

Using SA (Speech Analyzer)

This discussion supposes that you are using a computer with SA installed on it. First we will go through the process of doing something with an existing sampled speech file, and after that we'll see how to use SA to create files of speech from tape recorded data. I recommend downloading the file Samplesound.wav from the Recordings subdirectory of the class page: this will give you something to experiment with, so that you don't have to learn how to record before learning how to listen.

Opening a file.

You can open a file in the standard Windows manner, pressing ctrl-o or using File-Open on the menu. This program uses the old-style "open" interface (which means that it doesn't print out the long file name), so you may see names like "sampl~1.wav" which probably doesn't mean very much to you. I will try keep the file names of the recorded class materials short (e.g. 2309a.wav, 2309b.wav and so on) so that you can figure out that 2309e.wav is the 5th file from the data recorded on September 23, if any data is recorded that day). The program remembers which directory you were last using, so you may only need to navigate the directory structure once to find where you are keeping files. [I will make suggestions for managing files later].

The displays for SA can be configured in many ways; I'll describe how to set it up one way, and we will consider other possible setups later. The displays are set up under Graph-Types [shortcut F7]; you can select from their pre-set options, or enter a custom setting. It is probably best to use just one or two panels for your displays. Almost always, you will want to include the Waveform, and possibly Auto Pitch or Spectrogram A. After you fiddle around with the settings and get something you like, you can save those settings under Tools-Startup, and either make the changes permanent or temporary. The rest of the discussion will assume that you have the program set up to display Waveform and Spectrogram A, though we might monkey with the windows and throw in Pitch.

The spectrogram display is actually slightly special, since you have to actually hit the button in the lower lefthand of the Spectrogram window that says "Recalc" — that's because making the spectrogram is somewhat slow. You may want to modify the setting for how the spectrogram is made: that's done under Graph-Parameters, then select "Spectrogram". Usually, you want a "Coarse" spectrogram (so you can see formants, and so that segment edges are crisper). You may also prefer Monochrome over Color; also, you may want to change the Maximum Frequency (you can go up to 11,000, using the standard sampling rate — I'll explain that later, but for now, you can assume that everything will be at the standard rate). Playing around with Maximum and Minimum Threshold doesn't hurt (it is intended as a way to make the "noise" go away on the display), but I generally find that it doesn't help.

You have two cursors for the file, left and right. The standard left-mouse button sets the left cursor to be wherever the arrow is. You set the right cursor with shift-left-mouse. The left and right cursors correspond to the beginnings and ends of parts that you might do something to (play, delete, copy), and they are apparently different colors (possibly green and red: that is not my area of expertise). Sensibly, the left cursor has to come before the right cursor in time, so if you try to put the end before the beginning, it will reset both cursors, putting the left cursor just a

bit before the right cursor. You can also move the cursors by sneaking up on the line with the mouse and when it becomes a finger, you can slide that cursor left and right.

One useful thing to do with the program is play something. The Playback menu has a number of preset options, a number of which are assigned to keys. Some useful presets are F8: “play between cursors”, F11: “play window” and F12: “play file” which will play just the part between the set cursors, just the part that you can see, or the entire file. Another very useful item under the Playback menu is “Player”. Here, you can change the playback speed to make it faster (of marginal utility) or slower (useful). Also, if you check the “repeat” box, it will play the same thing over and over, waiting however many milliseconds you set on the slider between repetitions (by default, it gives you a 1 second pause). Finally, where it says “Play”, you can tell it to play “Entire File”, or “Between Cursors”, or various other options. Last thing on “Player”, you need to hit the “play” radio-button to get it to actually yap. If you hit “okay” that means “I’m done with this Player menu”, and it does not mean “Okay, now play it”. Hit the “play” radio button to play.

Suppose you’ve positioned the cursors around the word in the speech file, and you want to do something with it. You can “select” it (which makes the selected part blue and the waveform portion inside the selected area white) with ctrl-space. You can delete the selected part with ctrl-x, and you can copy it with ctrl-c. [You always have to select something in order to be able to copy or cut]. You can zoom in and out with Graph-Zoom, then select zoom in, out, cursors (shortcut: ctrl-enter), or full view (ctrl-home). The fast way to focus on a smaller part of a waveform is to set the cursors to be in roughly the right place, then hit ctrl-space ctrl-enter, and that zooms into the portion between cursors.

One thing you’ll notice is that if you have selected a smaller section of the file and zoomed in, the spectrogram display is not affected. You have to explicitly hit “recalc” to make the spectrogram display match the (modified) portion of the waveform.

When you copy something, you probably want to put it somewhere. Edit-Paste as new (ctrl-n) opens a new window and copies the selection from the clipboard into the new window. It automatically gives it a name like “SA1” or “SA5” — Window will show you what windows are open. If you want to take a large file and cut it up into bits, you select the individual bits, then ctrl-n them into their own windows (if you like the results, I suggest saving and closing the window, so you don’t get confused about what you’ve done).

Here is a summary of the main functions that you’ll use most often.

F7	controlling the display type
<u>G</u> raph- <u>P</u> arameters	setting spectrogram parameters (also pitch)
left-mouse	set left cursor
shift-left-mouse	set right cursor
ctrl-space	select the portion between the cursors
ctrl-c, ctrl-x	copy or delete the selection
ctrl-n	paste the copied selection into a new window (name generated automatically)
ctrl-enter	zoom in to the part between cursors
ctrl-home	show the whole file

As with most programs, if you change a file it will ask if you want the changes saved. Since you can have a lot of files open but (by default — which can be changed) it displays only the most recent open file, you should pay attention to what file name it's asking you about (and whether you really want to save the changes).

Recording speech

Recording is somewhat easy, but it probably requires some advanced fiddling with settings and cables. Once you've gotten the hang of the fiddling necessary, the "setup before recording" part becomes very easy.

Suppose you want to record and look at a couple dozen words of Guarani. First you locate a tape recorder and some tape. I also sort of recommend that you use an external microphone, because the built-in mike that comes on most tape recorders is pretty horrid. There isn't any need to get a very fancy studio mike costing \$75 or more (truth in advertising time: I have no idea what they cost nowadays. My collection of mikes cost about \$25.00 apiece, but that was 5-10 years ago). If you use a fancy recorder, a fancy mike, and record in a soundproof room, you will get excellent recordings. However, you can also get more or less serviceable recordings using a rotten recorder, just the built-in mike, and do the recordings in the hall. Recording environment is most important, followed by microphone, then recorder.

Let's now assume that you have some recorded material, and you want to put it on the computer. You need to connect the tape recorder output (either "line-out" if there is such a thing, or "headphone") to the line-in port on your computer. The line-in port is a little 1/8 inch round hole, and there could be four or more of them on your computer. Once in a while they actually etch the words "line-in" near the line-in port. Usually, they don't, and instead there is some kind of illegible and meaningless icon. The port might be in the front or the back: look for where your speakers connect (that isn't infallible: on some computers, they put the speaker connector on the front and the line-in on the back). A possibility: there might only be one "in" port, that works for both microphone and line-in. I'm going to pretend that you've plugged the recorder output into the computer line-in, but if things don't work, you should bear in mind the possibility that you plugged something into the wrong hole. Inputting into an output hole results in no recorded sound at all, but inputting into the wrong type of input hole will result in bad sound (a sometimes useful diagnostic tool).

You also have to make sure that the volume control settings are correct on the computer, so that you can actually record and play things back. Open Volume Control by clicking the loudspeaker icon at the lower right, and you should get a menu with a bunch of sliders. The first thing you want to do is make sure the recording control settings are right. Go to Options-Properties and click "recording" (you're in trouble if there's nothing to click), make sure that Line-in is checked, and click OK. The crucial thing to do is check under "Line-In Balance" that Line-In is "checked" (if it says "Select": but it should *not* be checked if it says "Mute": I don't know why, but different versions say different things). You may also have to explicitly fiddle with the settings for Microphone, to de-select the microphone (or mute it). You don't want to have both line-in and microphone being "recorded" at the same time, and you do want to have line-in being recorded. Also, check the volume slider for line-in: it might be set to the bottom, so be sure it's at least half-way up (at least at the beginning).

At this point, you're probably ready to record. Start running the program SA, and click the microphone icon on the upper left. This will give you a box with a button labeled REC, also

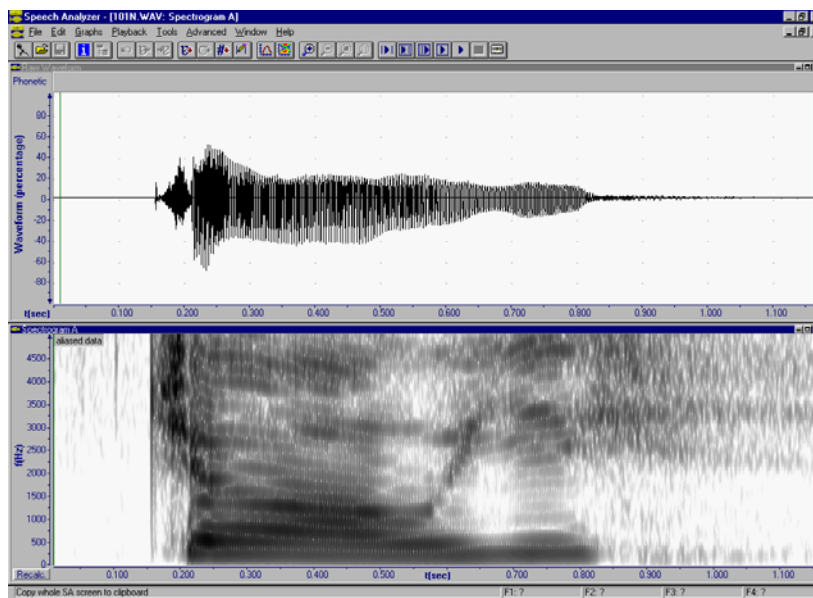
a graphic volume display, and it will print some settings. It should say “22050 Hz 16-bit mono highpass filtered”, If it doesn’t, click the Settings button and click buttons to select 22,050 Hz sampling rate, 16-bit and Highpass filter. If you don’t have a 16-bit option, that’s okay (albeit inconceivable). You can now start playing the tape, and as the signal comes through, you’ll have various indications that the signal is coming through. For example, you will hear something. Also, the volume display thing will start to twitch.

If you don’t hear anything, check that the speakers are plugged in and turned on, also check whether the playback sound is not muted, and so on. If you don’t hear anything and there is no twitch on the recording level indicator, maybe you’ve got the thing plugged into the wrong hole, so try one of the others, and check the “mute/select” buttons under volume control.

Supposing now that you’re getting a signal, it’s time to do actual recording. You start recording by clicking the REC button in SA, and stop by hitting Stop. If you hit the Play button, it will play back the stuff you just recorded. If it sounds reasonable and you want to use that bit of sound, press Apply. (That’s goofy, in my opinion, but the relevant word is “Apply”). When you do, it will then display the recorded waveform in a window (it creates a default name like SA1). You can now listen to it, edit it, save all of it or part of it.

When you first start making recordings, you may need to fine-tune the volume controls. Oh, I should give you a warning about blowing out your ears accidentally. If the input volume is too high and / or the playback volume is too high, and if you’re using earphones, you could hurt yourself. As a precaution, hold the earphones in your hands the first time you play it back, just to be sure that it won’t deafen you when you play it (you think I’m kidding?). One sign that you’ve got the recording volume too high is if you see a lot of black stuff that looks like it’s cut off straight across the top (it’s called “peak clipping”, which means it’s too loud). The solution is to drop the volume a notch or two. You should see nice smooth undulations on the waveform. The other possibility is, you have too little signal (so try cranking it up). Volume problems can arise if you have line-in plugged into the mike jack.

Here’s a picture of what a waveform with reasonable input levels might look like.



Note that this recording was made in non-lab conditions, with off the shelf equipment, and in a normal room

Once you have figured out which hole is the right one to plug the line-in into, and get the recording setting set up, you shouldn't have to go that more than once (unless you're sharing a machine with others and someone messes with the settings). At least you will be able to remember what you did to get it set up the last time.

A note on recording

We aren't going to focus on recording technology, but a few guidelines will help you to make decent recordings. One thing I suggest is "calibrating" your recorder, to see how much annoying noise you are going to get. Ideally, you want a sound-proof booth with a really good recorder and microphone, and no fidgeting. Rather often (ever is too often in my opinion), you will get interference from somewhere, which shows up as a faint hum. So check your equipment by recording nothing in silence (with the microphone plugged in and turned on). You ought to get a fairly flat line (usually there is a one bit fluctuation) doing this. If you get a hum, try a different recorder.

The microphone should not be too close to the speaker's mouth: a simple guide is, the distance from thumb to little finger. If you have a foam windscreen on the mike, this will probably eliminate the explosions as the lips open. The other important thing is that if the speaker holds the mike (probably necessary), try to be sure they hold it towards the bottom, to avoid hand-fidgeting being picked up by the mike.