

LING205 Practical Phonetics

Spring Session 2006

Laboratory Manual

Geoffrey Stewart Morrison

Department of Linguistics, University of Alberta

4 May 2006

Table of Contents

| | | | |
|--|----|---|----|
| Basic Operations | 3 | Measuring Voice Onset Time | 23 |
| Downloading Praat | 3 | Graded Assignment 1 - Voice Offset Time | 26 |
| Starting Praat | 3 | Noise | 30 |
| Getting Help | 4 | Graded Assignment 2 - Fricative Spectra | 31 |
| Loading a Sound | 5 | Graded Assignment 3 | |
| Removing an Object | 6 | - Synthesis of Vowel-like Sounds Using Sine Waves | 38 |
| Recording a Sound | 7 | Source and Filter Synthesiser | 42 |
| Configuring sound recording options for Windows XP | | Graded Assignment 4 - Measuring Vowel Formants | 47 |
| | 7 | Vowel-Consonant Perception Experiment | 53 |
| Recording with Praat | 8 | | |
| Saving a Sound | 9 | | |
| Editing a Sound | 10 | | |
| Recording problems: | 11 | | |
| Making a TextGrid | 12 | | |
| How to mark a boundary: | 14 | | |
| Other things to try: | 16 | | |
| Saving a TextGrid | 17 | | |
| Drawing Pictures | 18 | | |
| Running a Script | 21 | | |
| | | © 2005/2006 Geoffrey Stewart Morrison | |

Basic Operations

All the assignments described in this manual are performed using the *Praat* software system developed by Paul Boersma & David Weenink at the University of Amsterdam. It is recommended that students work through the earlier sections of the manual and perform the non-graded exercises so as to familiarise themselves with Praat before starting the graded assignments.

Downloading Praat

1. Go to <http://www.praat.org> and download the latest version of Praat appropriate for your operating system.
2. Follow the instructions on the webpage as to how to install Praat.
 - you may want to manually make a shortcut to Praat so that it easier to start the program
3. Download the *silipa93* font (if available for your operating system) and follow the instructions on the webpage as to how to install the font.

All the examples in this manual were produced using version 4.4.17 of Praat running under the Windows XP operating system.

Starting Praat

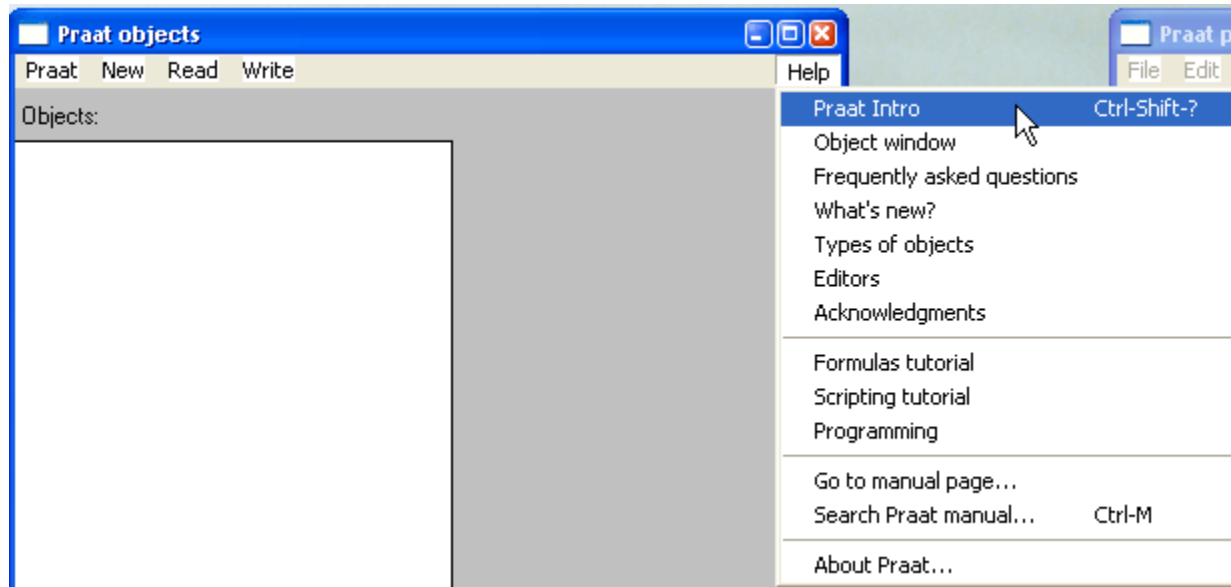


Double click on the Praat *program file*, or the *shortcut* if you created one.

- You must always start Praat and then open other files using the *Praat > Open Praat Script...* or the *Read* menu items.
- You cannot launch Praat by double clicking on a sound file, TextGrid file, etc.

Getting Help

Under the Praat help menu there are various links to different parts of the Praat Manual. Start by looking at the *Praat Intro* section.



Other sources of help:

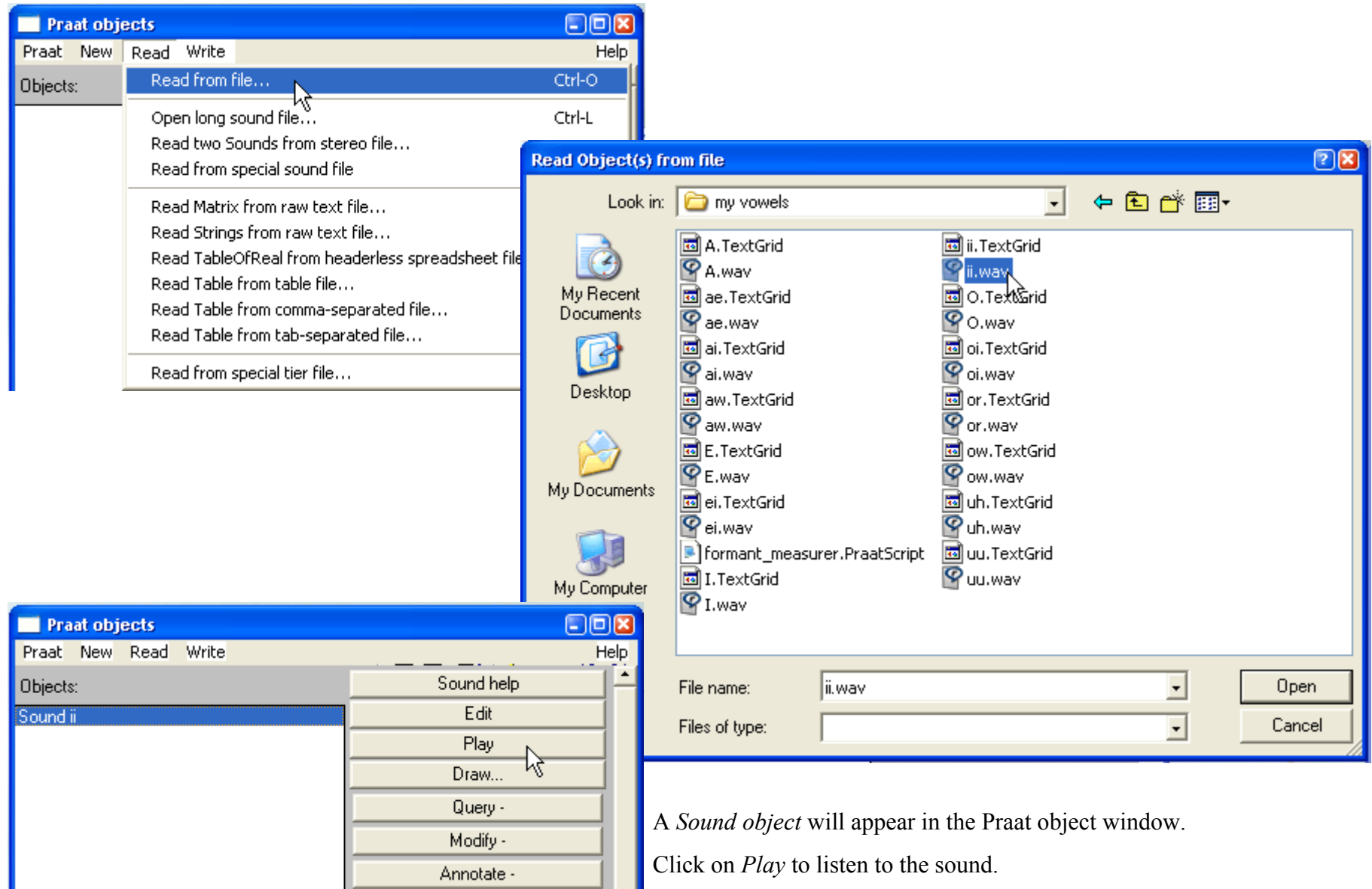
Sidney Wood, Lund University, *Beginners' Guide to Praat*

<http://www.ling.lu.se/persons/Sidney/praate>

Pascal H. H. M. Van Lieshout, University of Toronto, *Praat Tutorial*

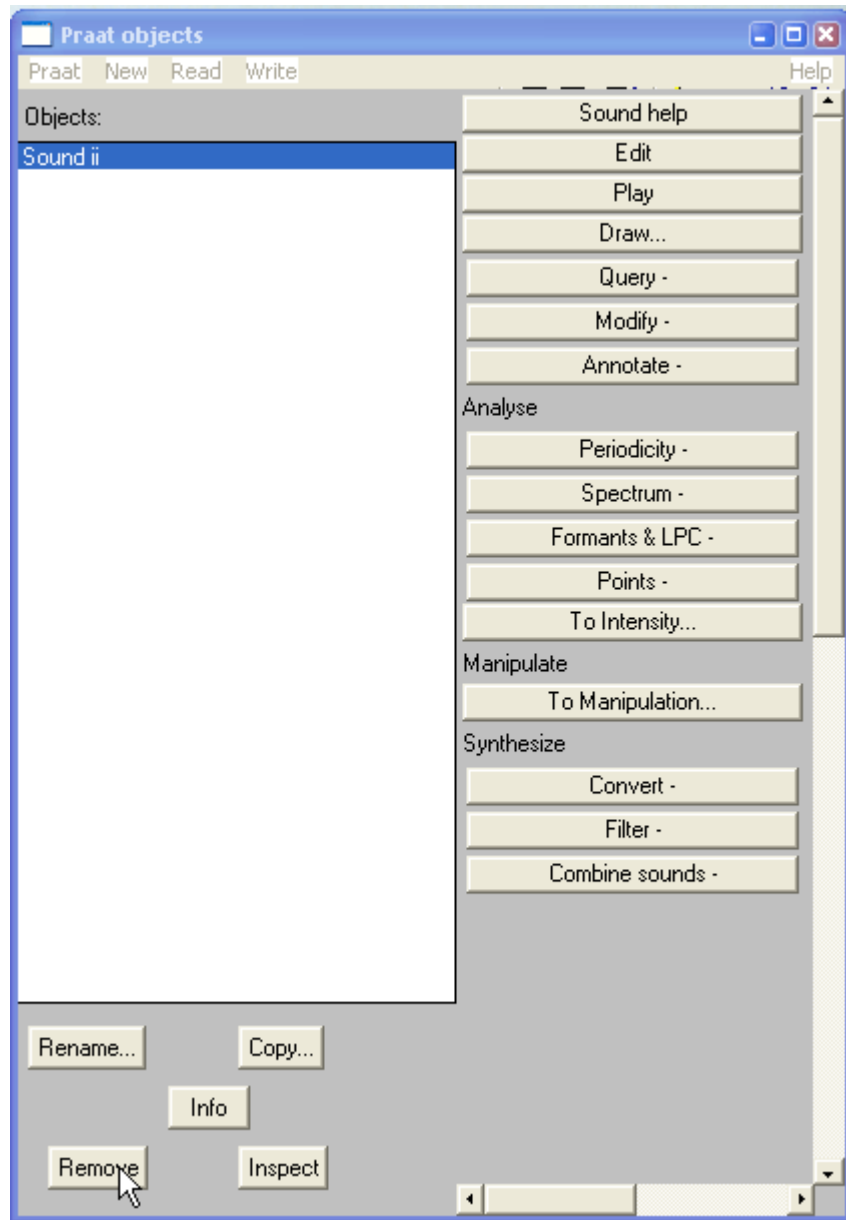
<http://www.slp.utoronto.ca/English/page-1-364-1.html>

Loading a Sound



A *Sound object* will appear in the Praat object window.

Click on *Play* to listen to the sound.



Removing an Object

- click on the object
- click on the *Remove* button

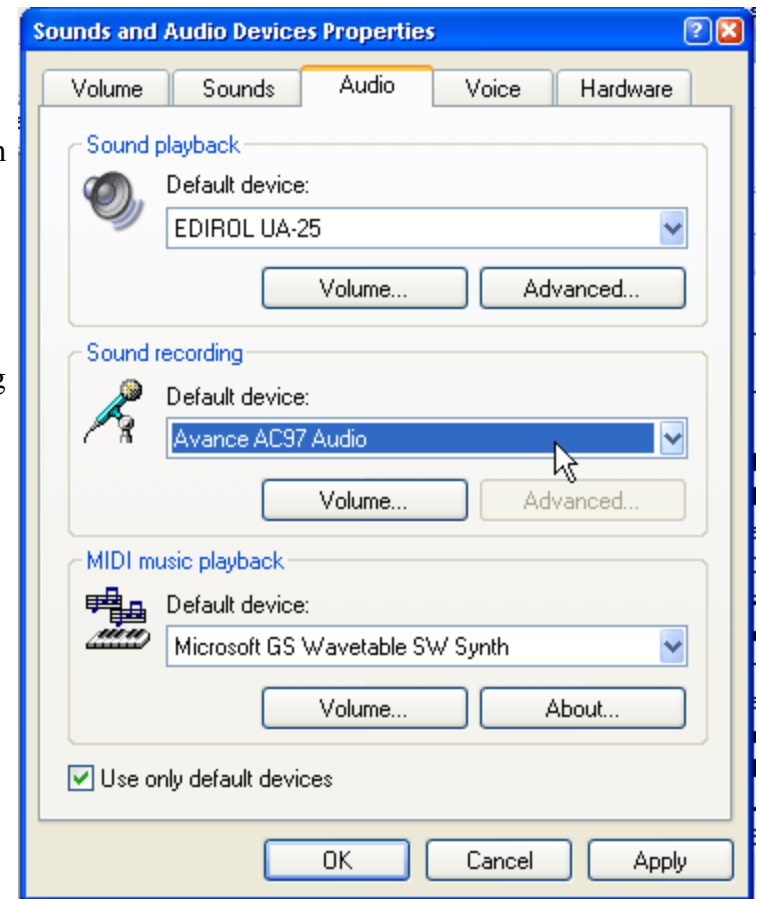
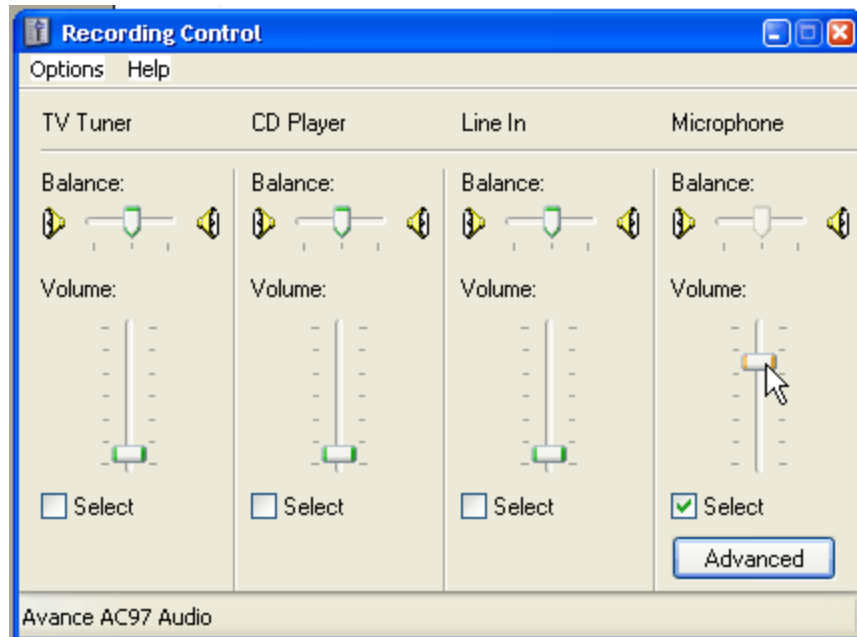
You can also rename or copy an object using the *Rename...* or *Copy...* button.

To remove multiple objects:

- hold down the CTRL or SHIFT key and click on the objects to be removed
- click on the *Remove* button

Recording a Sound

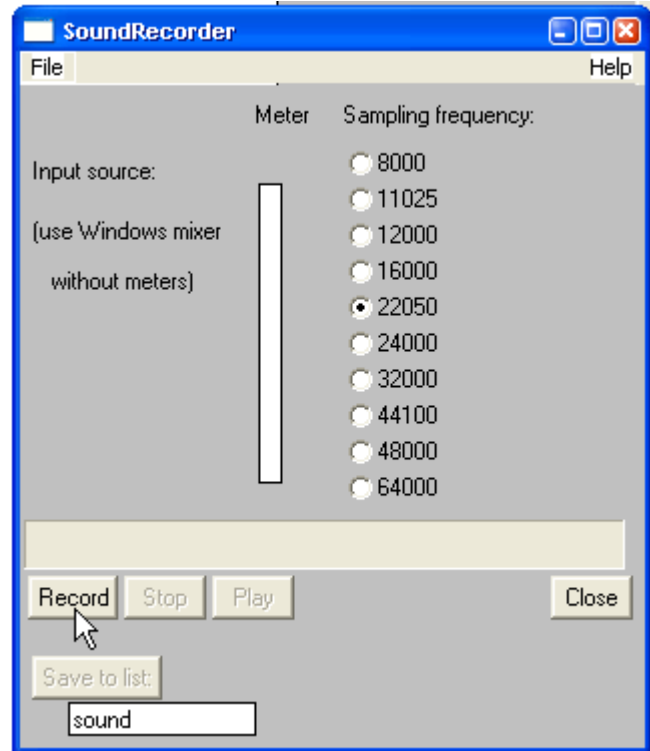
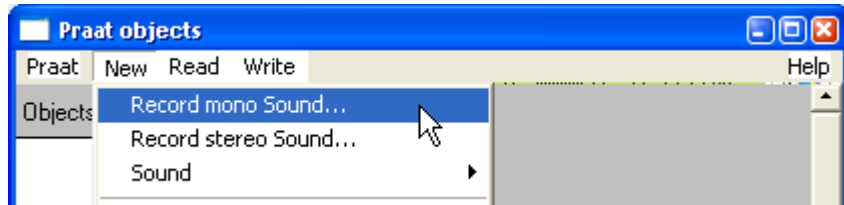
- plug in a microphone
- you may need to configure the sound recording options for your operating system
 - Configuring sound recording options for Windows XP
 - go to the *Control Panel*
 - select *Sounds and Audio Devices*
 - under the *Audio* tab, select the soundcard you are using for recording
 - click on the *Apply* button
 - click on the *Volume* button
 - click on the *Microphone Select* checkbox
 - adjust the recording *Volume*



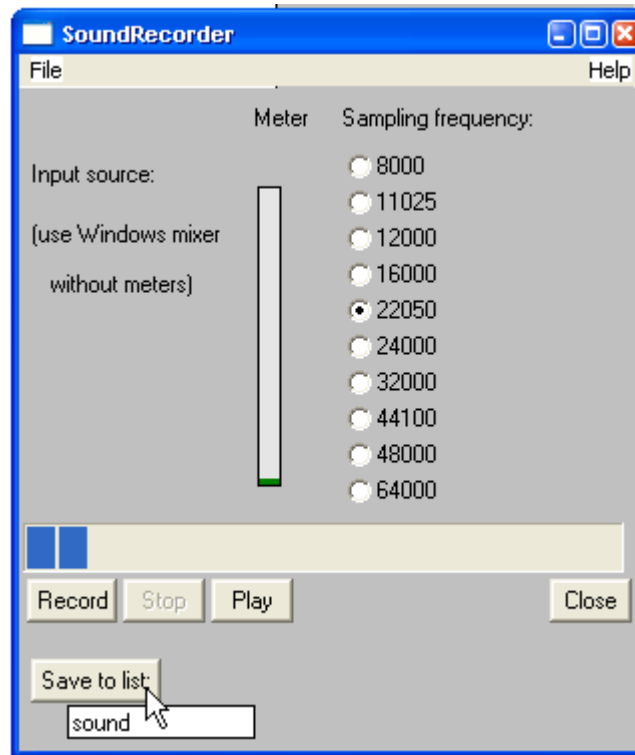
The instructions starting on the next page, explain how to record sound directly into Praat.

You may prefer to record sound files using some other program, then open the sound files in Praat. A free sound recorder you could try is *Audacity* <http://audacity.sourceforge.net/>

Recording with Praat



A sampling frequency of 22.050 kHz is usually adequate, or more than adequate, for phonetic analysis



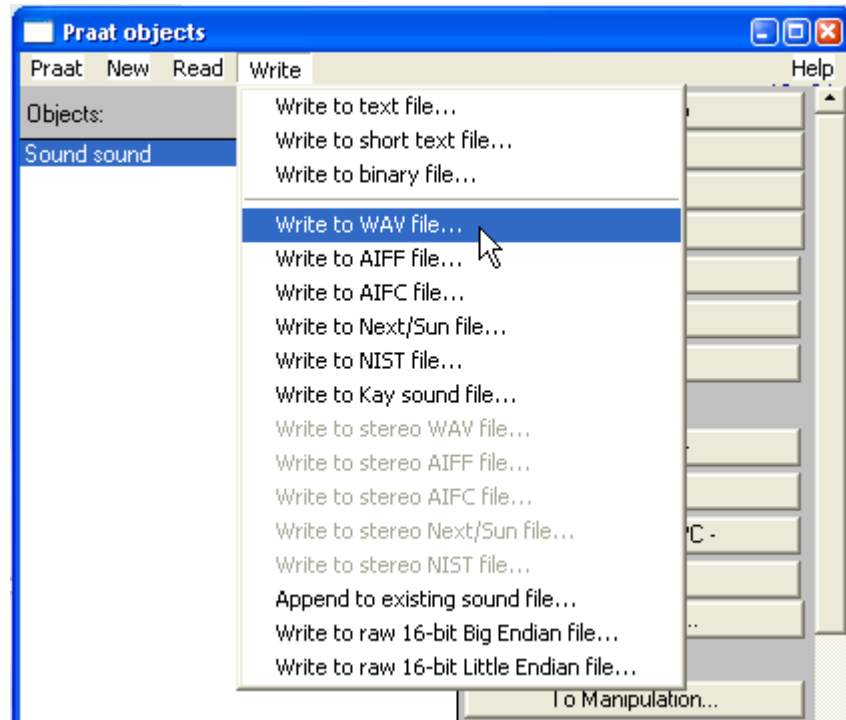
– click on the *Record* button to begin recording, and the *Stop* button to stop recording

– click on the *Save to list* button to create a sound object in the Praat objects window

– in the box below the *Save to list* button you can type in a name for the sound object, the default is “sound”

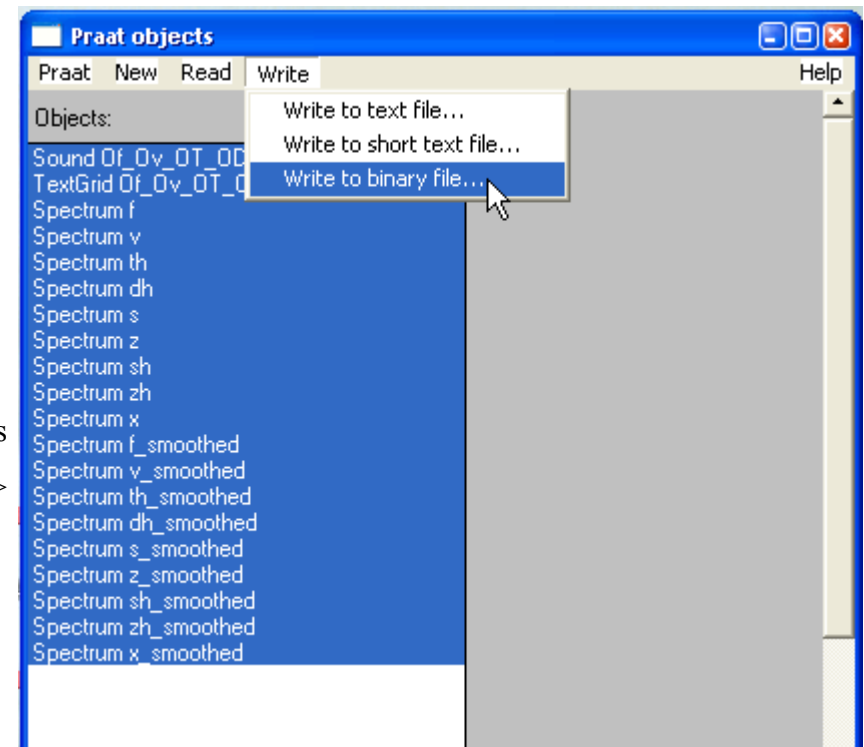
Don't forget to click on the *Save to list* button!

Saving a Sound

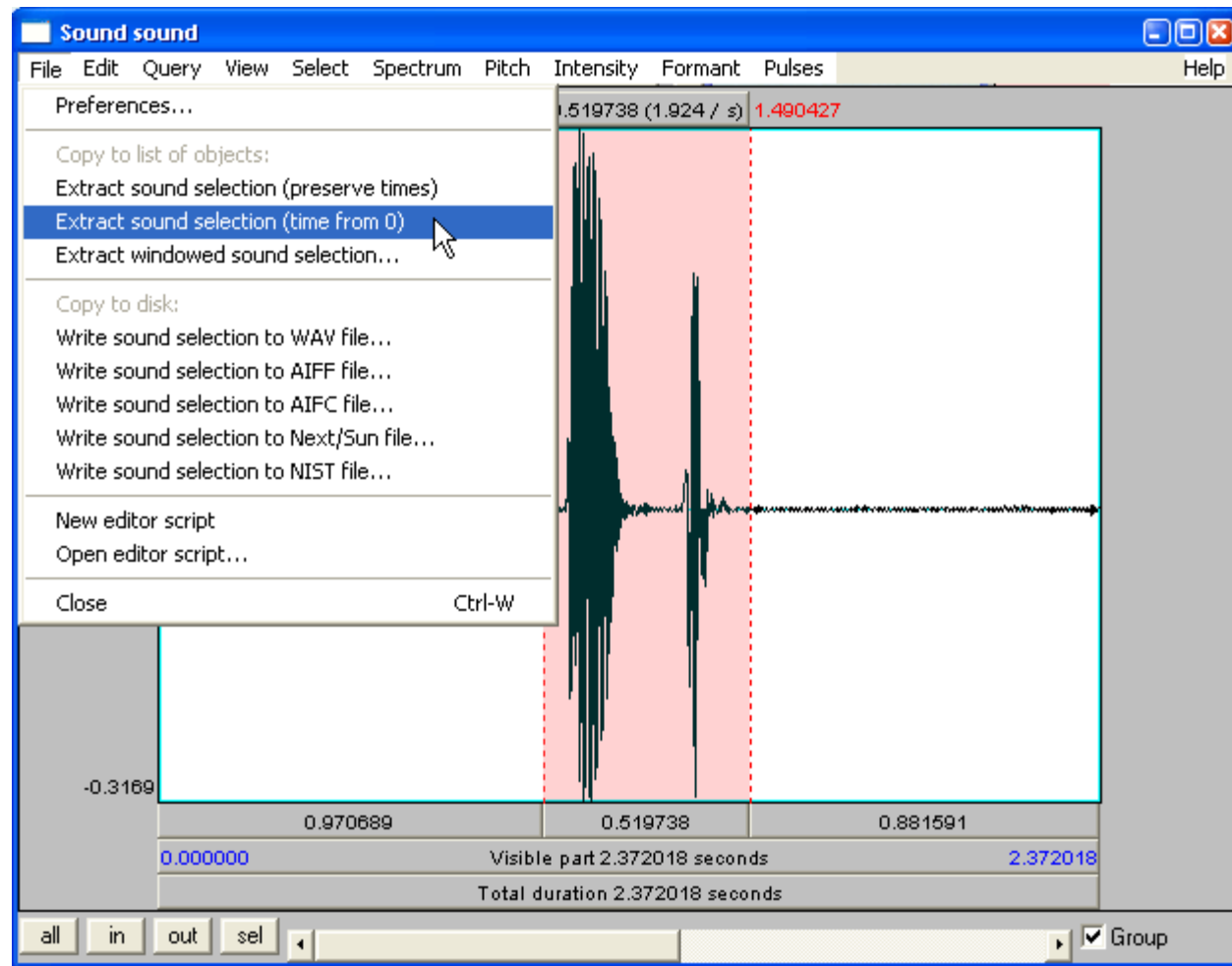


If you want to use the sound at some time in the future, you will need to save a copy to disk. For graded assignments, it is highly highly recommended that you save a copy of everything you do.

– choose a sound-file format such as *.wav*



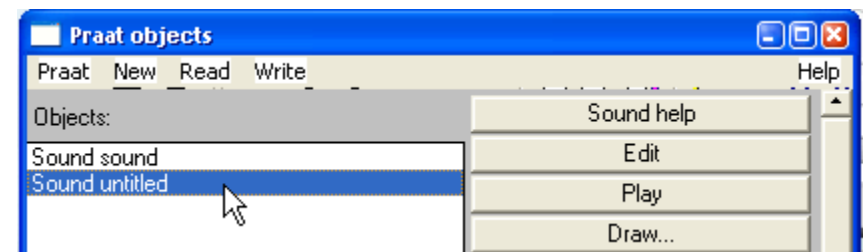
– you can save a group of Sound, TextGrid, Spectrum and other objects together in a single file, select all the objects you want to save then *Write > Write to binary file...* the file extension will be *.Collection*

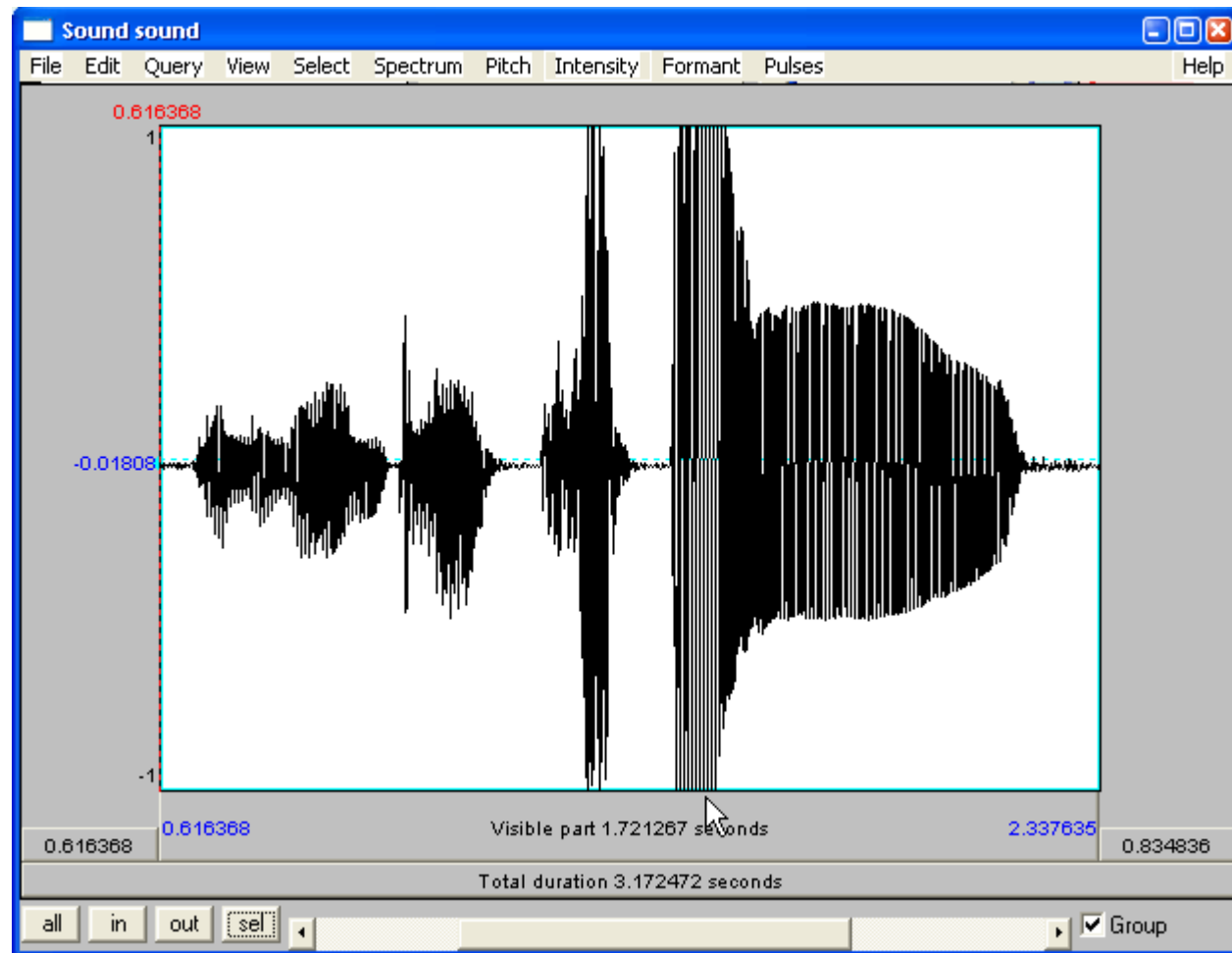


Editing a Sound

You may have some unwanted recording at the beginning and end.

- select the sound object in the object window
- click on the *Edit* button
- click and drag to select the part of the waveform that you want
- *File > Extract sound selection*
- in the object window, select the new sound object “untitled” and rename it something more meaningful
- save the new sound object as a sound file and remove the old sound object



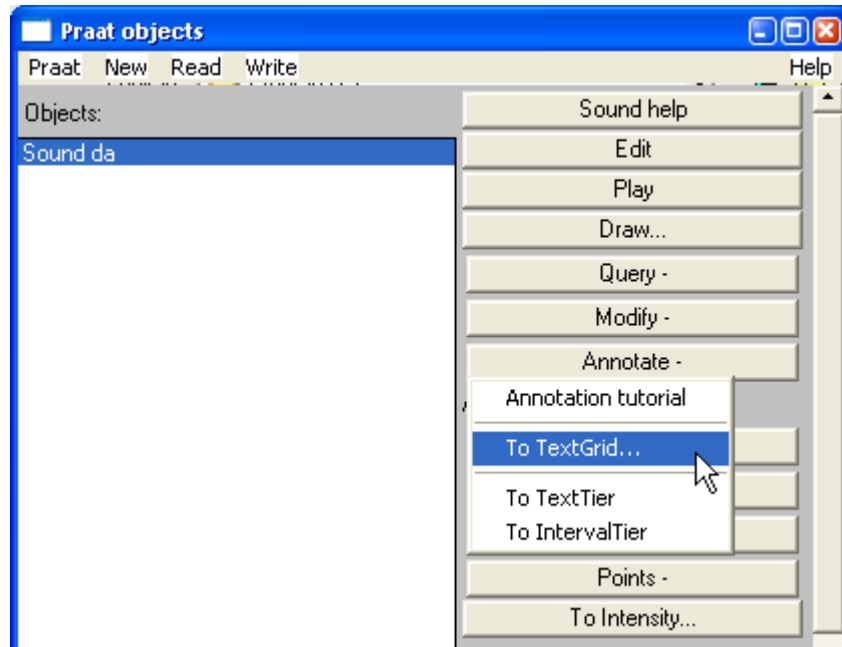


Recording problems:

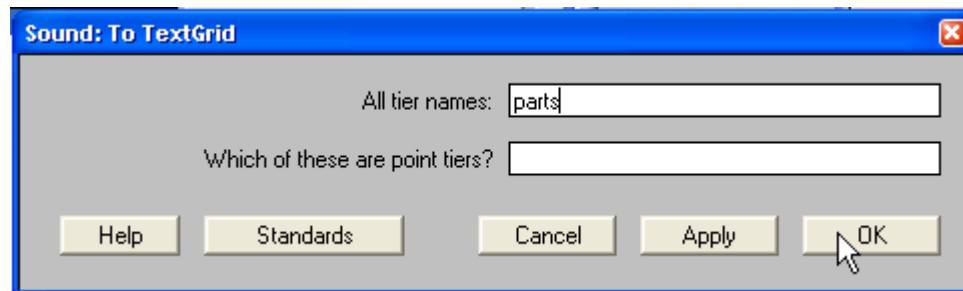
- One possible recording problem is that the recording is too quiet. Try turning up the operating system's microphone volume.
- Another problem is caused if the operating system's microphone volume is set too high. This causes *clipping*. A picture of clipping is given here. Recordings with clipping sound fuzzy. Try turning down the operating system's microphone volume.

- If the recording is only a little too quiet, you can select the sound object, then click on *Modify* -> *Scale Peak...* and set the *New Maximum Amplitude* to 0.99. This won't work well if the recording is very quiet because you are also amplifying any noise in the recording.
- Reducing the amplitude of a clipped recording won't help!

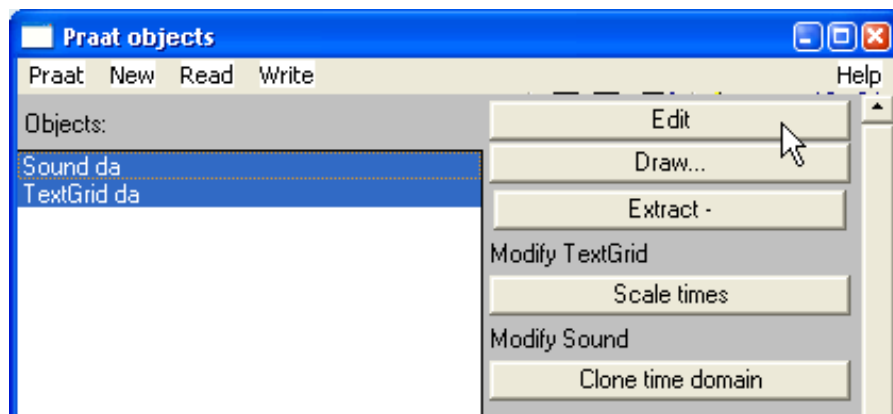
Making a TextGrid



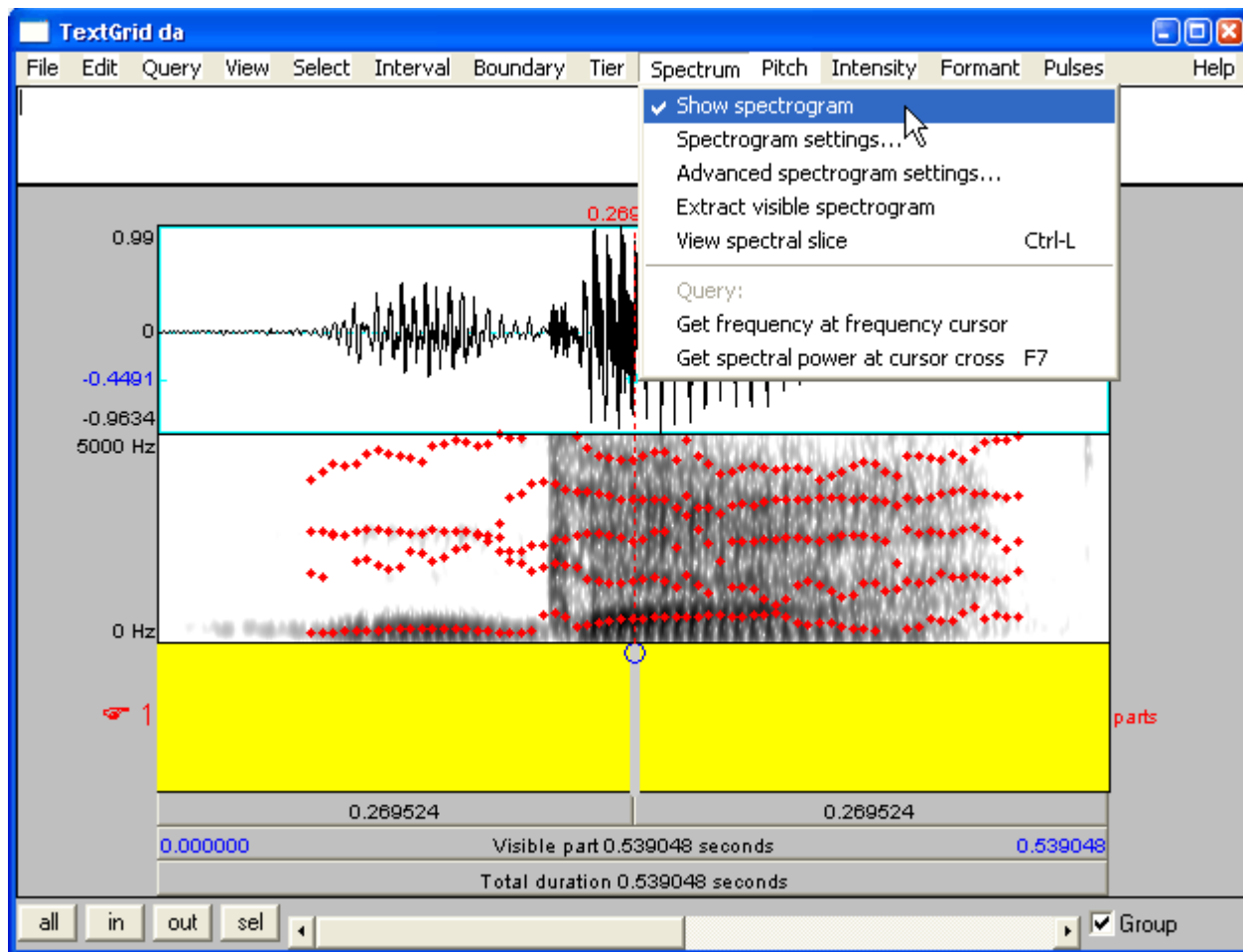
- record or load a sound object
- click on the sound object to select it
- click on the *Annotate* - button
- select *To TextGrid...*



- clear the default tier names
- type in a tier name
 - two names separated by a space for two tiers etc.

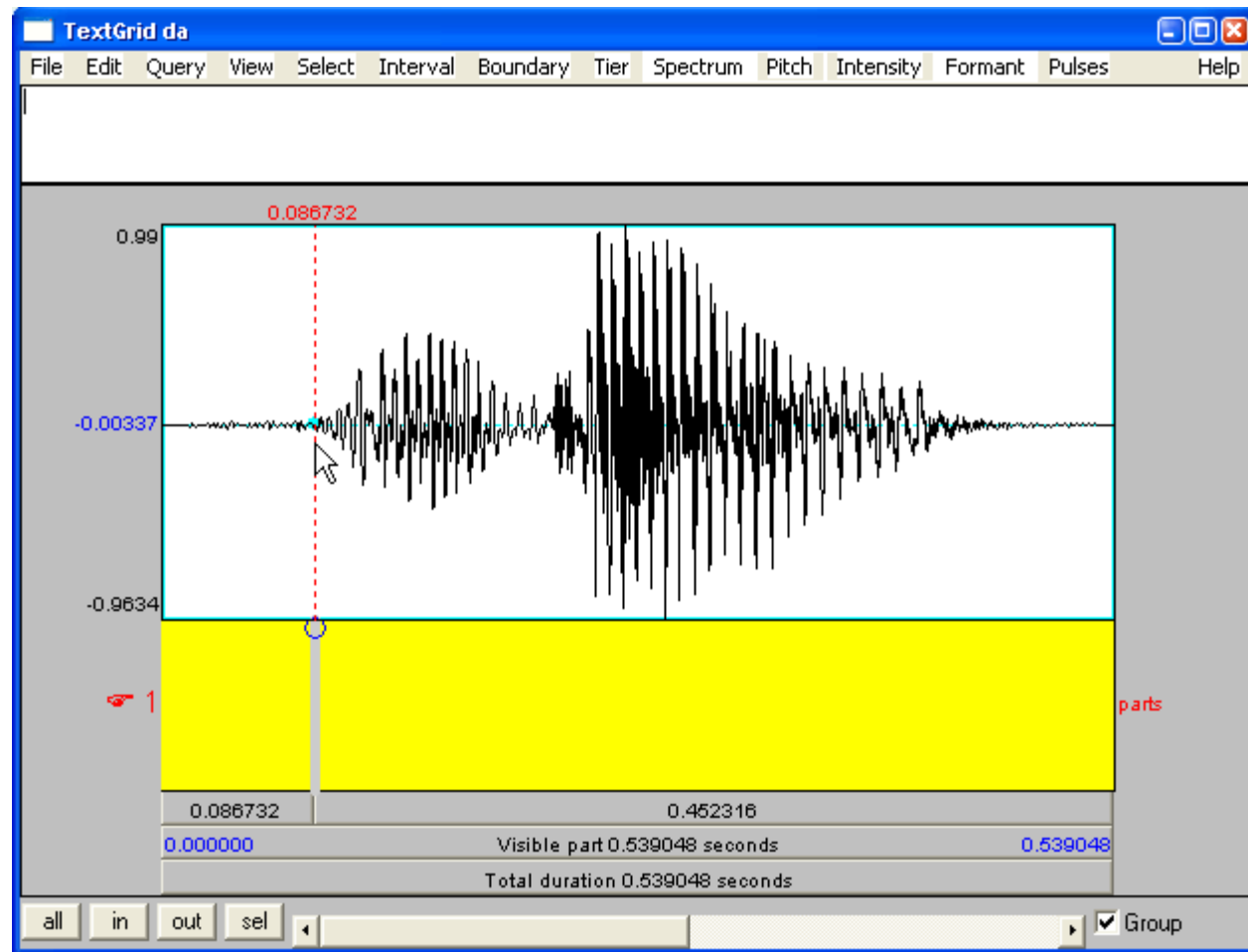


- select both the Sound object and the TextGrid object
- hold down the CTRL key and click on both objects
- click on the *Edit* button



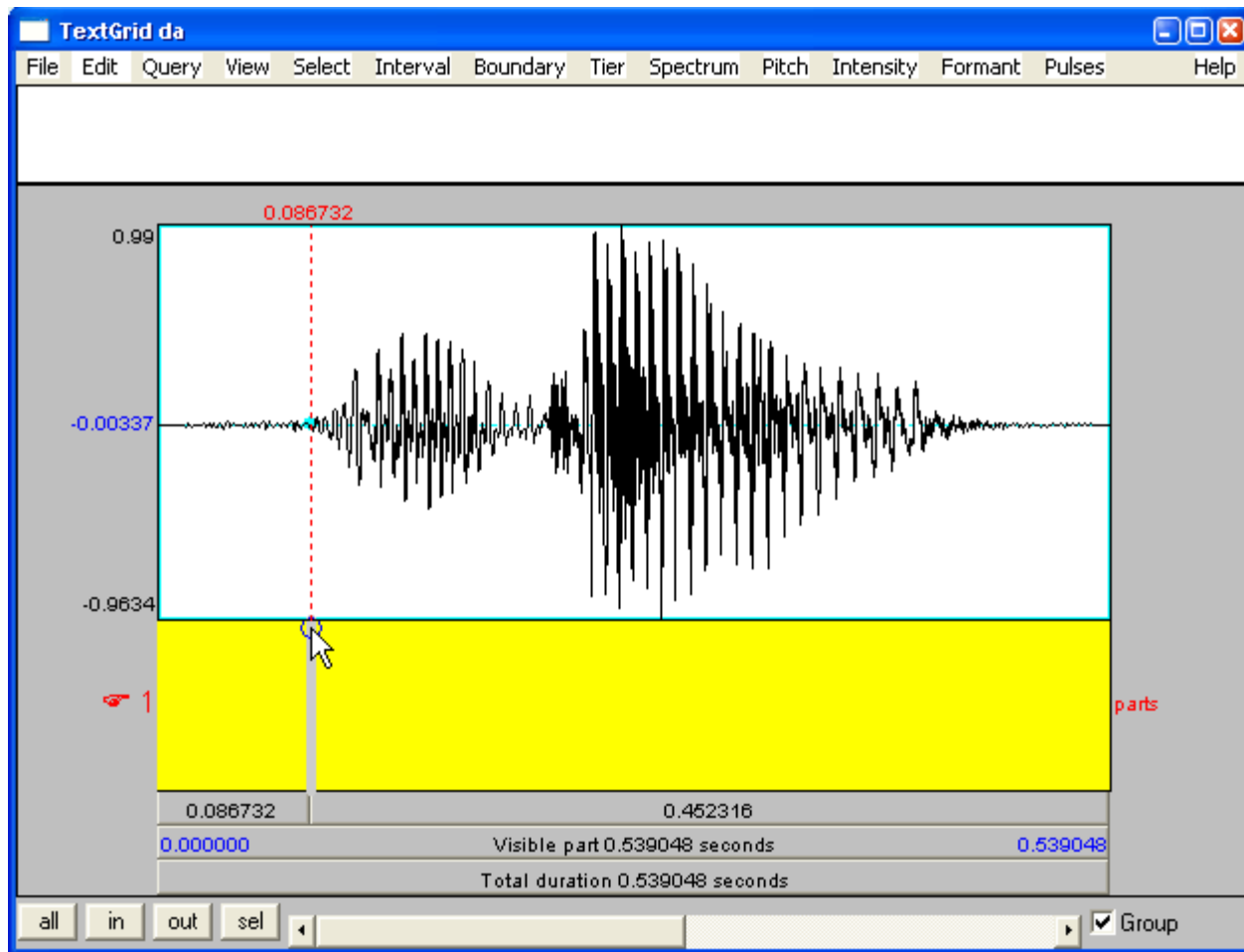
The default view may show several things we don't need right now.

- click on *Spectrum*
 - uncheck *Show Spectrogram*
- click on *Pitch*
 - uncheck *Show Pitch*
- click on *Formant*
 - uncheck *Show Formants*
- click on *Intensity*
 - uncheck *Show Intensity*
- click on *Pulses*
 - uncheck *Show Pulses*



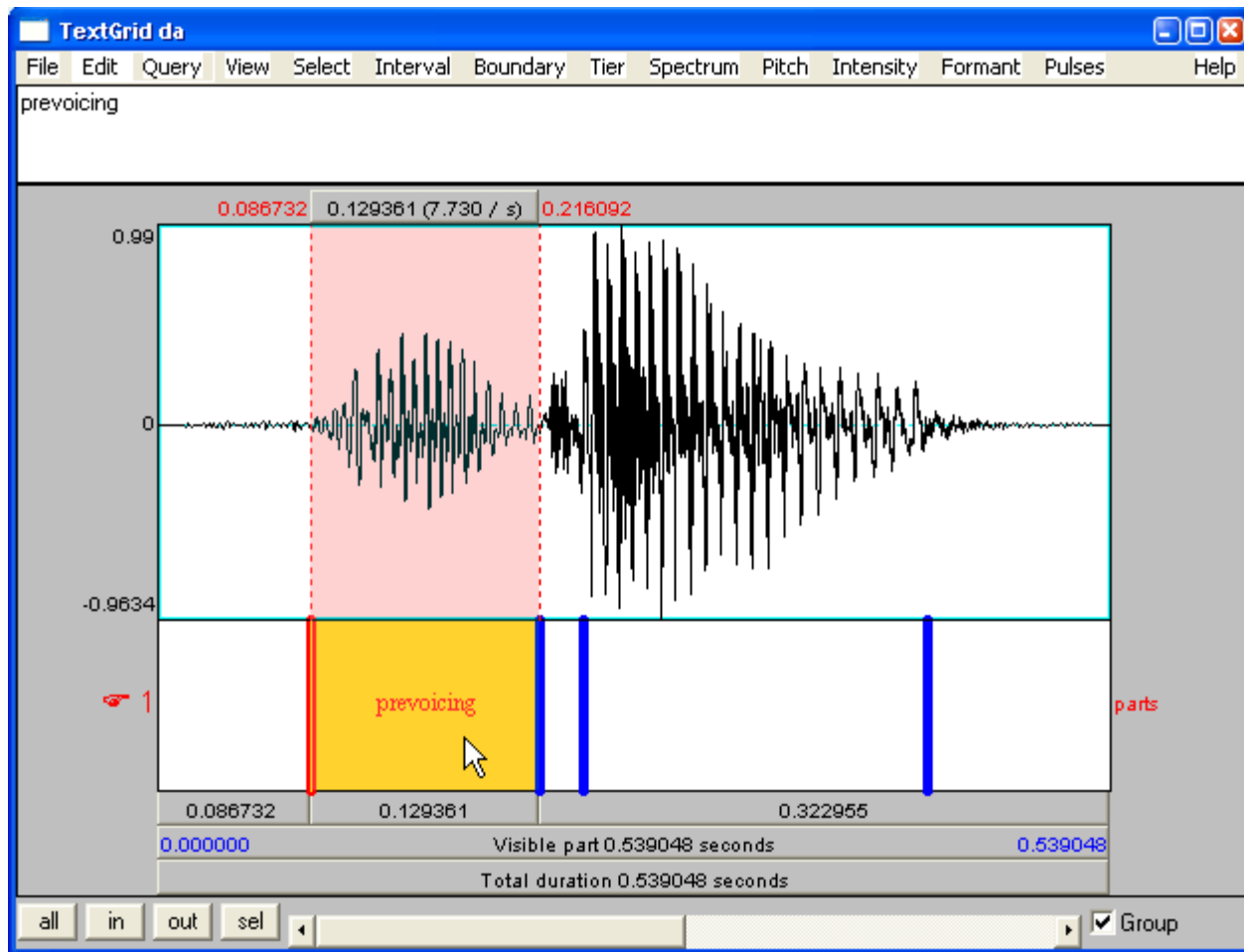
How to mark a boundary:

– click on the waveform part of the display



– click on the circle at the top of the text tier in the TextGrid portion of the display

– repeat for as many boundaries as needed

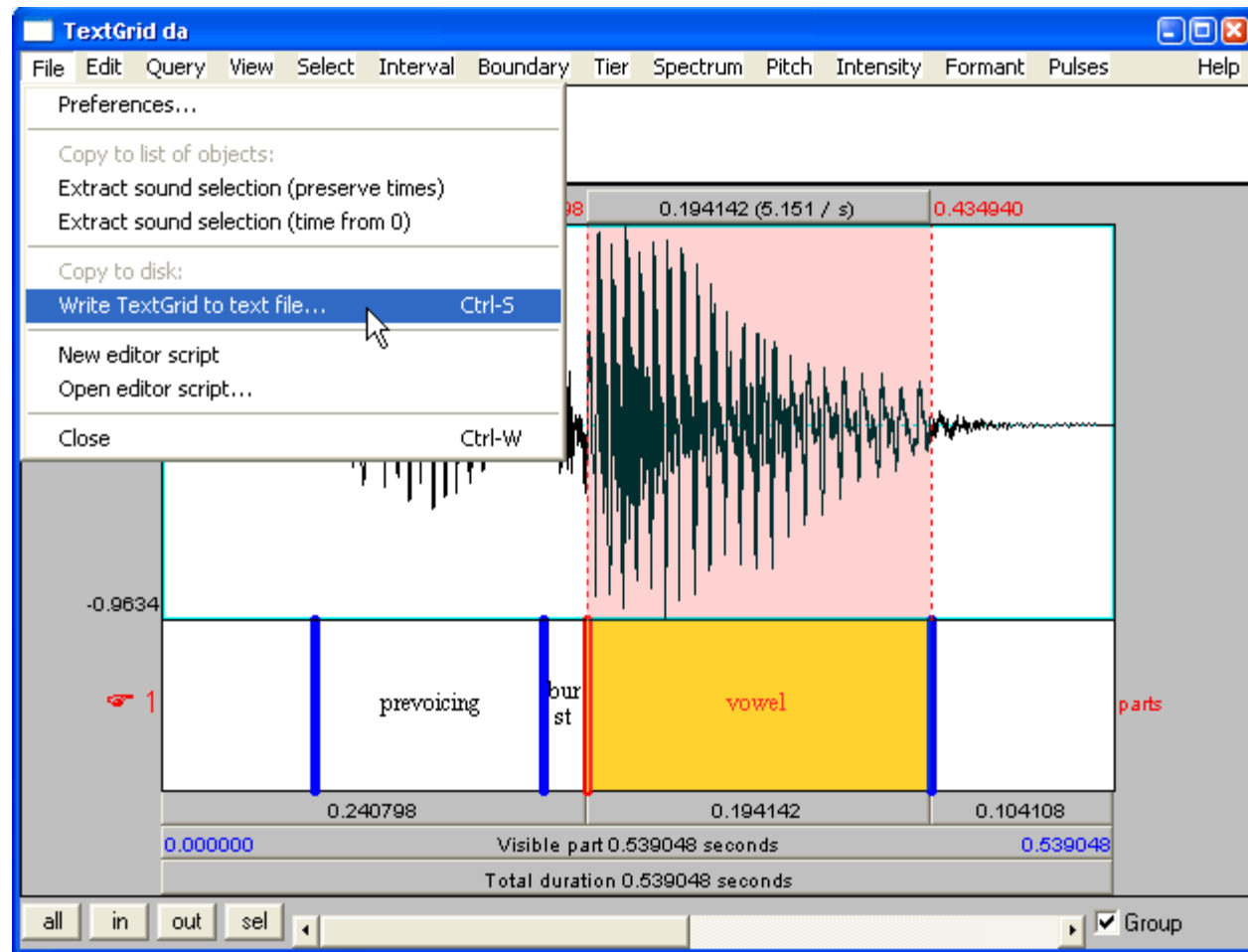


- click on an interval
- type in a label
- phonetic symbols can be typed, look under the *Help* menu item in the Edit window.

Other things to try:

- click and drag in the waveform portion of the display
- click on the *all*, *in*, *out*, and *sel* buttons
- click on each of the three bars just below the TextGrid portion of the display, including different portions of the topmost bar

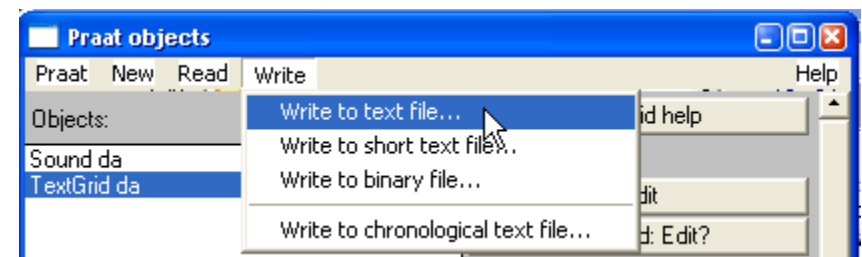
The numbers in the bars indicate the duration of parts of the sound, e.g., the part marked “prevoicing” is 129 ms long.



Saving a TextGrid

If you want to use the TextGrid at some time in the future, you will need to save a copy to disk. For graded assignments, it is highly highly recommended that you save a copy of everything you do.

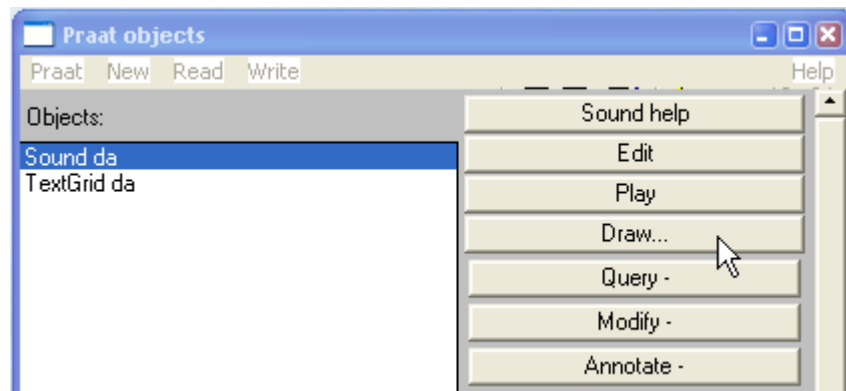
– I always add a *.TextGrid* extension to the filename



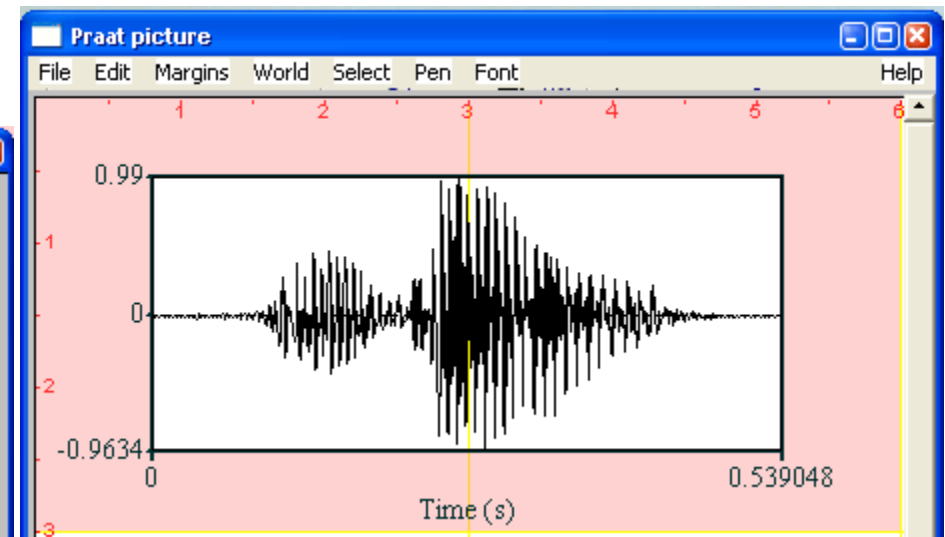
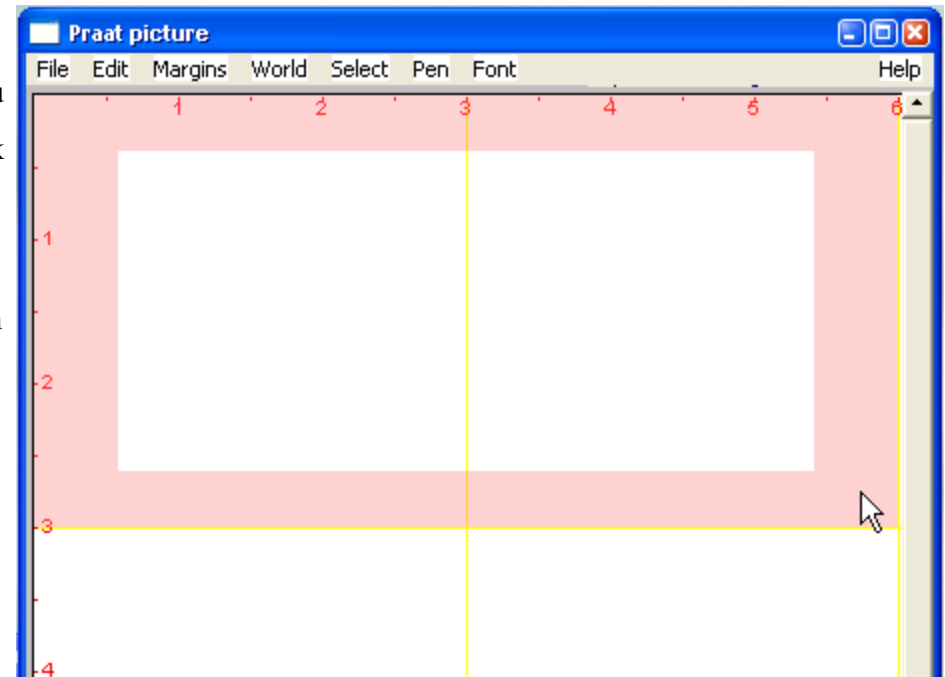
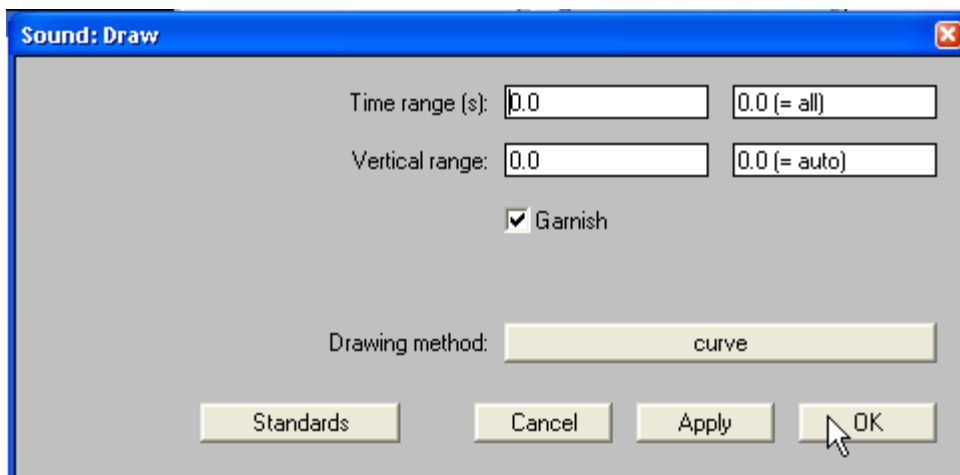
Drawing Pictures

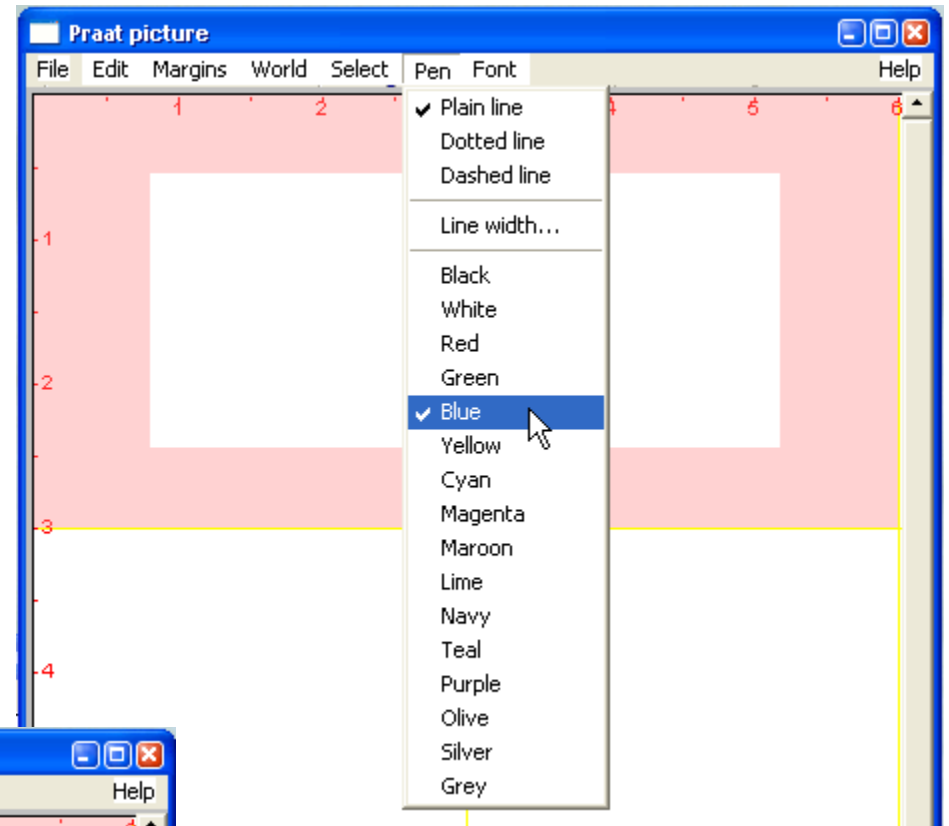
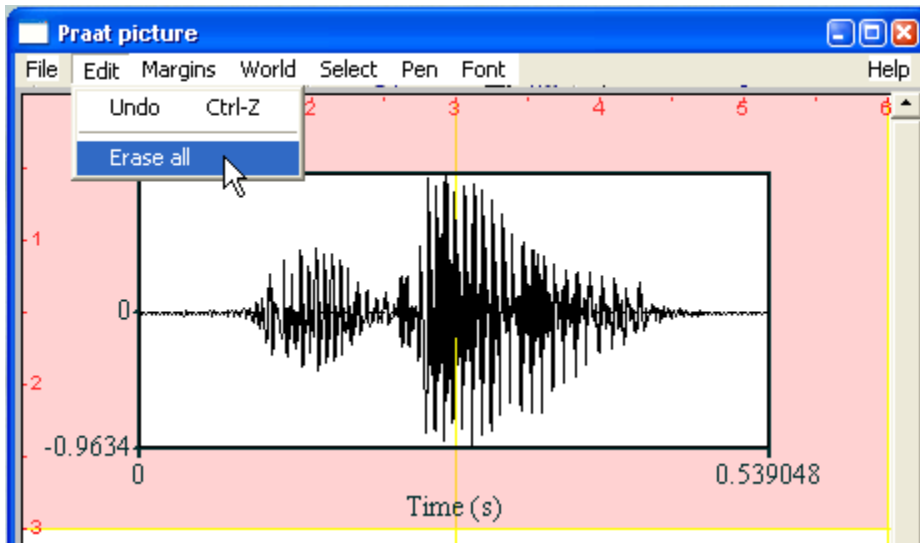
– In the Praat picture window, click and drag to select the area you want the picture to be drawn in. The picture will be drawn in the pink rectangle.

– select the object you want to draw and click on the *Draw...* button

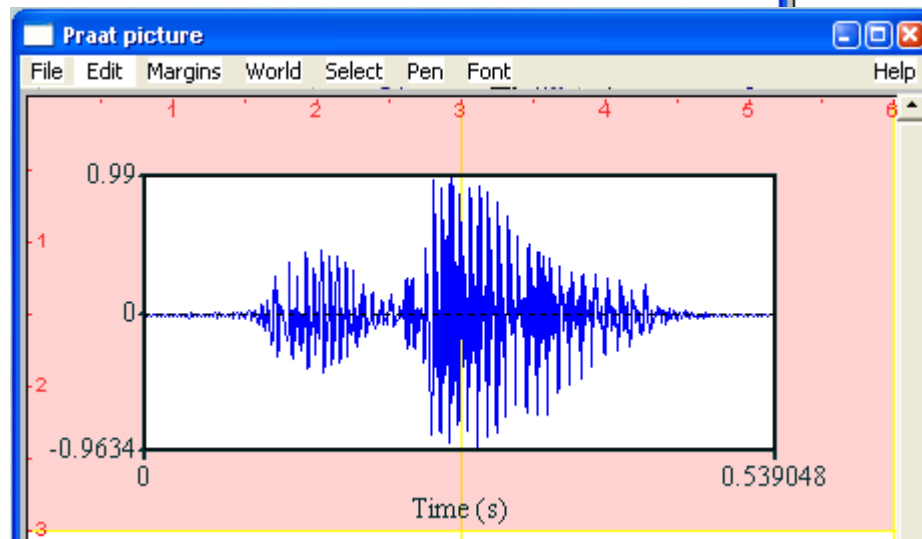


– set the options you want, or use the defaults

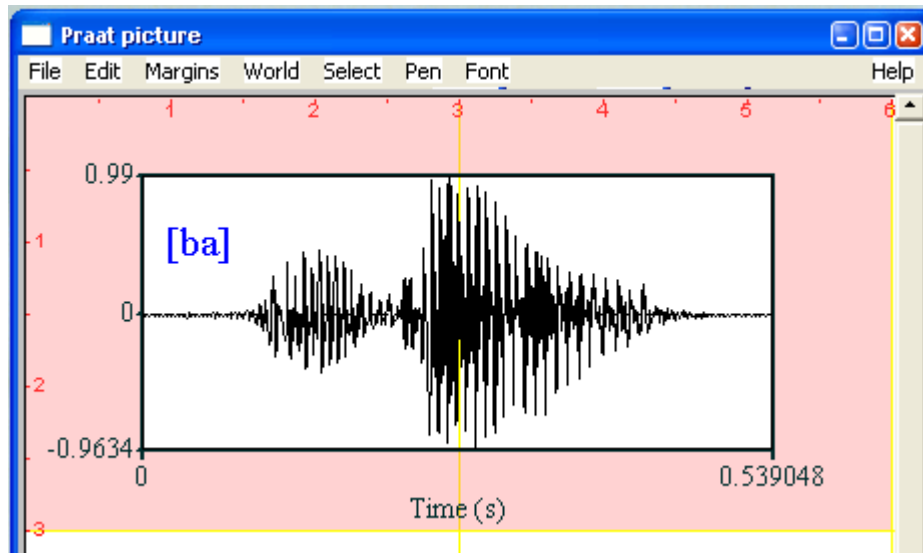




– set *Pen* and *Font* options before drawing

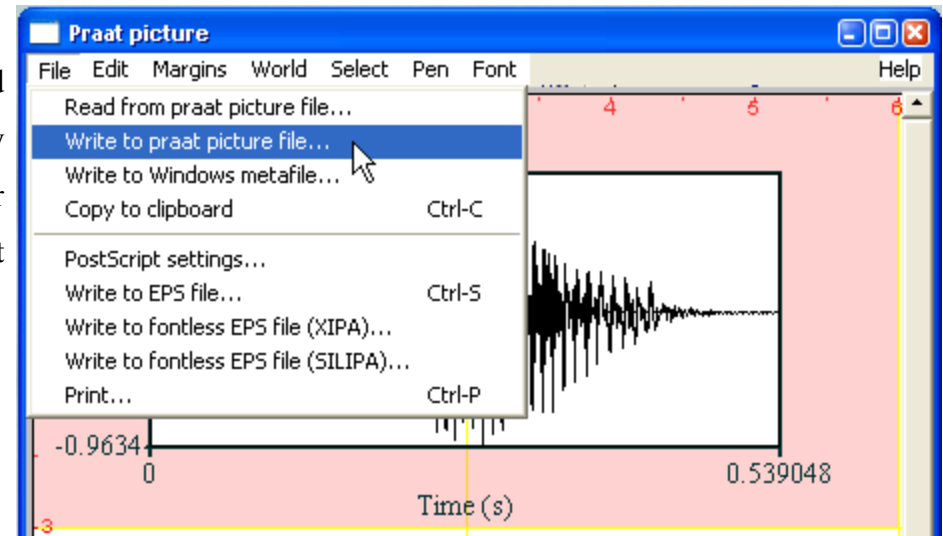


- If you select multiple objects and the *Draw...* button is visible, then you can draw these objects together.
- If you don't erase the picture then you can draw one object on top of another one at a time; try each object in a different colour, and unchecking the *Garnish* option after drawing the first object.

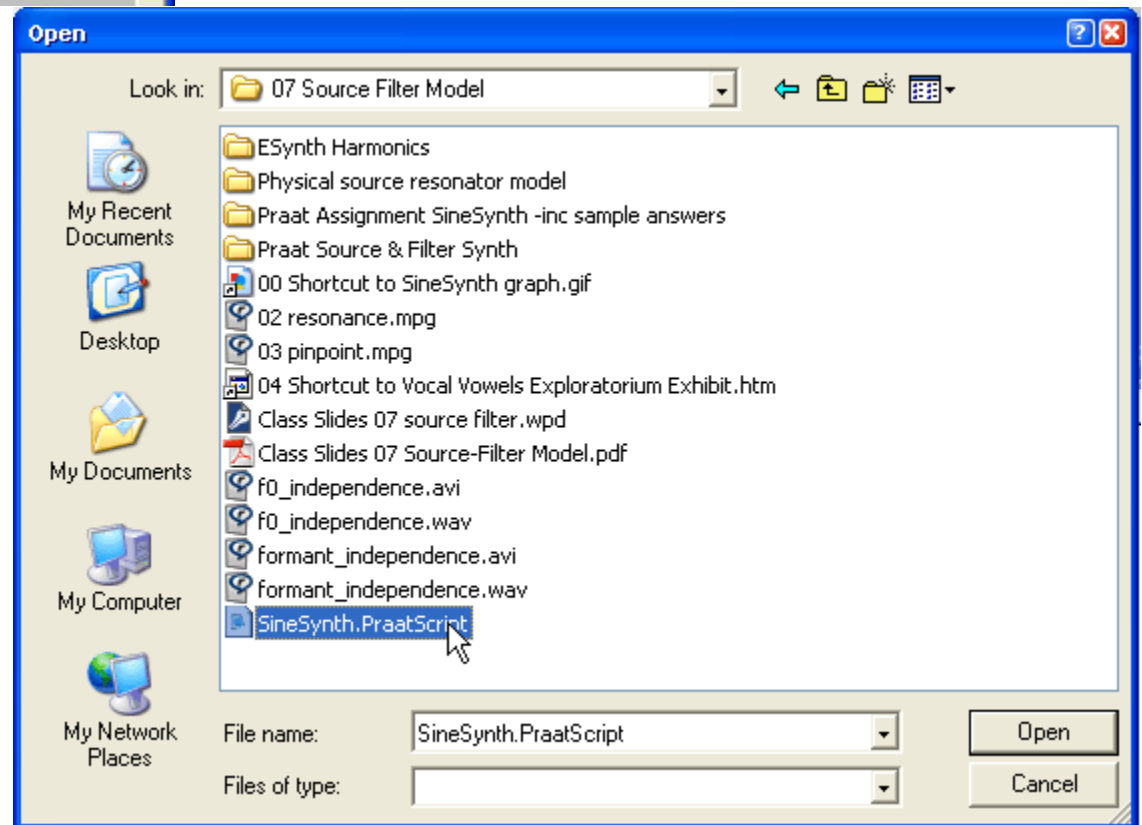
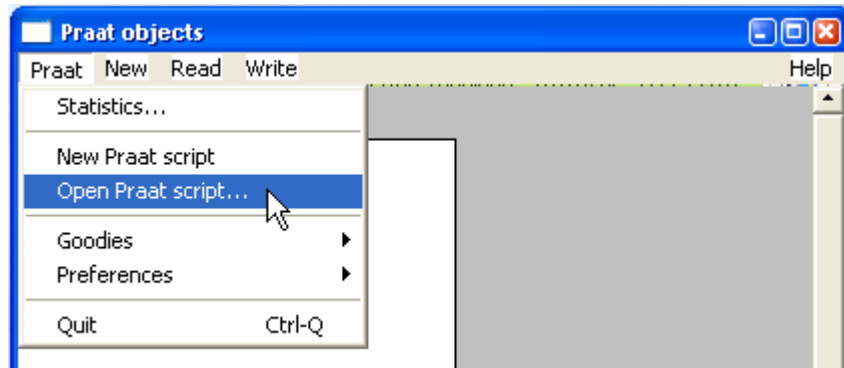


- To find out how to add text and other annotations, look under the *World* menu item and try different options.
- Look under *Help* to find out how to type in phonetic symbols.

If you want to use the picture at some time in the future, you will need to save a copy to disk. For graded assignments, it is highly highly recommended that you save a copy of everything you do. For assignment reports, you may also want to copy the picture and paste it into a word processor.



Running a Script



```
# SineSynth.E if Morrison July 2005 http://www.ualber
# Demonstrate can be combined to create vowel-like sounds
#
# This assignment is designed to give you some hands-on experience with s
# The intention is to give you some practical experience so that you will
# Performing this exercise conscientiously will make it easier to underst
# Please read the instructions carefully and take your time synthesising
# Play with the synthesiser until you think you have a good feel for the
#
# The default values (Fundamental Frequency) F1 = 100, (First Formant) F1
#
# The first three sine waves draw correspond to the Fundamental Frequency
#
# The fourth wave drawn (Combined_Frequencies_equal_amp) is the result of
# the amplitude (heigh on the graph) at each point in time in the C
#
# The fifth wave drawn (Combined_Frequencies_tilted_amp) is the result of
# the lowest frequency sine wave (the Fundamental Frequency) has th
# (in the fourth wave all the sinewaves added had the equal amplitu
#
# The sixth graph is a frequency domain representation of the fifth wave
# Note that there are spikes at the frequency of each of the compon
# The spectral amplitude tilt is -6dB per ocatve (the amplitude dec
#
# Try the following formant frequencies:
# F1 = 300 F2 = 870
# F1 = 730 F2 = 1090
# Which vowels do these sound like?
#
```

Measuring Voice Onset Time

– Record yourself saying a few plosive-initial words in a language you speak

e.g. for English possible words include:

/p/ /b/ /t/ /d/ /k/ /g/

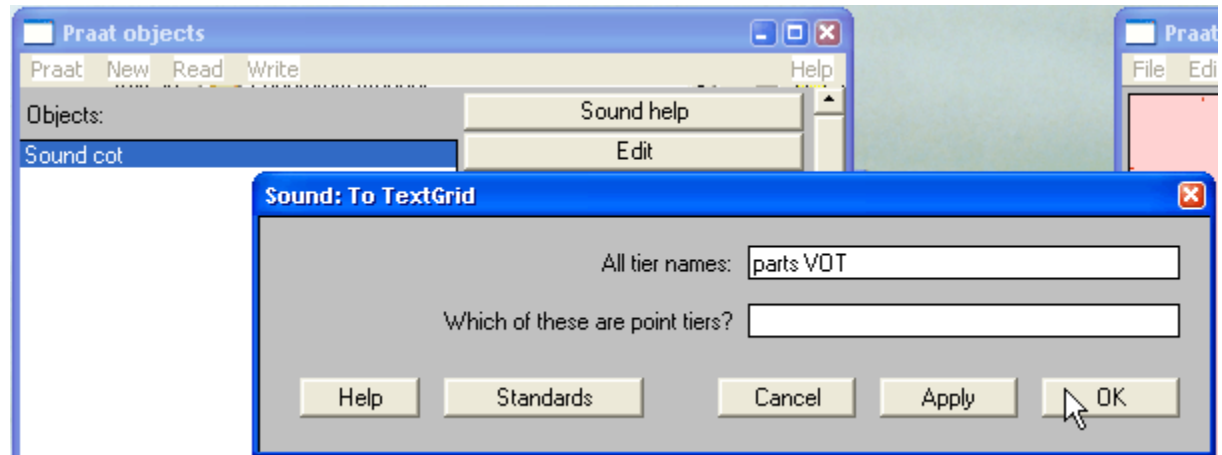
/i/ pea bee tea Dee key gee

/ɒ/ pot bot tot dot cot got

– say the words the way you would normally say them

– make a text grid with two tiers

– select both the sound and TextGrid objects and click on the *Edit* button



– mark the parts of the vowel on one tier
and the VOT on the other tier

– **Tips:**

– bursts might be hard to find in bilabials
– to find exact boundary points, click and
drag to select the relevant portion of
the waveform, then click on the *sel*
button to zoom in to the selected part
of the waveform

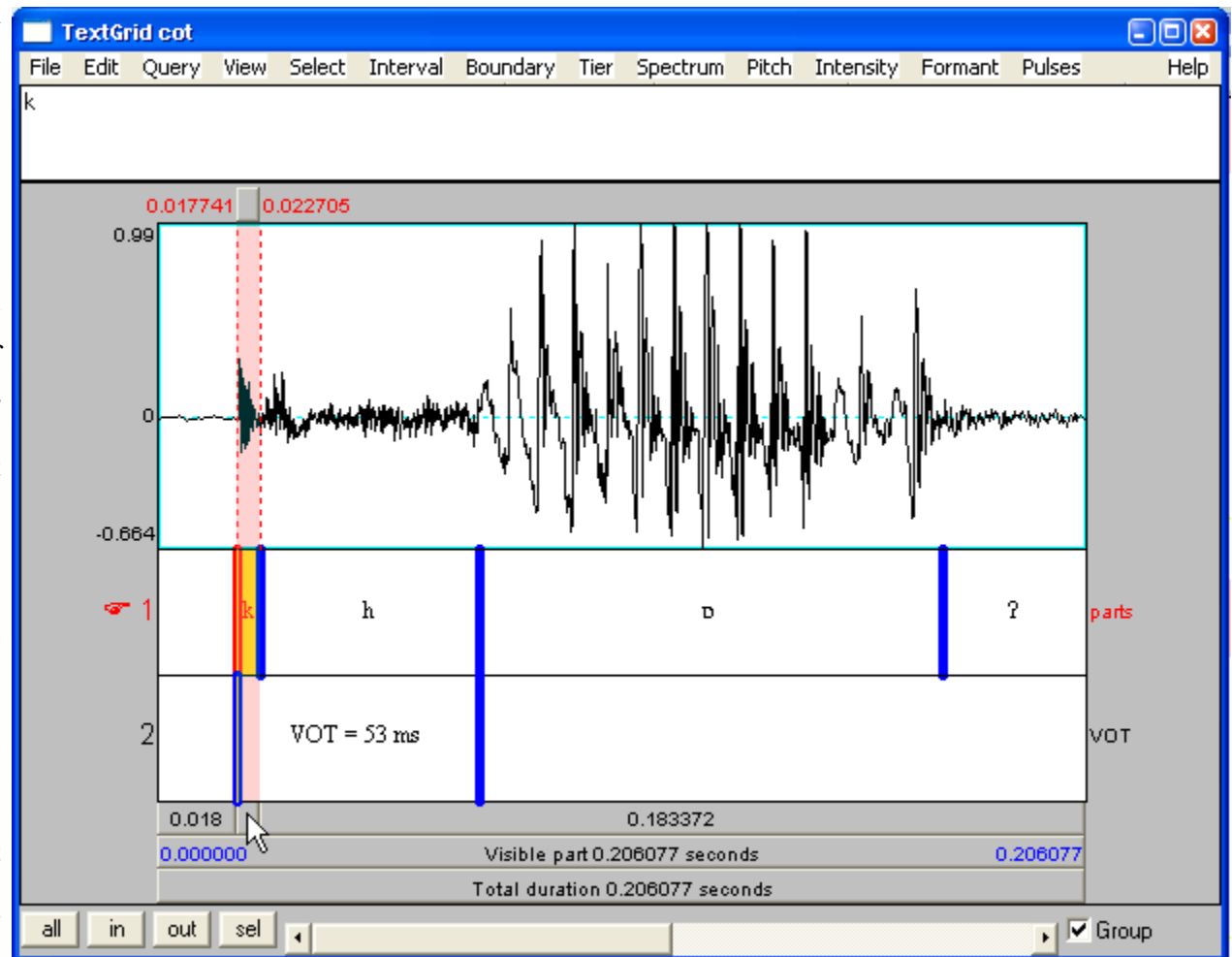
– pressing CTRL + 0 will move the vertical
line position selector to the nearest
zero crossing in the waveform

– to remove a boundary, click on it then
click on *Boundary > Remove*

– to hear a portion of the recording, select
a portion of the waveform and click
on the portion of the bar directly

below the bottom TextGrid tier, use your ears to help you identify the different parts of the word and learn to visually recognise the parts of the waveform

– clicking on a TextGrid interval will select it, the portion of the bar directly below the bottom TextGrid tier will give the duration of the selection in seconds



- you could mark text intervals with phonetic symbols, or written words such as “prevoicing”, “burst”, “aspiration”, and “vowel”
- Example soundfiles and TextGrids for prevoiced, voiceless, and voiceless-aspirated alveolar plosives were shown in class, and can be downloaded from the website.

Things you might want to investigate:

- Do you pronounce word-initial English /b/, /d/, and /g/ with prevoicing or with short-lag VOT?
 - If you produce prevoicing for English /b/, /d/, and /g/, which of the three has the most prevoicing?
 - Can you identify the place of articulation of the plosive by listening only to the prevoicing?
- Which of word-initial English /p/, /t/, and /k/ has the longest aspiration?
- Can you identify the place of articulation of the plosive by listening only to the burst?
- Can you identify the place of articulation of the plosive by listening only to the vowel?
 - If so, is it easier for voiced or voiceless plosives?
- Can you identify the vowel and the place of articulation of the plosive by listening only to the aspiration?

By the end of the course, you should be able to understand the implications of the answers to these questions.

This is an ungraded assignment, don't spend too long working on it, don't feel that you have to measure the VOT lots of recordings or answer all of the questions above.

Graded Assignment 1 - Voice Offset Time

Course Section 4 Maximum grade 10 points

Part A

Experiment

- Produce a vowel and then a voiced bilabial plosive but do not release the plosive.
- Try to maintain voicing for as long as possible.
- You should find that you can only maintain voicing for a short period of time compared to the length of time you can maintain voicing during a vowel.

Question

- Why can you only maintain voicing for a short period of time during a stop closure?
 - You will have to write about the air pressure conditions necessary for voicing.

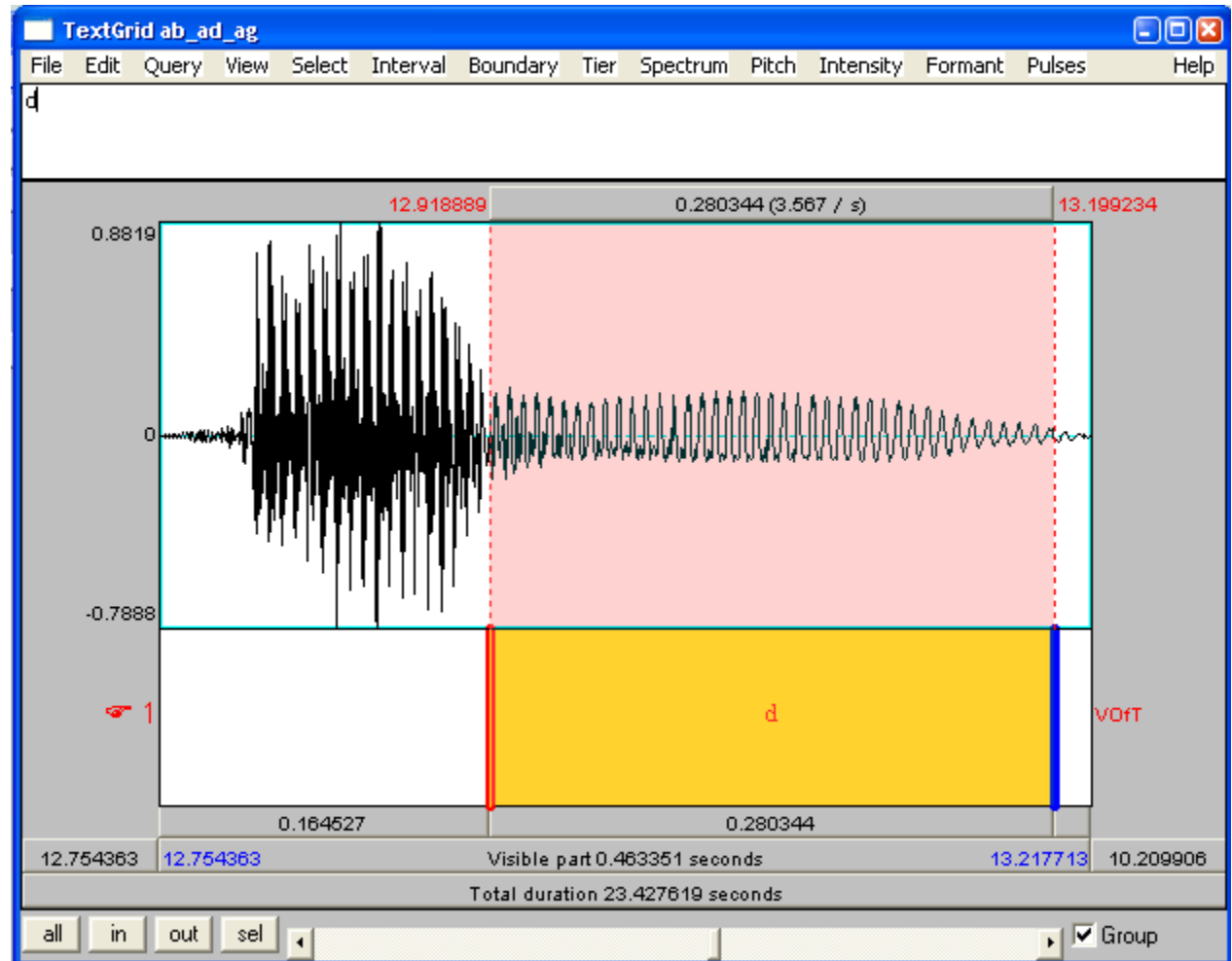
(my answer has 280 words)

Part B

Experiment

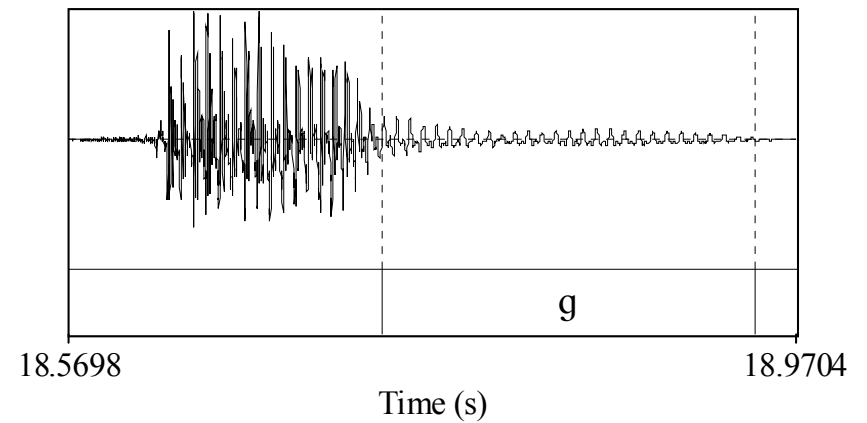
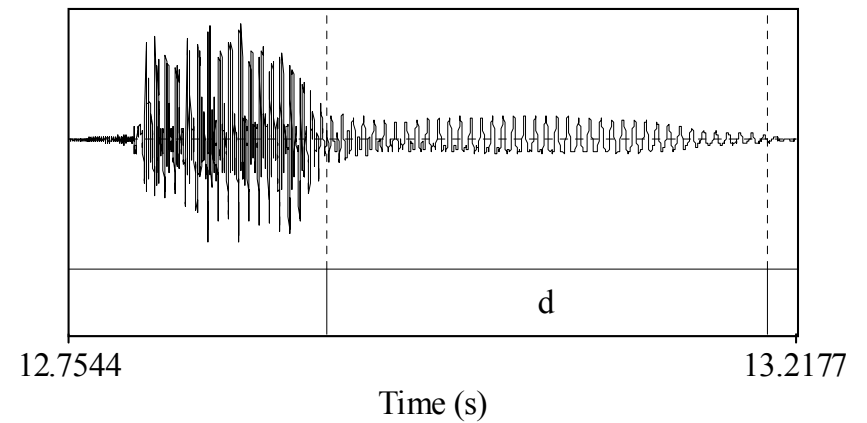
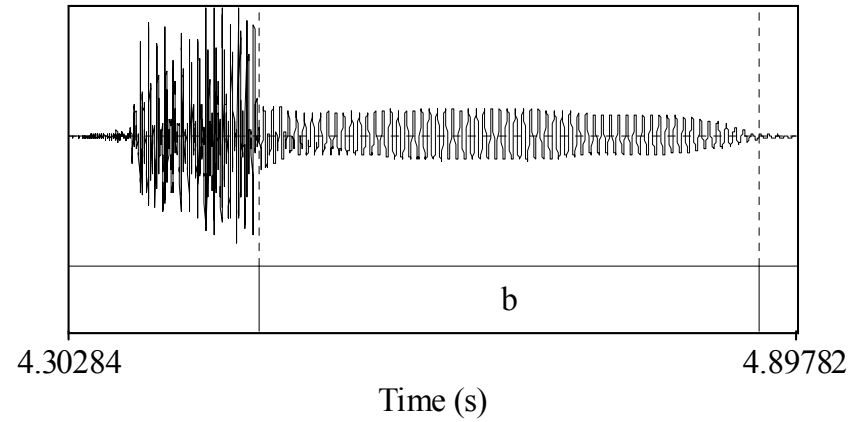
- Say a vowel then a bilabial plosive and try to maintain voicing for as long as possible during the closure. Time how long you can continue voicing for after the beginning of the closure. Repeat several times to get a range of values. Repeat for an alveolar stop and a velar stop.
- You could do this with a stopwatch, but you will get more accurate measurements if you record yourself and measure using *Praat* or some other piece of software. Below is an illustration of the measurements I made using Praat.

- I recorded myself saying [ab], [ad], and [ag] four times each, and saved the recording as a single file
- I loaded the file into Praat
- I made a TextGrid
- I marked the voicebar following each vowel
- I saved a copy of the TextGrid
- I selected the interval of each voicebar that I had marked, read off the duration, and wrote it in a table
 - Hint: zoom in to the word, select the interval for the voicebar, ALT + → will move to the next interval



Sample measurements:

| VC sequence | Duration (ms) | | | | |
|-------------|---------------|----------|----------|----------|------|
| | sample 1 | sample 2 | sample 3 | sample 4 | mean |
| [ab] | 388 | 375 | 410 | 375 | 387 |
| [ad] | 213 | 299 | 280 | 256 | 262 |
| [ag] | 259 | 205 | 190 | 270 | 231 |



– In your answer include:

- three pictures, one example for each place of articulation
- a table of measurements

Questions

- On the basis of my data above (and your data if it agrees), what is the relationship between place of articulation and duration of voicing?
- What might explain this? Do your best to come up with a theory.
 - Think about air pressure and volume.
- What do you do differently when you try to maintain voicing for as long as possible compared to when you let voicing end naturally?
 - Think about the bilabial case in particular and look in a mirror.

(my answer has 218 words)

- Please type your answer.
- Leave enough space for me to write in comments.
- Put your name at the top of every page and staple the pages together.
- Print only on one side of each page.
- Do not add an extra cover page.
- Use an IPA font. [g] is not an IPA symbol! (as a last resort write in the symbols by hand on a printout)
- Avoid writing about material which is irrelevant to the question, e.g., don't write about plosive releases.

Submit as hardcopy in class or as a pdf in the WebCT dropbox.

Noise

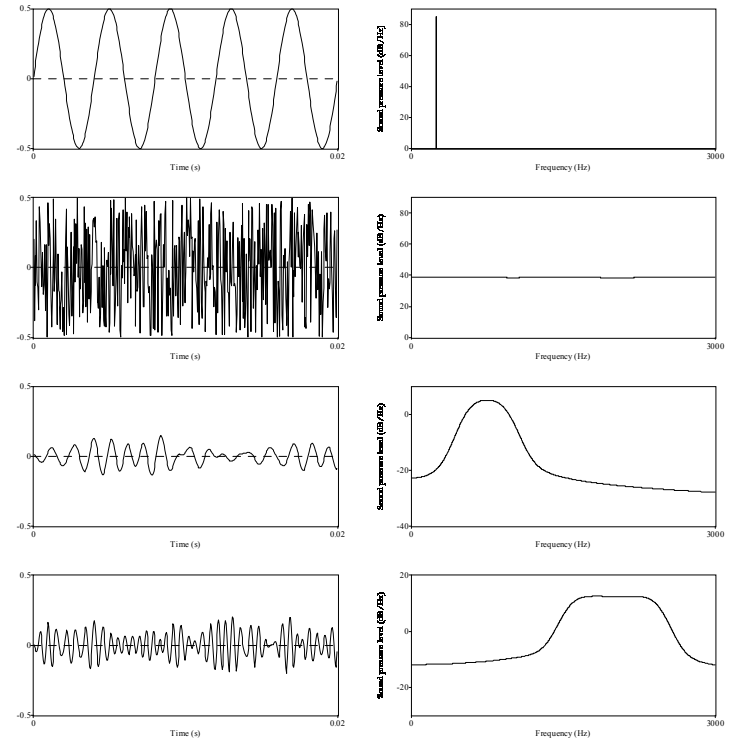
Download *noise.PraatScript* and *noise.prapic* Run the script.

The script generates the sounds which were demonstrated in class.

- listen to the sounds
- examine the sounds and the non-smoothed and smoothed spectra in the Edit window

Periodic sound

- In the time-domain representation (waveform), the sine wave repeats 5 times in 0.02 seconds
 - it has a period of (repeats once every) $0.02 \text{ s} / 5 = 0.005 \text{ s} = 5 \text{ ms}$
 - it has a frequency of $1 / 0.005 \text{ s} = 250 \text{ Hz}$
- In the frequency-domain representation (spectrum), there is a single spike at 250 Hz



White Noise

- In the time-domain representation, the amplitude of the sound jumps up and down randomly.
- In the smoothed frequency-domain representation, the amplitude is equal across all frequencies.

Band-Filtered Noise

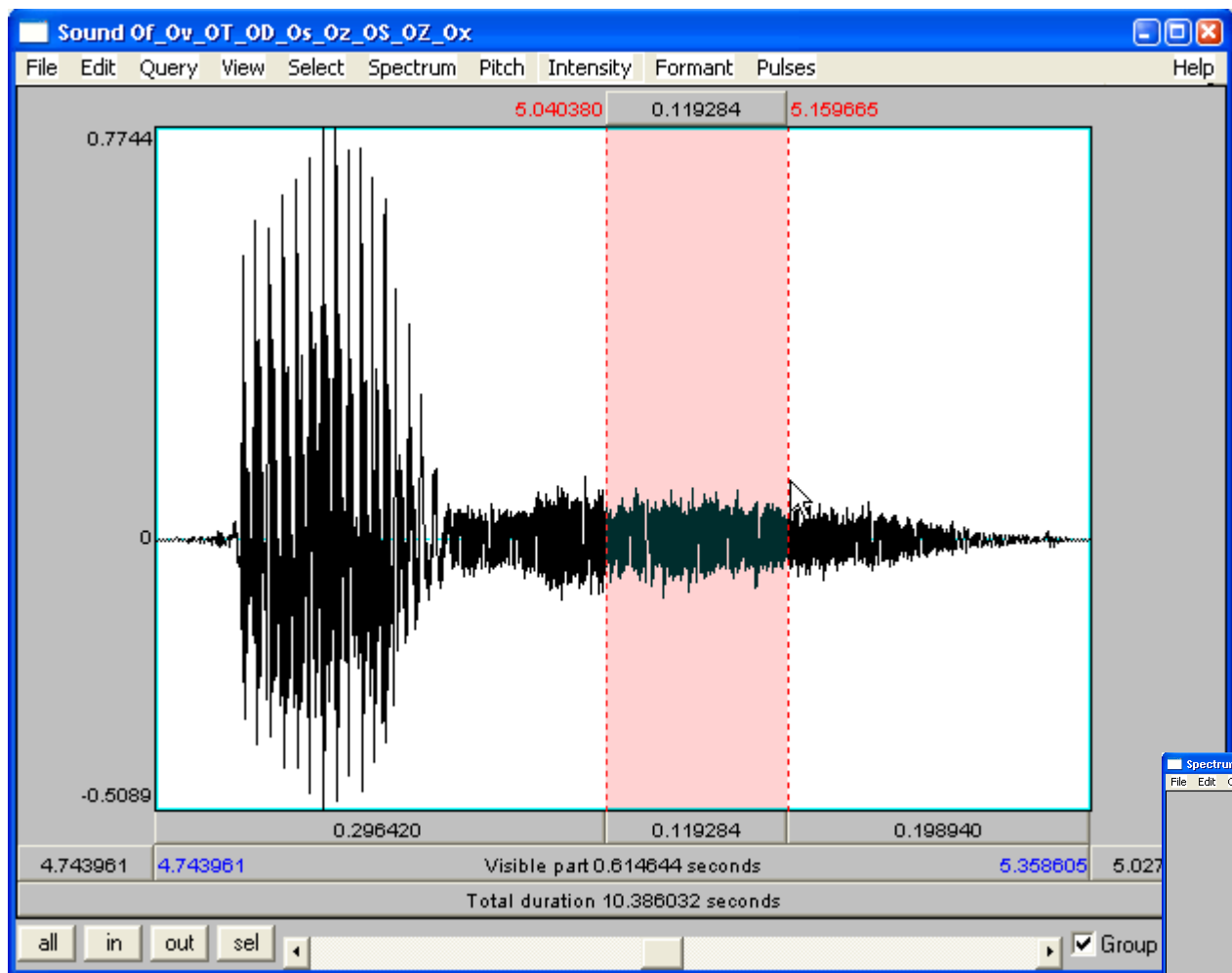
- Only a certain range of frequencies are included in the sound.
- The first band-filtered noise sounds like it has a low pitch because it has high amplitude for frequencies in the range 500–1000 Hz and low amplitude for all other frequencies.
- The second band-filtered noise sounds like it has higher pitch because it has high amplitude for frequencies in the range 1500–2500 Hz and low amplitude for all other frequencies.

Graded Assignment 2 - Fricative Spectra

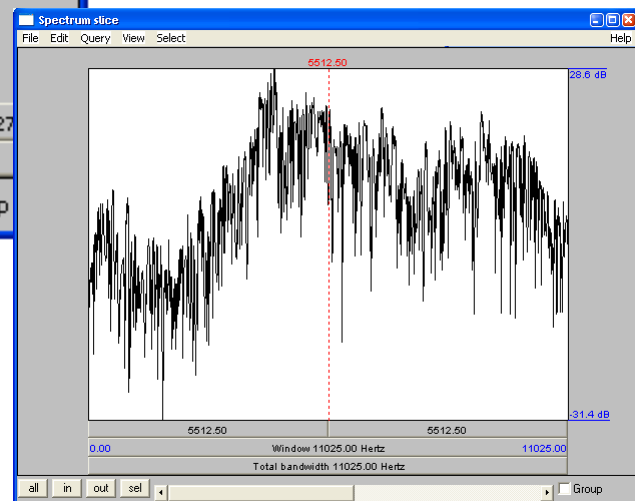
Course Section 5 Maximum grade 10 points

Experiment:

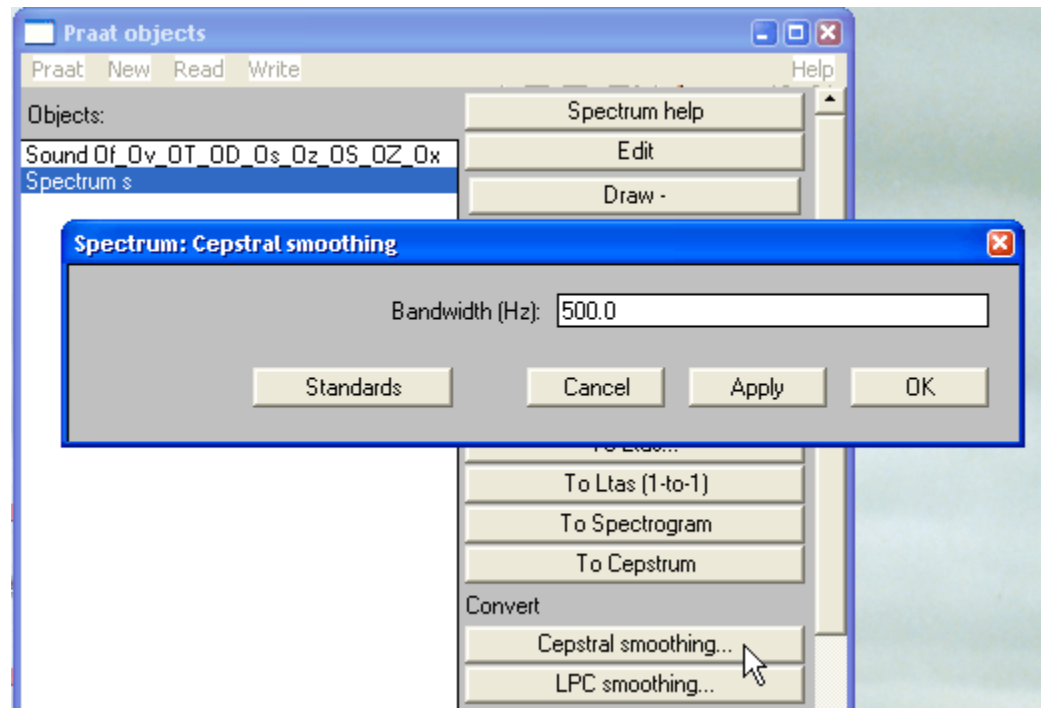
- Record yourself making the English voiceless and voiced fricatives [f, v, θ, ð, s, z, ʃ, ʒ], or a set of fricatives from another language that you speak (please indicate the language in your answer).
 - You may find it easier to say the fricatives if you produce a vowel before or after or both, but always make the same vowel, and make sure that your measurements are taken only from the central part of the fricative. In the examples below, I said [ɒf, ɒv, ɒθ, ɒð, ɒs, ɒz, ɒʃ, ɒʒ] and an additional vowel from Scottish English [ɒx].
 - If you record all the fricatives to a single file, you may wish to adjust the volume of the whole file, but do not adjust the volume of each individual fricative otherwise you will not be able to compare the relative amplitudes of different fricatives.
- Use Praat to make a smoothed spectrum for each fricative



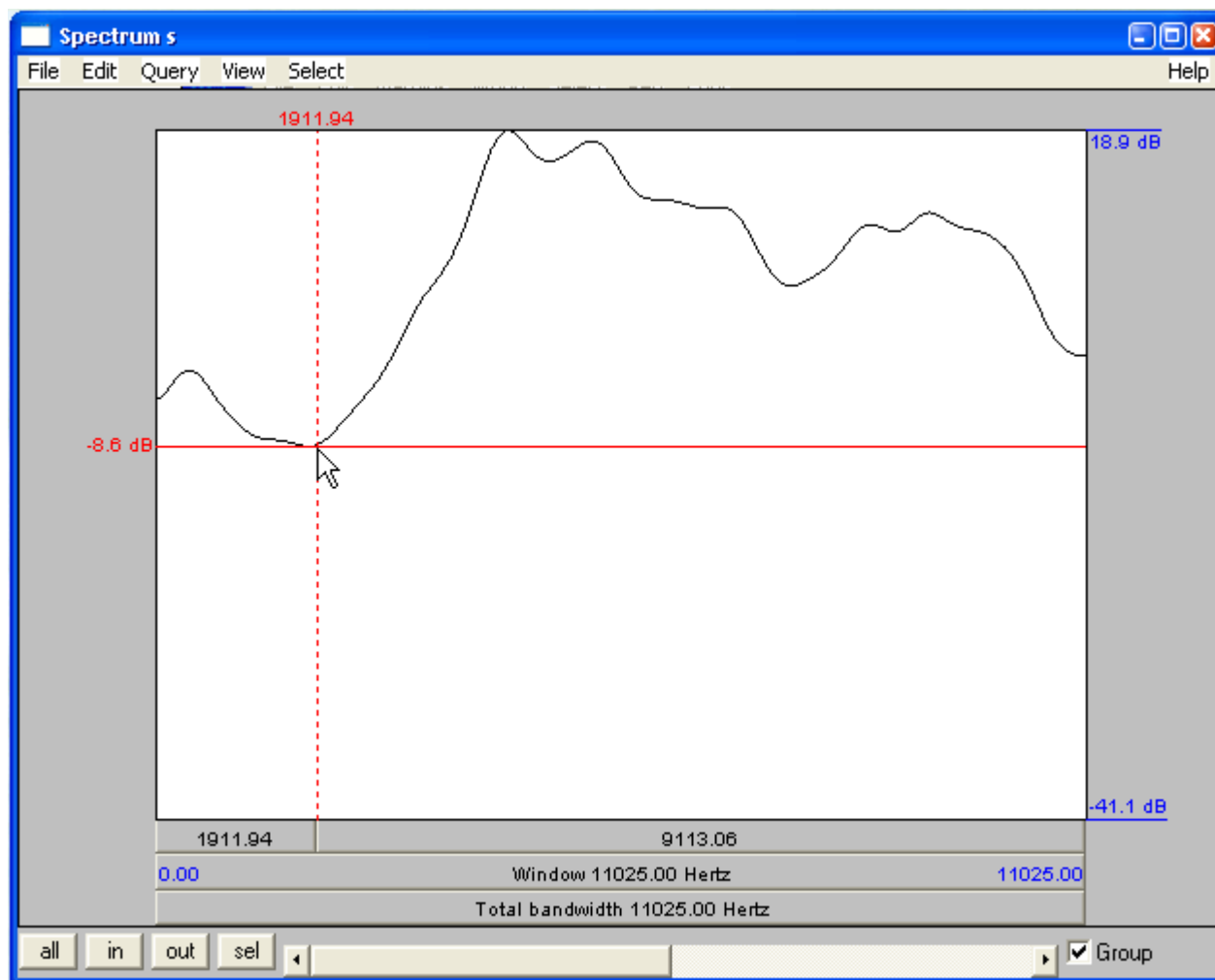
- In the Praat Edit window, for each fricative
- select approximately the middle half of the fricative (click and drag)
- click on *Spectrum > View spectral slice*
- a window opens showing a non-smoothed spectrum (Discrete Fourier Transform)
- close the spectrum window



- In the Praat objects window:
- rename the *Sound slice* object, e.g. *f*, *v*, *th*, *dh*, *s*, *z*, *sh*, *zh* for [f, v, θ, ð, s, z, ʃ, ʒ] respectively
- click on *Cepstral smoothing...*
 - if you increase the bandwidth, you will get more smoothing, try experimenting, but use the same bandwidth for all the fricatives you measure, I'm going to use 1000 Hz
- you will get a new Spectrum object with the same name as the original, you may wish to rename it, e.g. *s smoothed*
- Hint: to save time, you can make all the non-smoothed spectra, select all the Spectrum objects, and then click on *Cepstral smoothing...*



(a cepstrum is a Fourier transform of a frequency-domain spectrum, which is a Fourier transform of a time-domain waveform)



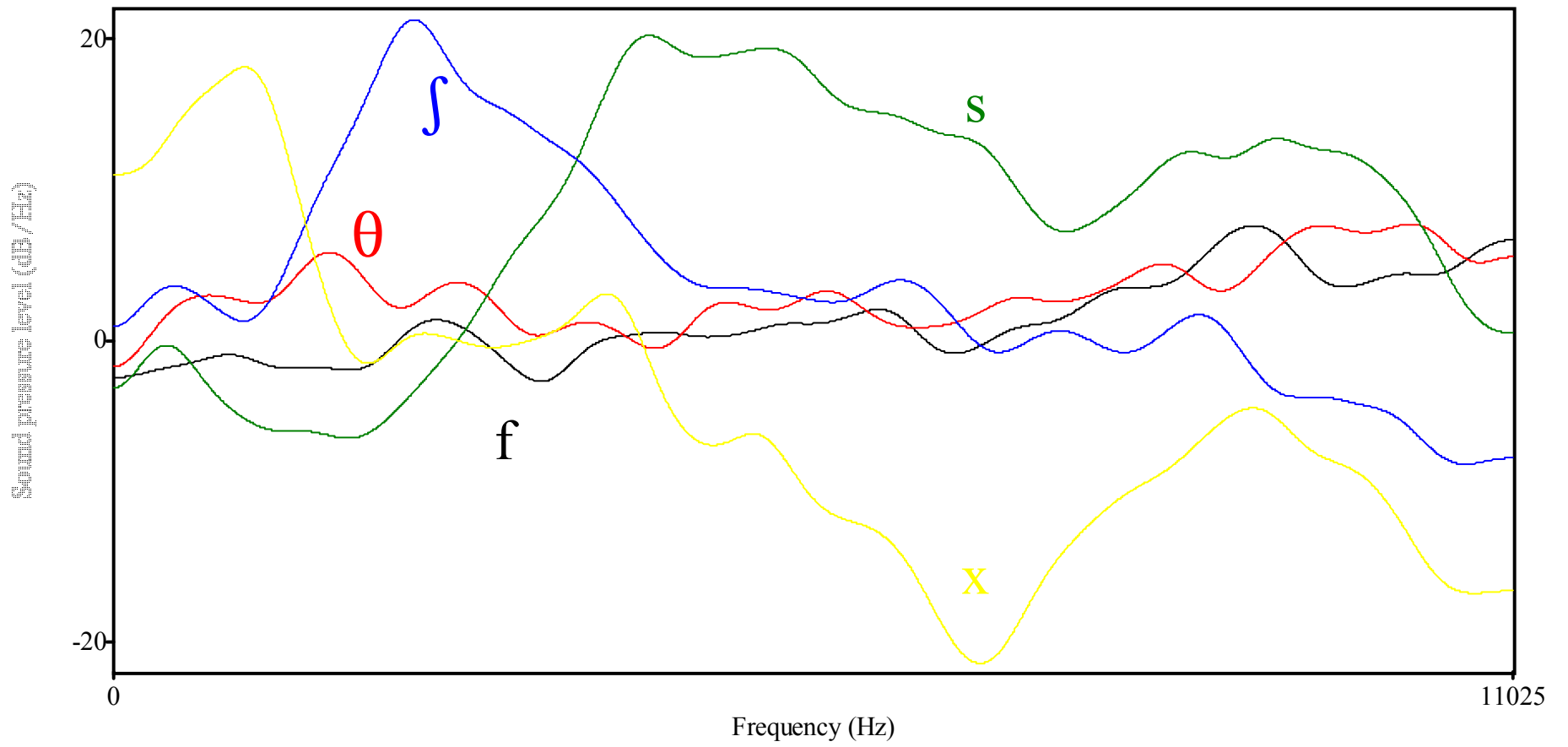
– if you open the smoothed spectrum in an Edit window, you can click to get the frequency and amplitude of peaks and troughs in the spectrum

–In the Praat object window, select each smoothed spectrum one-at-a-time, and draw it in the Praat picture window

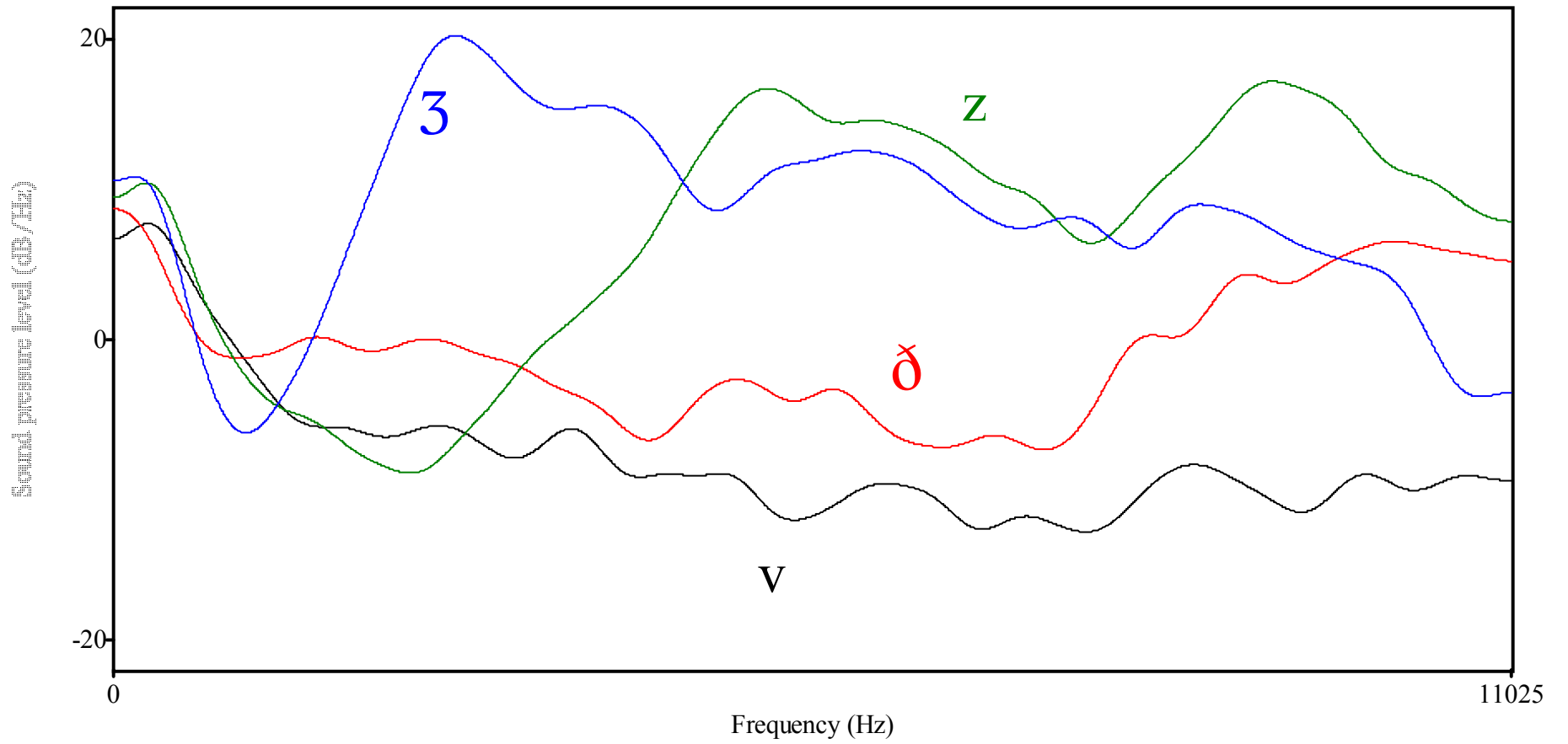
– set the minimum and maximum power (to find a good range for all spectra, select all the spectrum objects and draw them together, adjusting the range and redrawing until you find a good range)

- draw the spectra for all the voiceless fricatives on top of each other
- draw the spectra for all the voiced fricatives on top of each other
- only check *Garnish* for the first picture
- you may want to change colours and add labels to identify each fricative (add labels using Praat or hand draw labels on a printout)

English voiceless fricatives



English voiced fricatives



- Include the pictures of your voiceless and voiced fricative spectra with your answer to the assignment.

Questions:

Answer the following questions on the basis of the sample smoothed spectra that I provided above (you can ignore [x]). Sample sound files and smoothed spectra objects are also available from the course website. Listening to the sound files and looking at the details in the spectra objects may help when answering the following questions.

- What are the acoustic differences and similarities between the fricatives, and how do these differences and similarities correlate with the fricatives' perceptual properties (pitch and loudness), and with their articulation and the way in which turbulent flow is produced?

- Start with [f] and [θ]

- What are the similarities / differences between the spectra?
- What are the perceptual similarities / differences?
- What are the articulatory similarities / differences?
- Relate perception and pronunciation to the shape of the spectra.

- Next look at [s] and [ʃ]

- How are they similar or different to [f] and [θ]?
- How are they similar or different to each other?
 - same set of questions as above

- Next compare the spectra of the voiced fricatives with their voiceless counterparts

- What is the common difference between all the voiced and all the voiceless spectra?
- Relate perception and pronunciation to the shape of the spectra.

- My answer has 312 words

Graded Assignment 3

- Synthesis of Vowel-like Sounds Using Sine Waves

Course Section 7

Maximum grade 5 points

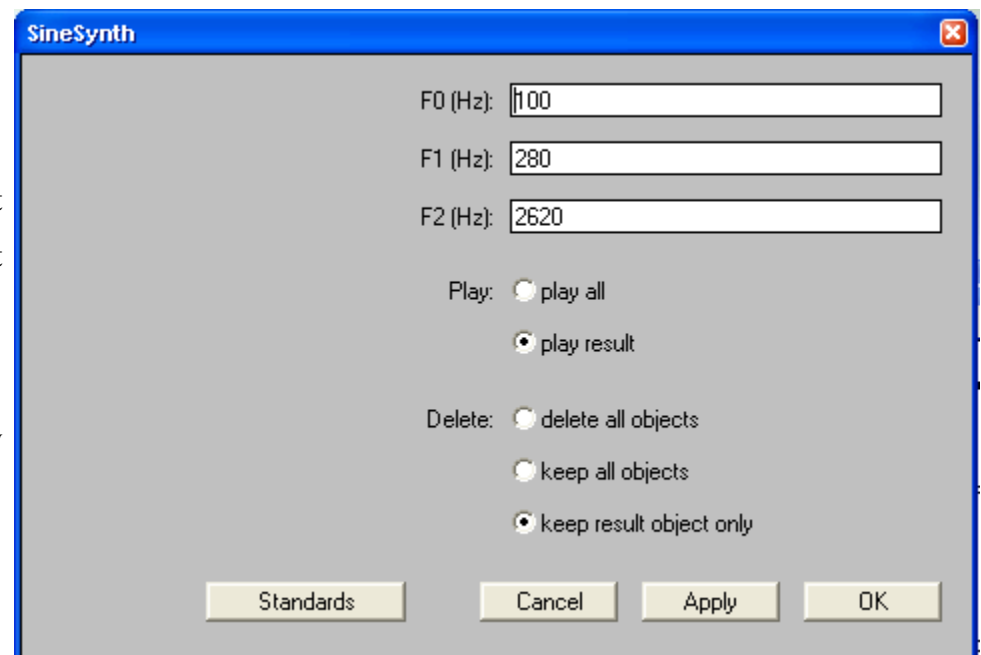
Complete this assignment before we cover Section 7 in class.

This assignment is designed to give you some hands-on experience with synthesising vowel-like sounds. The intention is to give you some practical experience so that you will have a feeling for what we are talking about when we cover the related theory in class. Performing this exercise conscientiously will make it easier to understand the theory when we cover it in class. Please read the instructions carefully and take your time synthesising the vowel-like sounds. Play with the synthesiser until you think you have a good feel for the relationship between F1, F2, and the resulting vowel-like sound.

– Download and run the script: *SineSynth.PraatScript*

– The default values (Fundamental Frequency) F0 = 100, (First Formant) F1 = 280, and (Second Formant) F2 = 2620 should result in a sound that sounds something like /i/, the vowel in ‘heed’.

These are not really formants, we will cover what formants really are later in class.



– The first three sine waves drawn correspond to the Fundamental Frequency, First Formant, and Second Formant.

– The fourth wave drawn is the result of adding the three sine waves together:

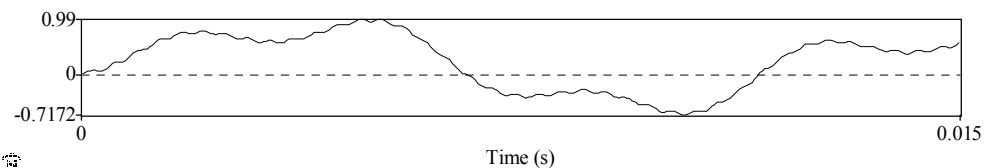
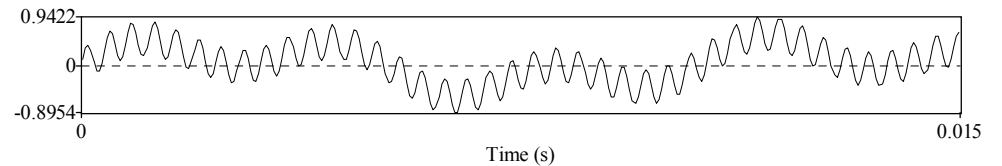
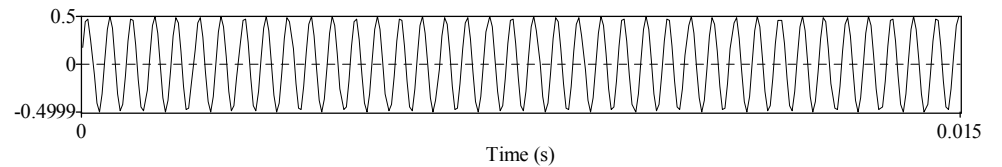
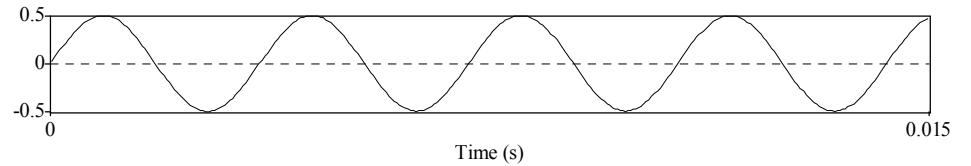
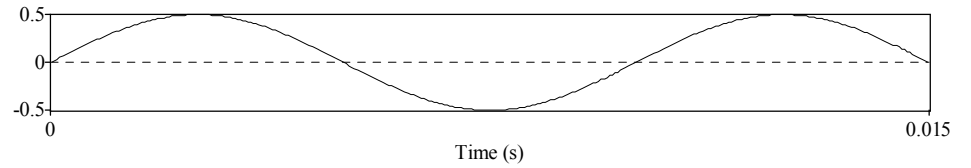
– the amplitude (height on the graph) at each point in time in the Combined Frequencies is the sum of the amplitudes of the three sine waves at the same point in time. All the component sine waves added had equal amplitude.

– The fifth wave drawn is the result of adding the three sine waves together, but each sine wave has a different amplitude:

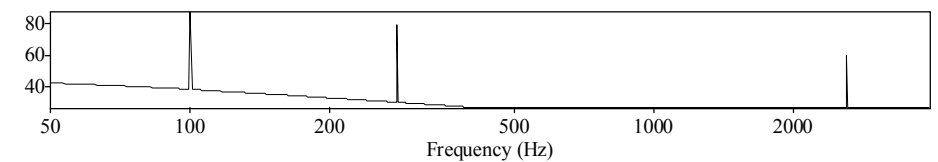
– the lowest frequency sine wave has the highest amplitude and the highest frequency sine wave has the lowest amplitude

– The sixth graph is a frequency-domain representation of the fifth wave (all the previous graphs are time-domain representations).

– Note that there are spikes at the frequency of each of the component sine waves, and the amplitudes of the spikes is such that the higher the frequency the lower the amplitude (this is called *spectral tilt*). The spectral tilt is -6dB per octave (the amplitude decreases by 6dB every time the frequency doubles).



Amplitude Spectral Density (dB/Hz)



– Run the script again and try the following formant frequencies:

F1 = 300 F2 = 870

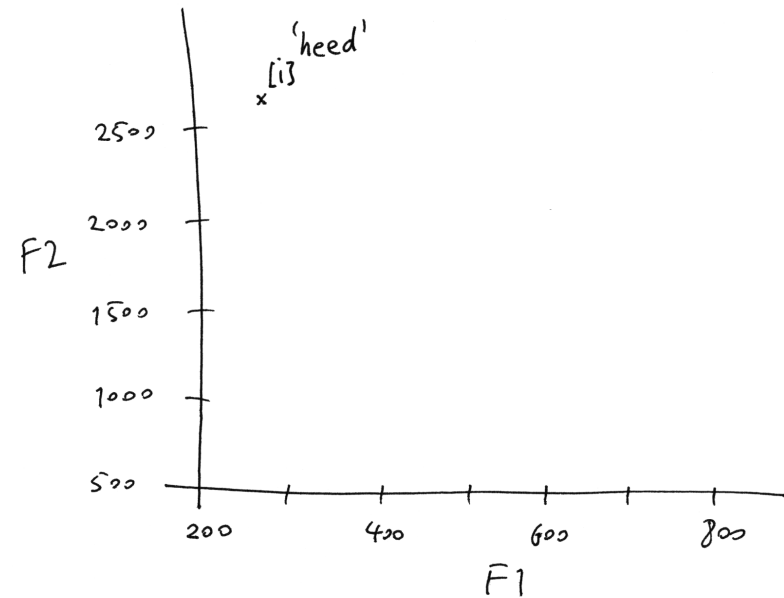
F1 = 730 F2 = 1090

– Which vowels do these sound like?

– Try different frequencies and see if you can make other vowels such as [æ] ‘hat’, [ɛ] ‘head’, [ɪ] ‘hid’, and [o] ‘hope’.

– if you are having difficulty, try some of the number from Rogers (2000) p 154

– Draw a graph with F1 on one axis and F2 on the other axis, and plot the points corresponding to each vowel that you have synthesised.



Please use the answer sheet on the next page and draw the graph on the back.

- Important point to remember: You can create a complex waveform such as a vowel-like sound by adding simple sine waves together.
- Reverse the process: Fourier analysis a technique which takes a complex waveform such as a vowel and determines which sine waves would have to be added together in order to make that sound. These are the frequency components of the sound.

Alternative software: *Esynth* by Mark Huckvale at University College London <http://www.phon.ucl.ac.uk/resource/software.html>

Assignment 3 (Section 7) - Synthesising Vowel-like Sounds Using Sine Waves

– Synthesise vowel-like sounds using the following F1 and F2. Which vowels do they sound like? (Vowel symbols are not essential, a sample English word will do.)

F1 = 280 F2 = 2620 [i] ‘heat’

F1 = 300 F2 = 870 _____

F1 = 730 F2 = 1090 _____

– What F1 and F2 correspond to the following vowels?

[æ] ‘hat’ F1 = _____ F2= _____

[ɛ] ‘head’ F1 = _____ F2= _____

[ɪ] ‘hid’ F1 = _____ F2= _____

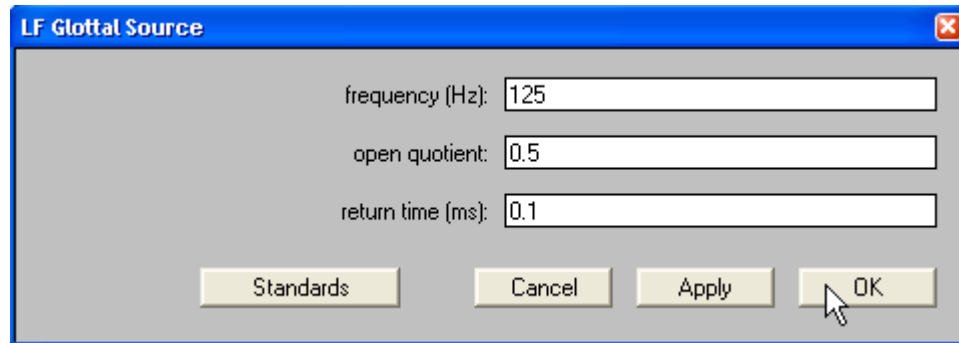
[o] ‘hope’ F1 = _____ F2= _____

– On the reverse side of this piece of paper, draw a graph with F1 on one axis and F2 on the other axis, and plot the points corresponding to each vowel-like sound that you have synthesised

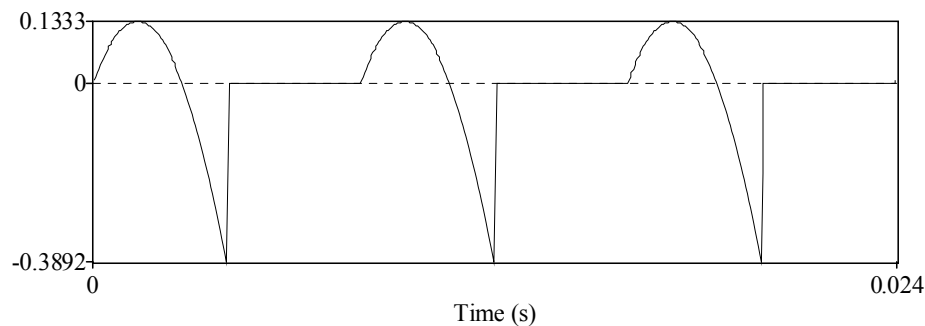
Source and Filter Synthesiser

Download the scripts: [DiphthSynth.PraatScript](#) [FilterSynth.PraatScript](#) [FilterSynth_no_pause.PraatScript](#) [SourceSynth.PraatScript](#)

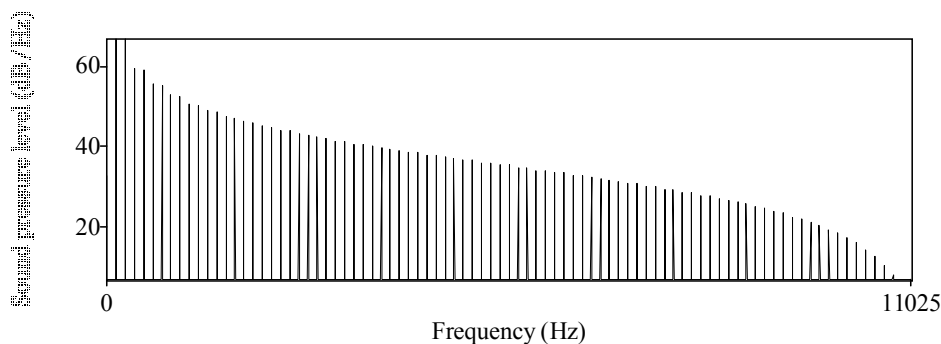
Open and Run *SourceSynth.PraatScript*



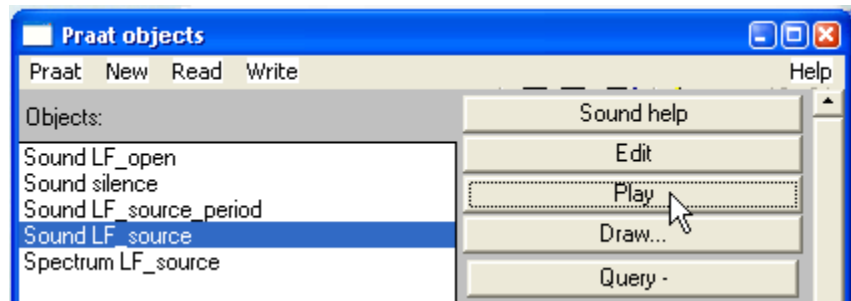
In the Praat picture window you will see a time-domain and frequency-domain representation of the synthesised glottal source (volume acceleration). Note the similarity in shape to the glottal source shown on page 22 of the Lecture Notes.



- the frequency in Hertz is the number of times the waveform repeats every second
- the open quotient is the proportion of each period during which the vocal folds are open, the time during which the waveform gradually rises and falls
- the return time is the time to go from the bottom of the spike in the waveform back to zero



- in the spectrum, each line represents a *harmonic* frequency component, the first harmonic is at the *fundamental frequency* (f_0), and higher harmonics are equally spaced at intervals of f_0
- the amplitude of the harmonics decrease as frequency increases (spectral tilt)

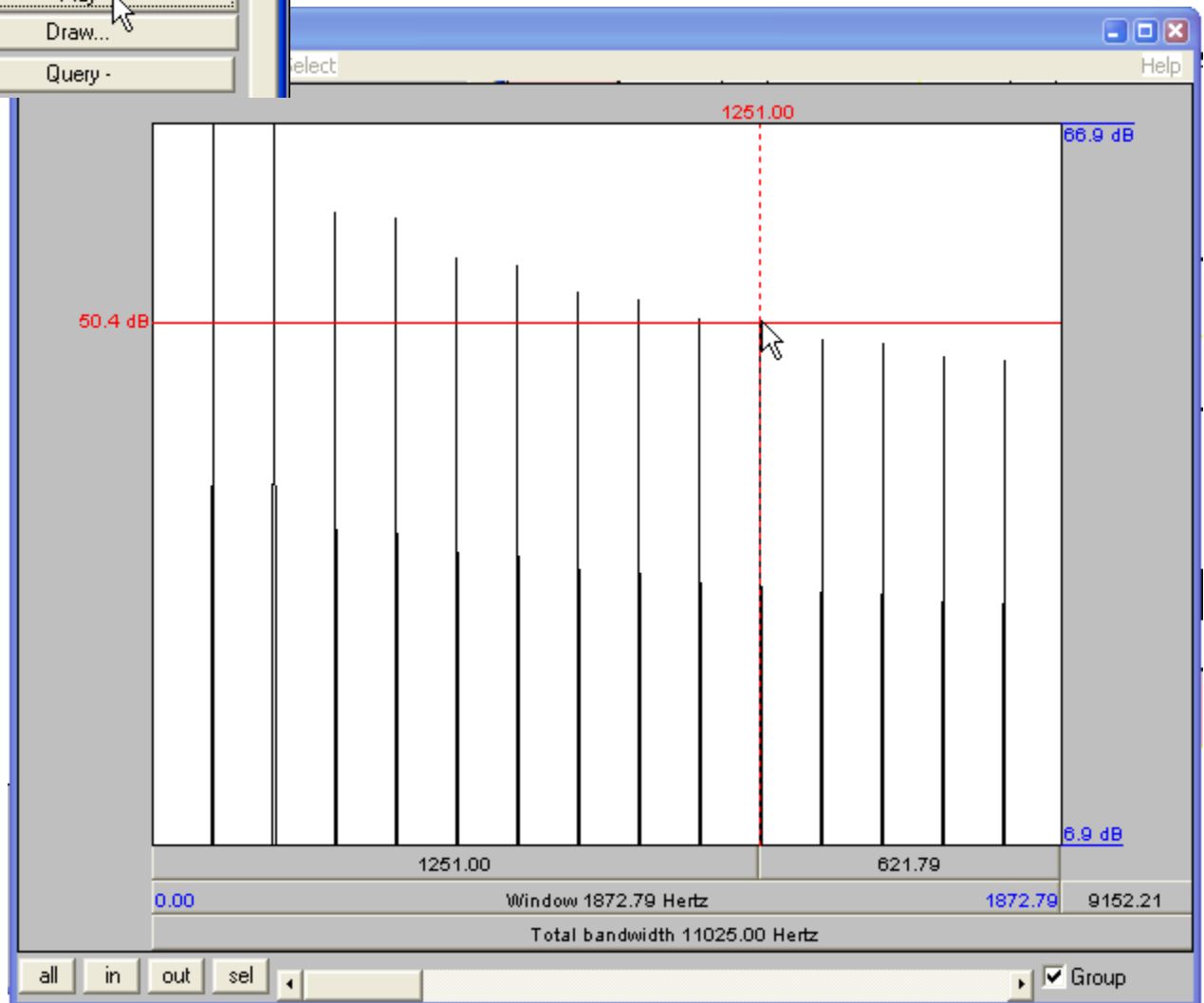


– select the *LF_source* Sound object and listen to it

– open the *LF_source* Sound and Spectrum objects to investigate details, e.g., the tenth harmonic has time times the fundamental frequency

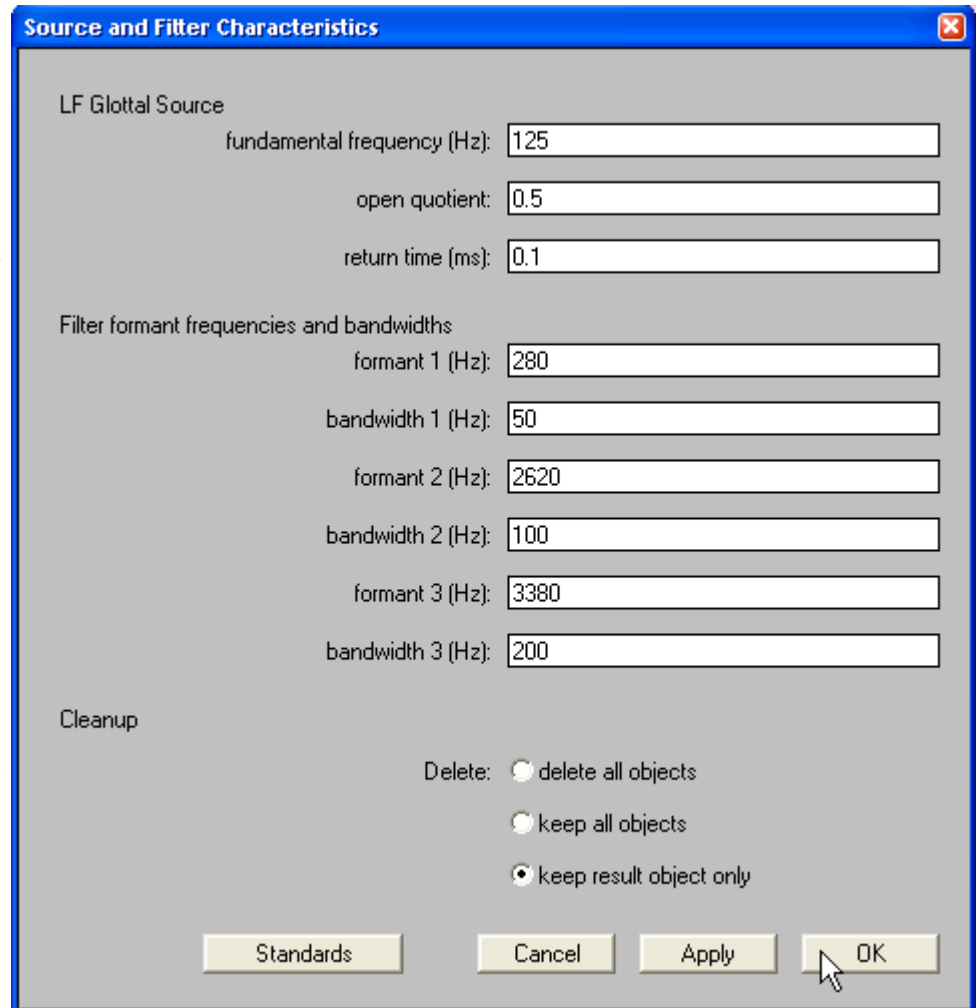
– run the script again and experiment with different frequencies, open quotients, and return times

This glottal source model is described in detail in Stevens (1998) p97–100



Open and Run *FilterSynth.PraatScript*

- select glottal source properties as before
- choose formant frequencies appropriate for the vowel you want to synthesise, you can leave the bandwidths at their default values
- click OK
- you can examine the resulting sound and spectrum objects by opening them in the edit window
- several panels will appear in the Pratt picture window, see next page for an explanation



[Alternative software: *Esystem* by Mark Huckvale at University College London <http://www.phon.ucl.ac.uk/resource/software.html>]

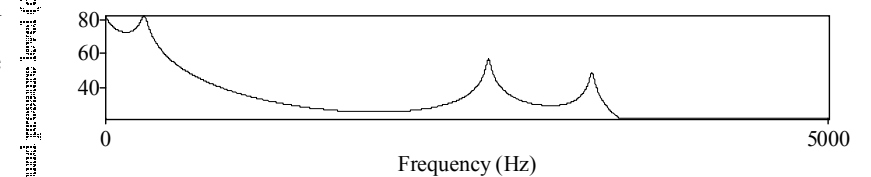
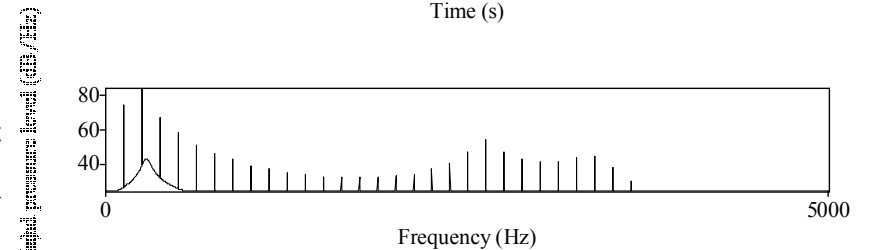
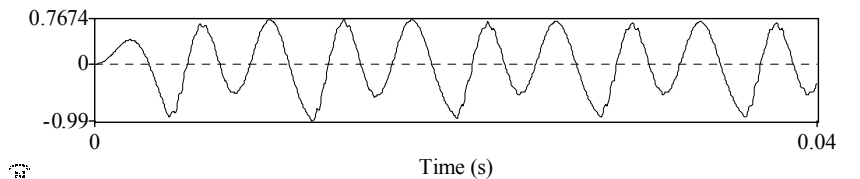
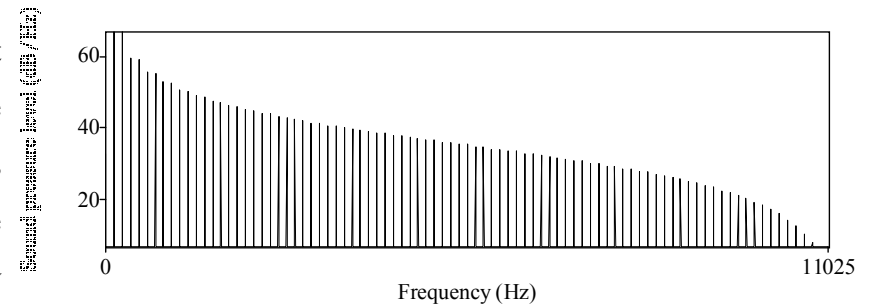
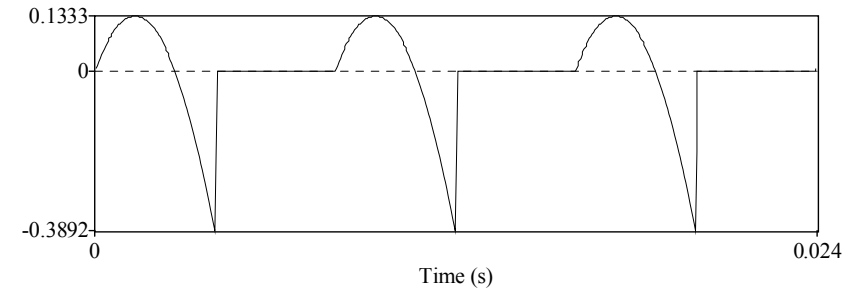
In the Praat picture window, you will see the same time and frequency-domain representations of the glottal source as above plus:

- Bottom panel: A frequency-domain representation of the *filter* simulating the resonance properties of the supralaryngeal vocal tract.

- Second-from-bottom panel: A frequency-domain representation of the result of passing the *glottal source* through the *filter*. The frequency spacing of the *harmonics* is unchanged, but their amplitudes have been changed. Harmonics near *formants* (the resonant frequencies of the supralaryngeal vocal tract) have relatively high amplitudes, and harmonics far from formants have relatively low amplitudes.

- Third-from-bottom panel: A time-domain representation of the result of passing the *glottal source* through the *filter*.

The spectral tilt of the glottal source is usually assumed to be -12 dB/oct., but depends on factors such as the open quotient. Radiation at the lips results in a spectral tilt of $+6$ dB/oct. (not drawn here but see p50 of the *Lecture Notes*), resulting in an overall spectral tilt of -6 dB/oct. for the vowel. The bottom panel actually has the -6 dB/oct. spectral tilt of the vowel rather than the 0 dB/oct. of the filter.



Open and Run *DiphthSynth.PraatScript*

- select glottal source properties as before
- choose formant frequencies appropriate for the diphthong you want to synthesise, the first value is the frequency of the formant at the beginning of the vowel and the second value is the frequency of the formant at the end of the vowel
- the result includes a *spectrogram* showing the change in formant values across time

Source and Filter Characteristics

LF Glottal Source

fundamental frequency (Hz):

open quotient:

return time (ms):

Filter formant start and end frequencies and bandwidths

formant 1 (Hz):

bandwidth 1 (Hz):

formant 2 (Hz):

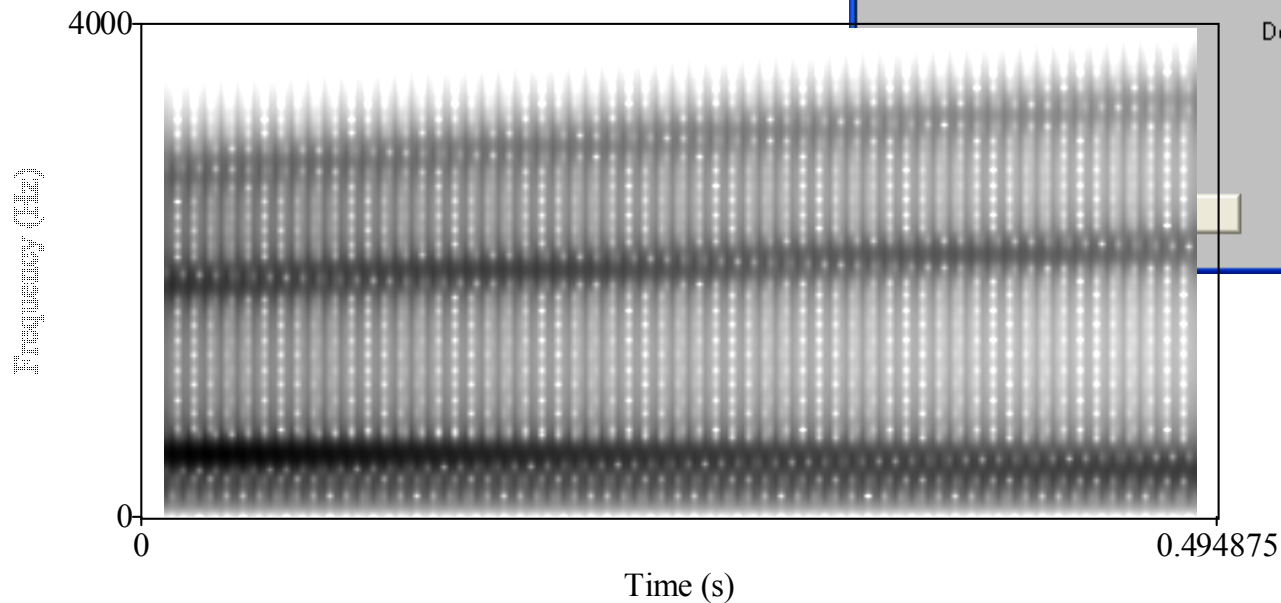
bandwidth 2 (Hz):

formant 3 (Hz):

bandwidth 3 (Hz):

Cleanup

Delete: delete all objects
 keep all objects
 keep result object only



Graded Assignment 4 - Measuring Vowel Formants

Course Section 8 Maximum grade 10 points

- Record yourself saying at least 5 vowels
 - these can be vowels from any language you speak
 - include at least one diphthong in your set of vowels
 - say the vowels either in isolation or in words, but if in words make sure the preceding and following consonants are always the same, e.g. English [h_t] *heat, hit, head* etc.

 - Measure the duration, and the first and second formant frequencies at 25% and 75% of the duration of each vowel
 - detailed instructions below
 - if your language has front rounded vowels you might also need to measure F3 to help distinguish them from non-rounded front vowels

 - Plot F1 and F2 for each vowel
 - put F1 on one axis and F2 on the other axis
 - draw each vowel as an arrow with the tail at the F1 & F2 you measured at 25% of the duration the vowel and the head at the F1 & F2 you measured at 75% of the duration the vowel

 - label each vowel with its IPA symbol

 - Draw a bar graph of the vowel durations.

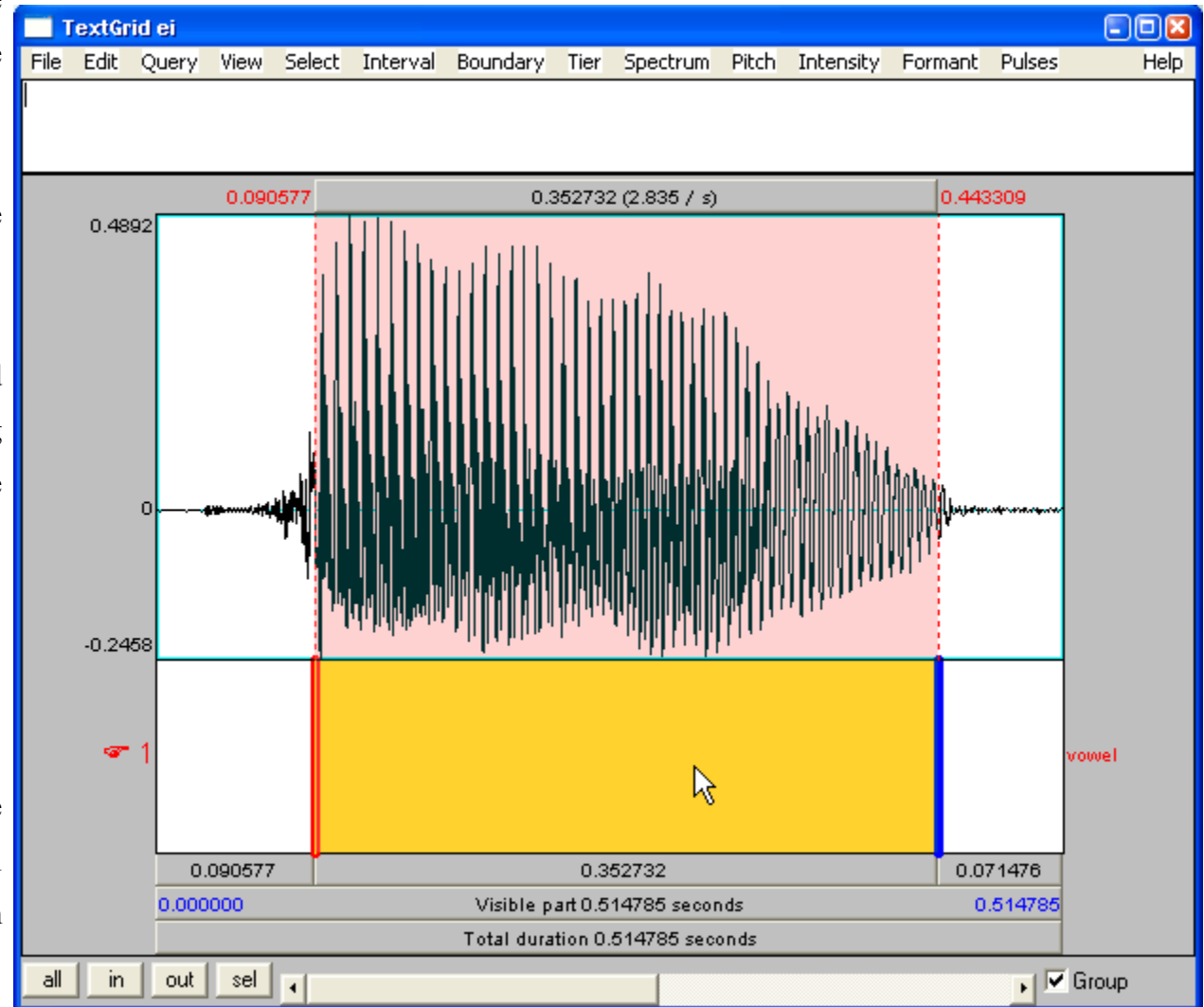
 - Submit the following for grading:
 - a table of duration, and first and second formant values for each vowel
 - indicate each vowel using an IPA symbol and an example word, and state the language spoken
 - the F1–F2 vowel plot
 - the bar graph of vowel durations
 - a picture of the beginning and end spectra for each vowel
 - label each set of spectra with an IPA symbol
 - indicate the F1 and F2 peak in the smoothed spectra (draw an arrow to or a circle around the relevant peaks)
- See *Lecture Notes* p 52–53, 63 for examples of F1–F2 plot, bar graph, and spectra

Detailed instructions for making measurements:

There are several ways to get formant measurements using Praat, the following method should be reasonably transparent in light of our discussion of formants in class, it should also be reasonably robust. Feel free to play with Praat or other software (e.g., *WaveSurfer*) and explore other ways of measuring formants, but please complete the graded assignment according to the instructions below.

- put each vowel recording in a separate sound file and save the files
- make a TextGrid for each recording and place one boundary mark at the beginning of the vowel and one at the end of the vowel
- save the TextGrid

[Alternative software: *WaveSurfer* by Kåre Sjölander & Jonas Beskow, at the Royal Institute of Technology (KTH) Stockholm
<http://www.speech.kth.se/wavesurfer/>]



Open and Run *formant_masurer.PraatScript*

- the script will use the first sound and TextGrid object in the Praat object window
 - it will complain if the TextGrid doesn't have two boundaries
 - you will get silly results if the boundaries in the TextGrid don't match the beginning and end of the vowel
 - it will delete any other objects in the list so make sure you save them before running the script

- the script measures the time between the first and second boundary in the TextGrid

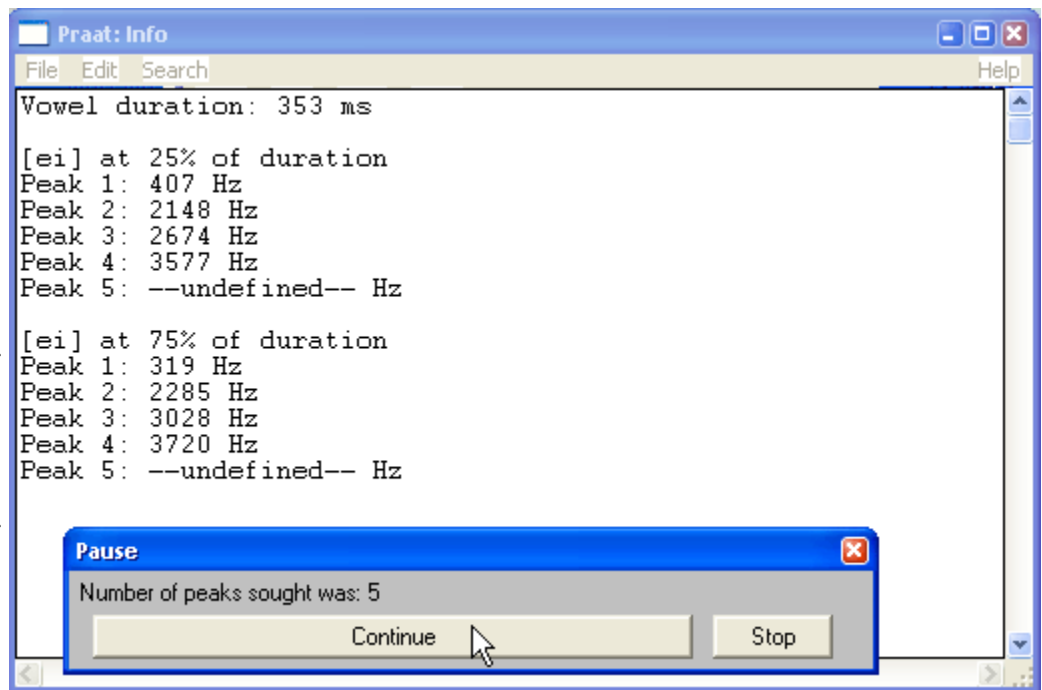
- then, 25% and 75% of the way between the two boundaries,

it extracts a 15 ms wide window and performs a Fourier analysis to calculate the spectral components of the portion of the vowel inside the window

- in the Praat picture window it draws a non-smoothed spectrum (in red) at 25% (top panel) and 75% (bottom panel) of the duration of the vowel

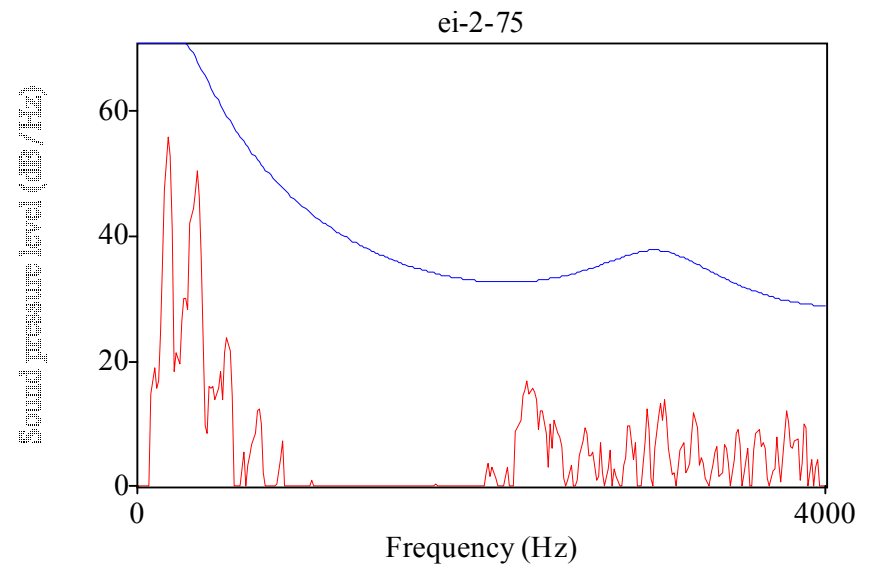
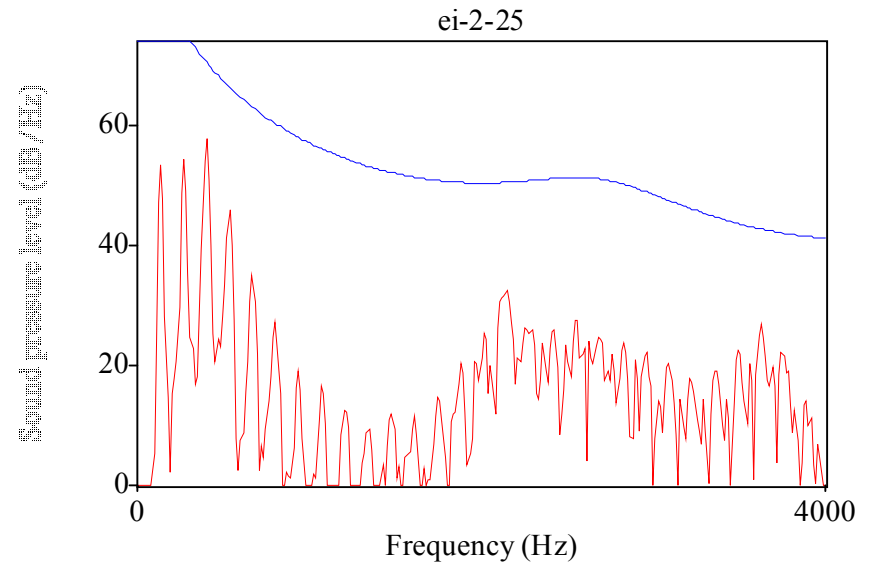
- it performs a Linear Predictive Coding (LPC) smoothing of each spectrum and draws it (in blue) on top of each non-smoothed spectrum

- in the Praat info window, the script prints the duration of the vowel and the frequency at each peak in the LPC smoothed spectra

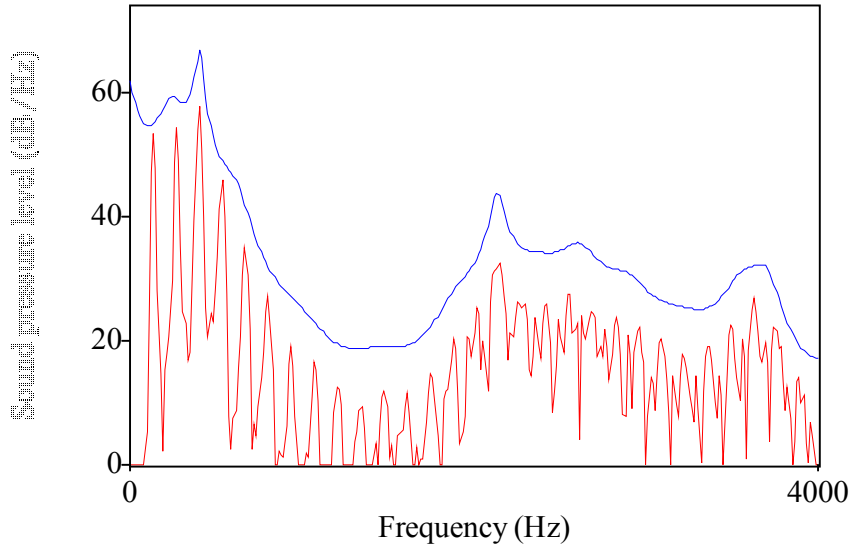


- a problem for measuring formants it to determine exactly how much smoothing to apply
 - too much smoothing misses formant peaks
 - too little smoothing wraps peaks around individual harmonics
- the script helps solve this problem by first resampling the sound recording to 8 kHz so that it only looks for peaks below 4 kHz
 - the maximum signal frequency which can be encoded by an 8 kHz sampling frequency is 4 kHz (this is called the Nyquist frequency)
 - F1 and F2 will always be well below 4 kHz
- the script then gradually increases the number of LPC coefficients, which decreases the amount smoothing
 - the script tries a low number of coefficients
 - it pauses to give the user the opportunity to look at the smoothed spectra
 - if the user clicks on the *Continue* button, the script increase the number of coefficients, then pauses again

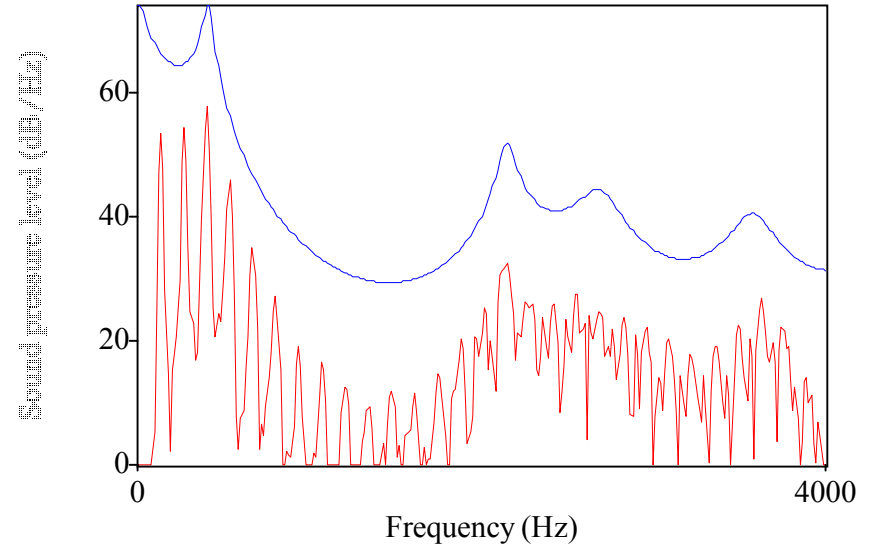
These spectra are too smooth



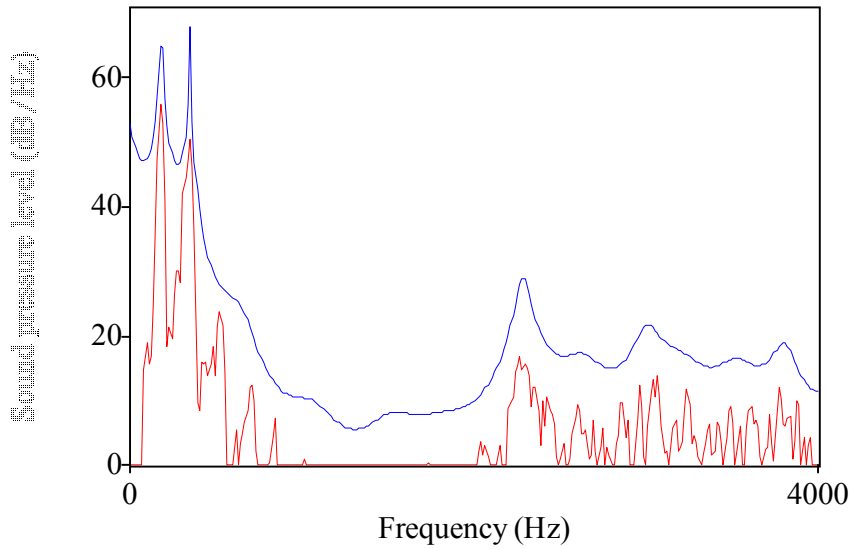
These spectra aren't smooth enough
ei-13-25



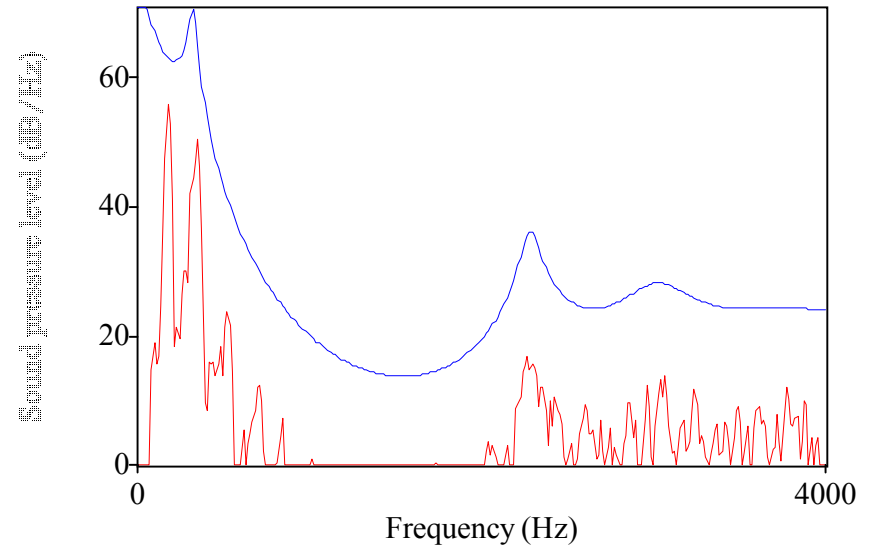
These spectra are just right
ei-5-25



ei-13-75



ei-5-75



How to get reasonable answers:

– Stop the script when it looks like it has picked the right amount of smoothing. Do this by visually comparing at the non-smoothed and smoothed spectra. Plan to go past the ideal number of peaks the first time you run the script, then run the script again. A rule of thumb for using the script is to visually count the number of formant peaks you think you can see in the non-smoothed spectra, and stop when the LPC smoothing is one greater than this number of peaks (in the “just right” example above the number of peaks sought was 5).

– At this point you should have some a priori ideas about what reasonable formant values are for the vowel you are measuring, e.g., for [i] you expect F1 and F2 to be far apart, for [u] you expect both F1 and F2 to be relatively low, for [ɑ] you expect F1 to be relatively high and F2 to be relatively low. If you have too much or too little smoothing you may get formant measures that don't correspond to your expectations.

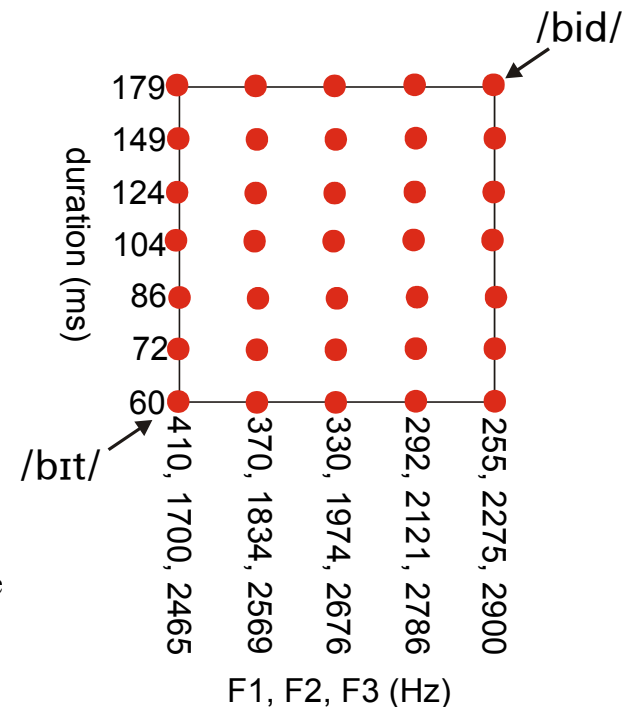
– The script plays the vowel and the windowed portions of the vowel. Make sure that what you hear corresponds to what you think you should be hearing for that vowel. You could do an auditory check of your results by feeding your measured formant values into *DiphthSynth.PraatScript*. Vowels in isolation often trail-off gradually getting quieter sometimes with a schwa at the end, try rerecording or try moving the second boundary marker towards the centre of the vowel.

– Male voices are often easier to measure than female or child voices because the lower fundamental frequency provides more closely spaced harmonics which result in a better sampling resolution for the resonance envelope. Some speakers may have creaky voice or other properties which make them difficult to measure. If you are having difficulty measuring your own voice, try downloading and measuring the male-voice samples available on the course website. The settings for the Nyquist frequency, and the window width are specified near the beginning of the script, you could also try manually changing those.

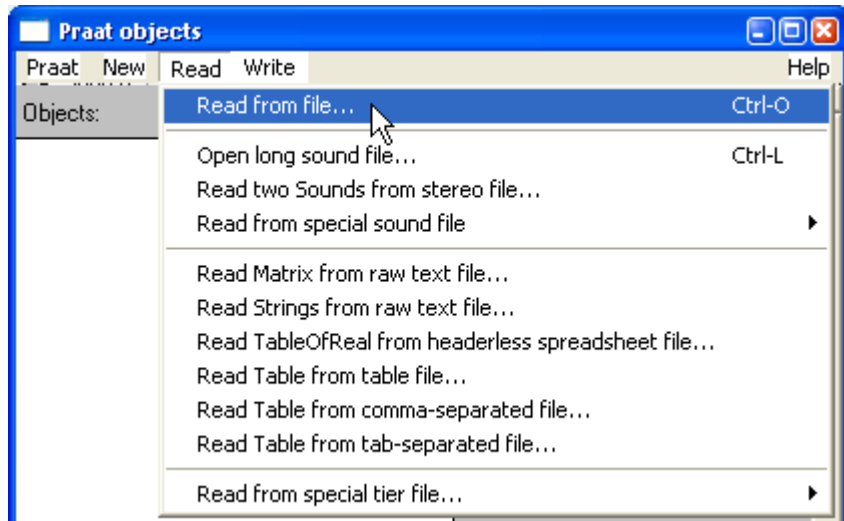
Optional: If you want to measure fundamental frequency, there are several ways to do this. One way is to open an unsmoothed spectrum in the Edit window, then place the cursor on the spike corresponding to the first harmonic and read off the frequency.

Vowel-Consonant Perception Experiment

- This is a demonstration of a perception experiment.
- A speaker read sentences containing the words /bit/, /bid/, /bit/, /bid/. The vowel duration and F1, F2, and F3 were measured.
- The stimuli were produced via LPC resynthesis of one of the speaker's /bit/ productions.
- Two properties were manipulated
 - vowel duration
 - vowel spectral properties
- The stimuli have 7 different durations and 5 different spectral properties combined to produce a matrix of 35 stimuli.
 - The shortest stimuli had a vowel duration typical of the speaker's /bit/ productions.
 - The longest stimuli had a vowel duration typical of the speaker's /bid/ productions.
 - Intermediate stimuli were spaced an equal duration apart on a logarithmic scale.
 - The stimuli with the highest vowels had F1, F2, and F3 properties typical of the speaker's /bit/ productions.
 - The stimuli with the lowest vowels had F1, F2, and F3 properties typical of the speaker's /bid/ productions.
 - Intermediate stimuli were spaced an equal distance apart on a straight line between the production values for /bit/ and /bid/ in the F1–F2–F3 space using a mel scale. (F1, F2, and F3 formant tracks were raised or lowered an equal amount along the whole of their trajectory.)

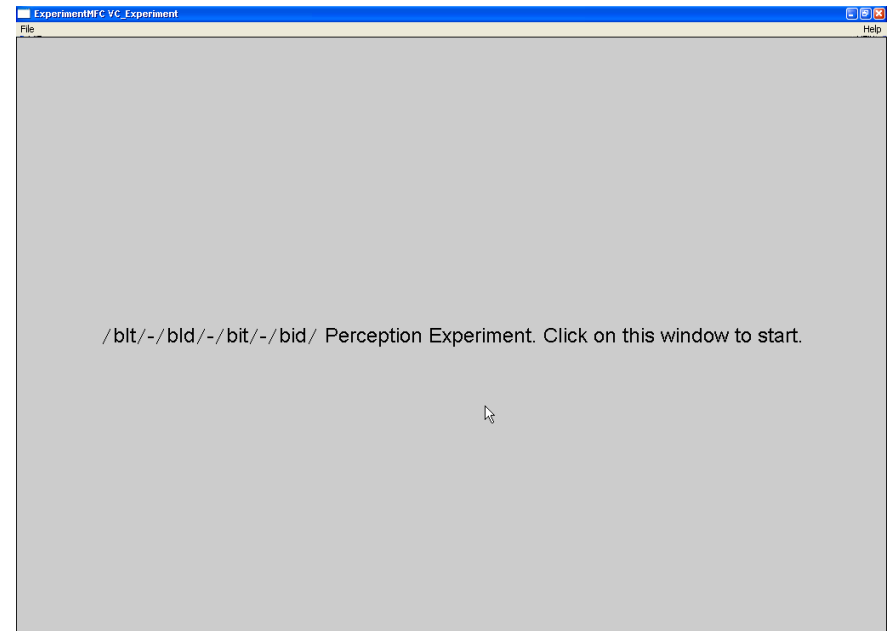
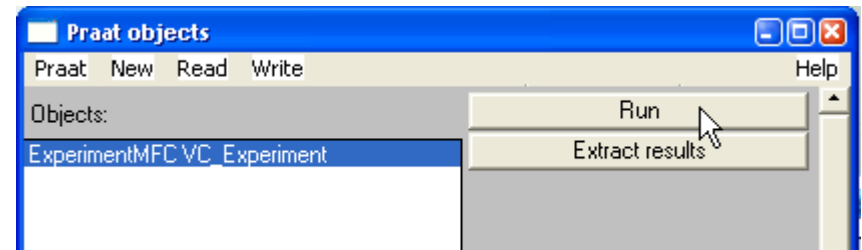


- Download and unzip *VCexperiment.zip*
- Open Praat and then open *VC_Experiment.ExperimentMFC* using *Read > Read from file...*



- Click on the window to start the experiment

- Run the experiment



– You will hear the sentence “What they’re wearing are _____ suits.”

– The word in the blank is one of the LPC resynthesised stimuli.

– Choose the response /bɪt/, /bɪd/, /bɪt/, or /bɪd/ that is closest to the word you heard.

– 0.25 seconds after you answer, the computer will play the next stimulus.

– The 35 stimuli will be played in random order.

– When running a real experiment, it is normal to run through the stimuli several times and get several responses for each stimulus.

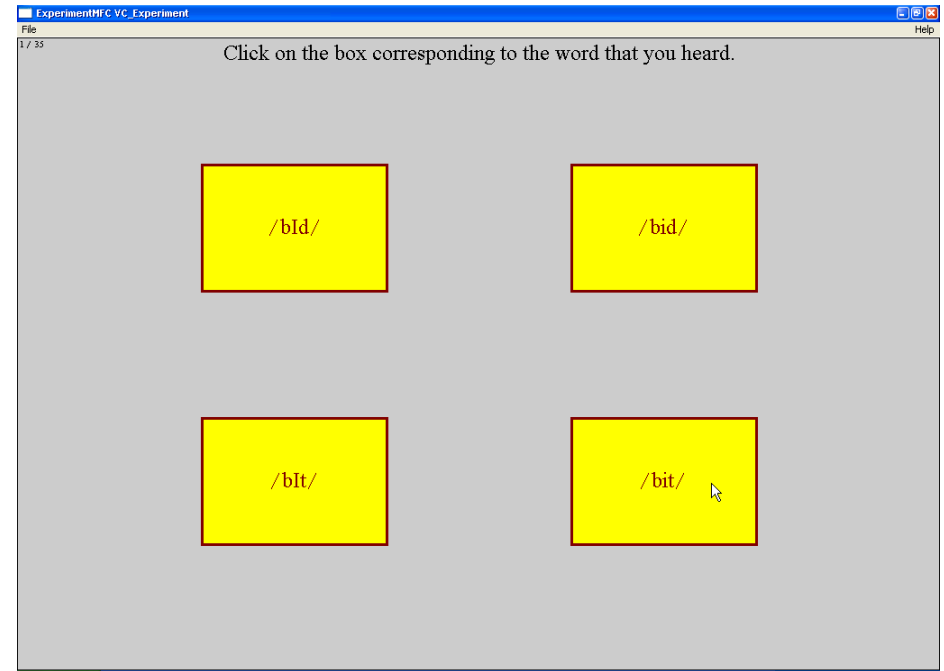
Stimuli near category boundaries may receive one response the first time and a different response the second time.

– For this demonstration you will only identify each stimulus once.

– Close the window at the end of the experiment.

– Open and run *VC_Analyse.PraatScript*

– A picture similar to the one below will appear in the *Praat picture* window. It shows which response you gave for each stimulus.



These are the results that I got when I completed the experiment.

Your answers may or may not be similar to mine.

– I distinguished the vowels only on the basis of spectral properties:

- the dividing line between /bit/–/bid/ and /bit/–/bid/ is perpendicular to the spectral axis, higher vowels were identified as /i/ and lower vowels as /ɪ/.

– I distinguished the consonant voicing on the basis of a mixture of duration and spectral properties:

- the dividing line between /bit/ and /bɪd/ was on a diagonal to the axes, stimuli with longer higher vowels were identified as ending in a voiced consonant, and stimuli with shorter lower vowels were identified as ending in a voiceless consonant.

– There was an interaction between vowel and consonant identification:

- When the vowel was identified as /i/, its duration had to be longer for the consonant to be identified as /d/, compared to when the vowel was identified as /ɪ/.

I ran a full-scale version of this experiment with L1-English and L1-Spanish L2-English speakers. The results are available under *Spanish Listeners' perceptual patterns for English /i/ and /ɪ/* on my website <http://www.ualberta.ca/~gsm2/>

