speech.lab".

You can read the file "answer1_speech.wav" into WaveSurfer or Praat if you would like to see a spectrogram of the acoustic waveform and check the word boundaries that are in answer1_speech.lab.

Question 1: Notice (figure 1) the oral airflow trace in the word "he". This word begins with an [h] and is followed by a word that begins with [h]. How does [h] affect the amount of airflow that escapes during each pulse of the vocal folds? What factors may have caused the airflow pattern that we see in this word?

EGG is derived from the electrical resistance of the neck. Electrodes are placed on either side of the neck and a small current is passed through. When the glottis is closed, more current can flow through then when the glottis is open. We will also be using the differential of the EGG (DEGG) to find the points of fastest glottal closing and glottal opening. See Childers and Lee (1991) for more discussion of EGG for different phonation types, and on the use of DEGG to estimate the open and closed portions of the glottal cycle (see their figure 10).

Figure 2 shows a blow up of the waveforms during syllabic [l] of "applicable". We marked the EGG trace with "closed" at the bottom and "open" at the top. As indicated in this figure during each glottal cycle - each puff of air through the glottis - there is a closing motion and an opening motion. The glottis vibrates open and shut. Figure 2 also shows how to use the DEGG signal to measure the closed portion of a glottal cycle (labeled "C") and the open portion of the cycle (labeled "O"). A downward peak in the DEGG marks the moment of glottis closing and an upward peak marks the moment of glottis opening. One important measure of the relative breathiness or creakiness of the voice is the amount of time that the glottis is open during each cycle. This is called the open quotient and can be calculated at O/(C+O).

Question 2. Measure the open quotient over the course of the [h] of "has" from time 0.72 to 0.82 in answer1_4chan.wav. How does OQ change over this interval?

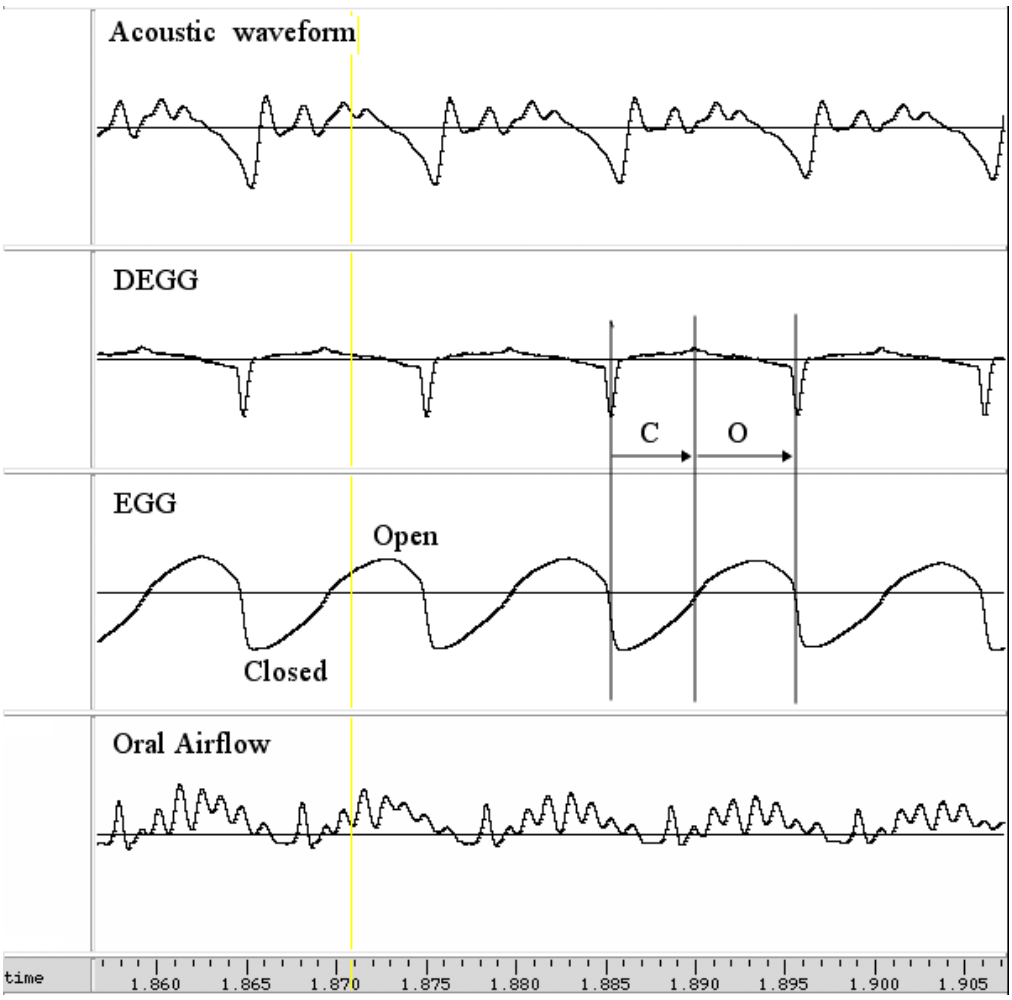

Figure 2. Five glottal cycles during the syllabic [l] of "applicable".

Question 3. Look at the pattern of EGG, airflow, and acoustics for the [h] of "who" in the interval from about 0.65 to 0.72 seconds. Is this fricative voiced? How do you explain the oscillations in the airflow and acoustic records?

Question 4. Now look at the [z] in "has" - examining the EGG, airflow and acoustics in the interval from 1.0 to 1.04 seconds. Is this fricative voiced? How do you explain the patterns seen here and how they differ from the pattern you examined in the [h] of "who"?

Question 5. The sentence that we chose for this exercise has some vowel - vowel sequences that might be realized with a glottal stop or some laryngealization: "the only" and "only applicable" and "applicable answer". Looking at the data, we do not see gross evidence of glottal stops in these VV transitions. Looking more closely though at the OQ is there evidence of laryngealization at these transitions?

Question 6. The sentence also has initial stops in "applicable", "keep", and "to" These are "voiceless" stops. Using the EGG, airflow, and acoustic data describe the sequence of events (in the traces themselves and then perhaps in the speaker's pronunciation) and measure the voice onset time of these

stops.

Question 7. Finally, we also have a couple of examples of final stops in "keep" and "it". Is there evidence that either of these final stops is accompanied with a simultaneous glottal stop? What pattern would you look for as evidence of this?

The distance of the microphone from the speaker's mouth. Notice that Childers and Lee (1991) identify the down peak in the DEGG signal as the "main excitation" of the vocal tract during voicing. The moment of closure produces the loudest noise of voicing. Note that there is a delay in these temporally aligned recordings so that the DEGG peak occurs slightly ahead of the largest (negative) peak in the acoustic signal. The delay between the two negative peaks is an indication of the distance (given the speed of sound) of the microphone from the vocal folds.