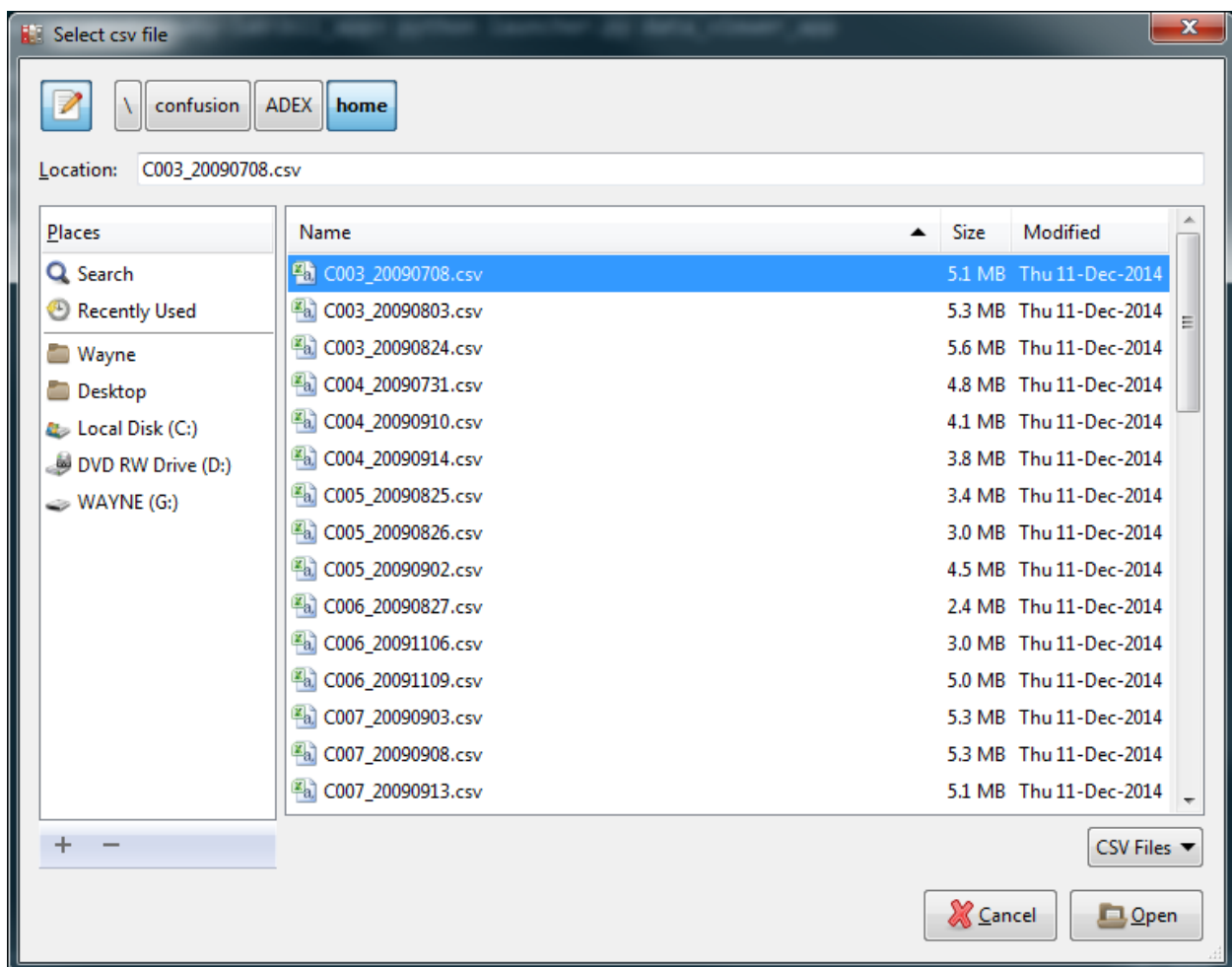


## Data Viewing Application

This program allows you to listen to the segments in an ADEX-like spreadsheet file. You can check off segments that have a particular significance. It is also possible to sort or filter the data. The results (including the checkboxes) can then be exported to a new spreadsheet file.

### Running the program

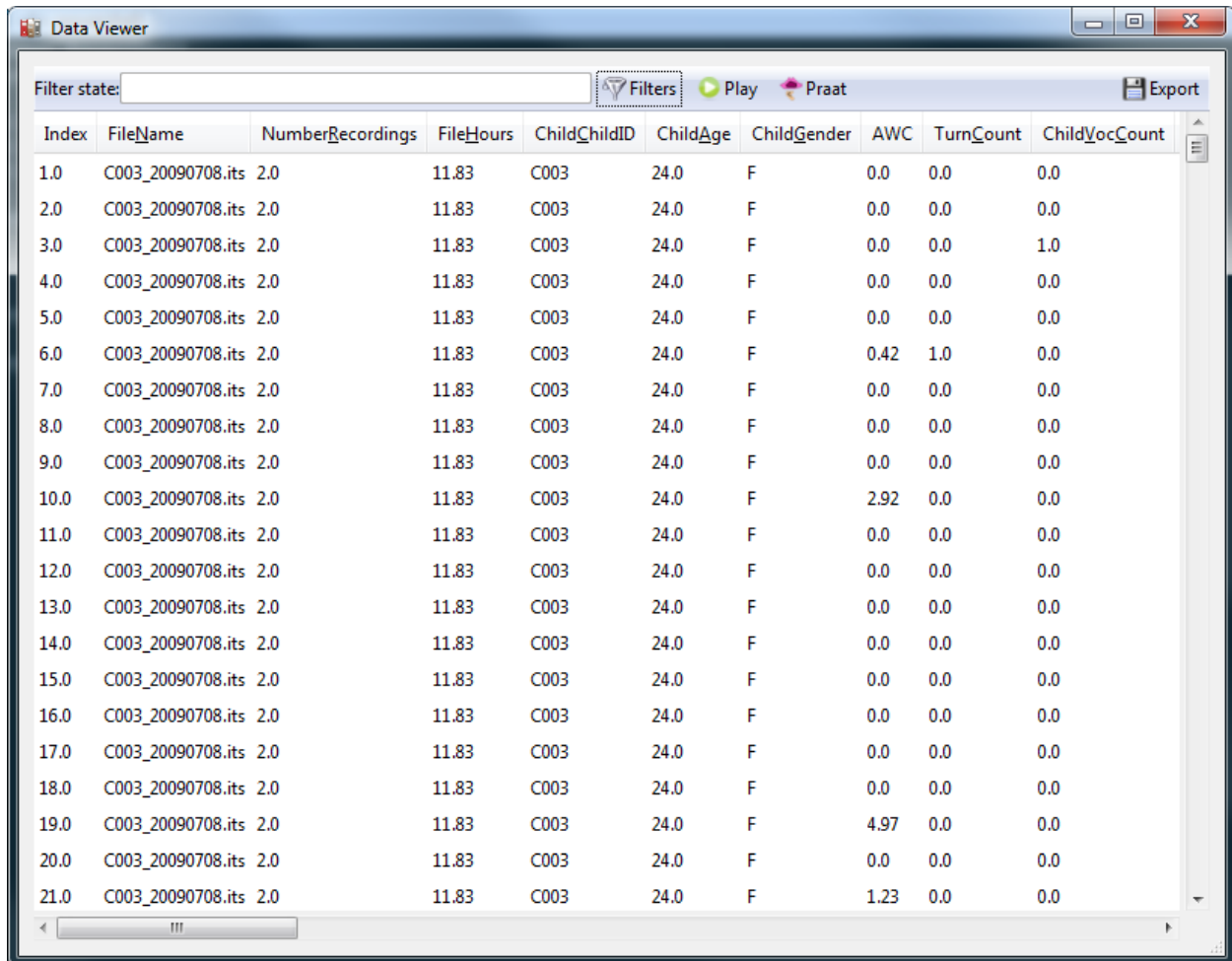
The LENA computer contains a desktop shortcut for the application. Starting the program brings up an “open csv file” dialog box. This window allows you to select the ADEX-produced spreadsheet file that you would like to open. This file must be in comma-separated value (.csv) format. Select a file and click “Open.”



At this point another similar window should open, this time asking for the audio file. Select the .wav file that corresponds to the ADEX spreadsheet you just picked. Most of the time, these can be found on the Drobo, under F:\Transcriber Files\Transcriber Files\Original Files\.

Once you’ve selected your file, click “Open”.

A progress bar is displayed as the file is opened. After the processing is complete, you should see a window that looks something like this (columns will differ depending on your spreadsheet):



Index	FileName	NumberRecordings	FileHours	ChildChildID	ChildAge	ChildGender	AWC	TurnCount	ChildVocCount
1.0	C003_20090708.its	2.0	11.83	C003	24.0	F	0.0	0.0	0.0
2.0	C003_20090708.its	2.0	11.83	C003	24.0	F	0.0	0.0	0.0
3.0	C003_20090708.its	2.0	11.83	C003	24.0	F	0.0	0.0	1.0
4.0	C003_20090708.its	2.0	11.83	C003	24.0	F	0.0	0.0	0.0
5.0	C003_20090708.its	2.0	11.83	C003	24.0	F	0.0	0.0	0.0
6.0	C003_20090708.its	2.0	11.83	C003	24.0	F	0.42	1.0	0.0
7.0	C003_20090708.its	2.0	11.83	C003	24.0	F	0.0	0.0	0.0
8.0	C003_20090708.its	2.0	11.83	C003	24.0	F	0.0	0.0	0.0
9.0	C003_20090708.its	2.0	11.83	C003	24.0	F	0.0	0.0	0.0
10.0	C003_20090708.its	2.0	11.83	C003	24.0	F	2.92	0.0	0.0
11.0	C003_20090708.its	2.0	11.83	C003	24.0	F	0.0	0.0	0.0
12.0	C003_20090708.its	2.0	11.83	C003	24.0	F	0.0	0.0	0.0
13.0	C003_20090708.its	2.0	11.83	C003	24.0	F	0.0	0.0	0.0
14.0	C003_20090708.its	2.0	11.83	C003	24.0	F	0.0	0.0	0.0
15.0	C003_20090708.its	2.0	11.83	C003	24.0	F	0.0	0.0	0.0
16.0	C003_20090708.its	2.0	11.83	C003	24.0	F	0.0	0.0	0.0
17.0	C003_20090708.its	2.0	11.83	C003	24.0	F	0.0	0.0	0.0
18.0	C003_20090708.its	2.0	11.83	C003	24.0	F	0.0	0.0	0.0
19.0	C003_20090708.its	2.0	11.83	C003	24.0	F	4.97	0.0	0.0
20.0	C003_20090708.its	2.0	11.83	C003	24.0	F	0.0	0.0	0.0
21.0	C003_20090708.its	2.0	11.83	C003	24.0	F	1.23	0.0	0.0

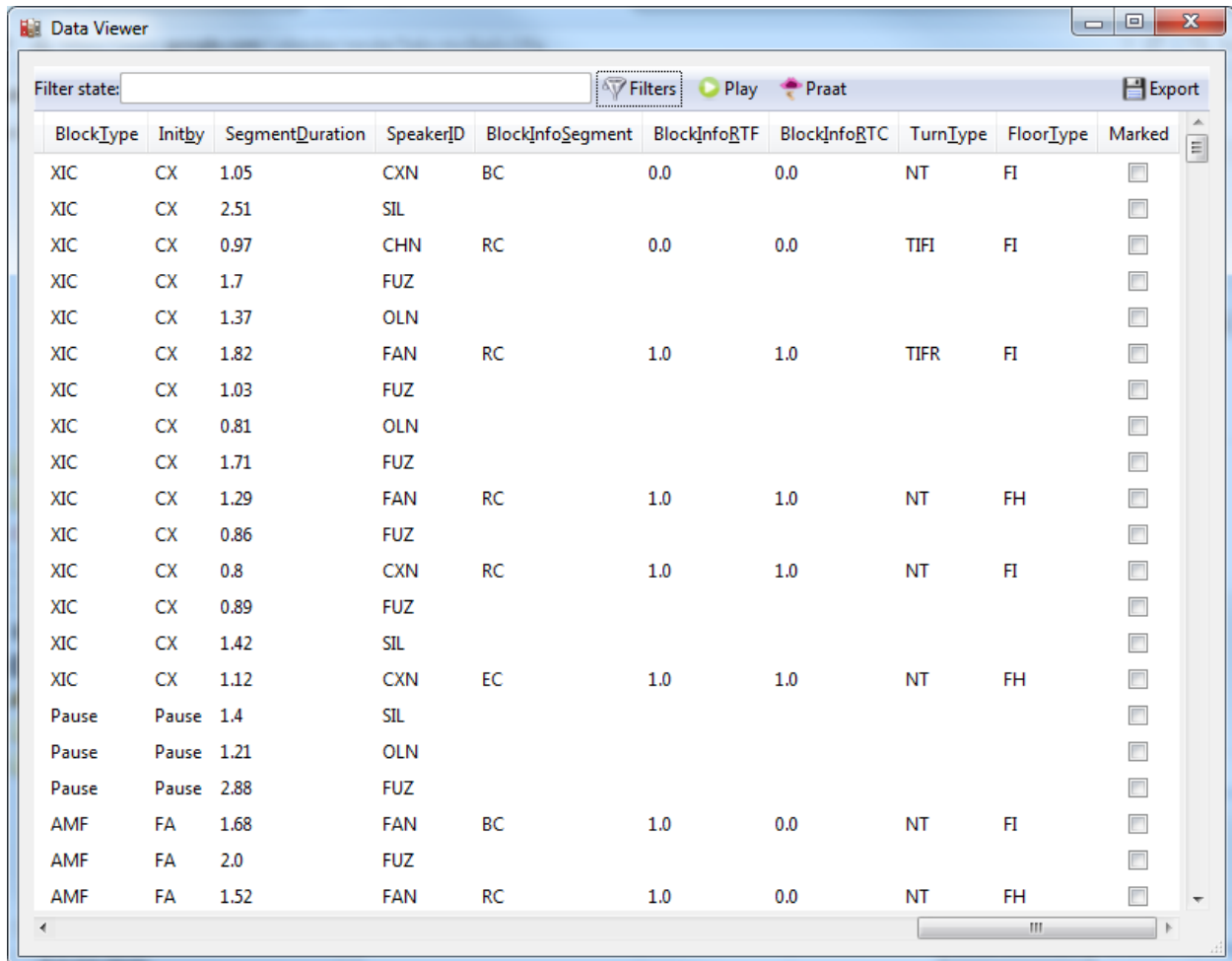
### Listening to audio clips

To listen to a segment, select its row, and click the “Play” button in the toolbar along the top of the window.

In order for this to work, the spreadsheet must contain two columns (named “Wav.Begin” and “Wav.End”) that give the absolute starting and ending times (in seconds) for each row. If your file doesn’t contain these columns, you can create them fairly easily in Excel using the “Elapsed Time” and “Segment Duration” columns. Just be aware that LENA can sometimes reset the “Elapsed Time” column. (You should set things up so that the start time of each row (except the first) is equal to the end time of the last row. The end time should be equal to that start time you just calculated, plus the segment duration.)

### Checking off rows

If you're looking for something in the audio, you can check off the segments that contain it. If you scroll to the far right, you'll see a column of checkboxes called "Marked":



The screenshot shows the 'Data Viewer' window in Praat. It contains a table with the following columns: BlockType, Initby, SegmentDuration, SpeakerID, BlockInfoSegment, BlockInfoRTF, BlockInfoRTC, TurnType, FloorType, and Marked. The 'Marked' column contains checkboxes for each row. The table lists various audio segments, including speech (XIC, AMF) and pauses, with associated speaker IDs and segment durations.

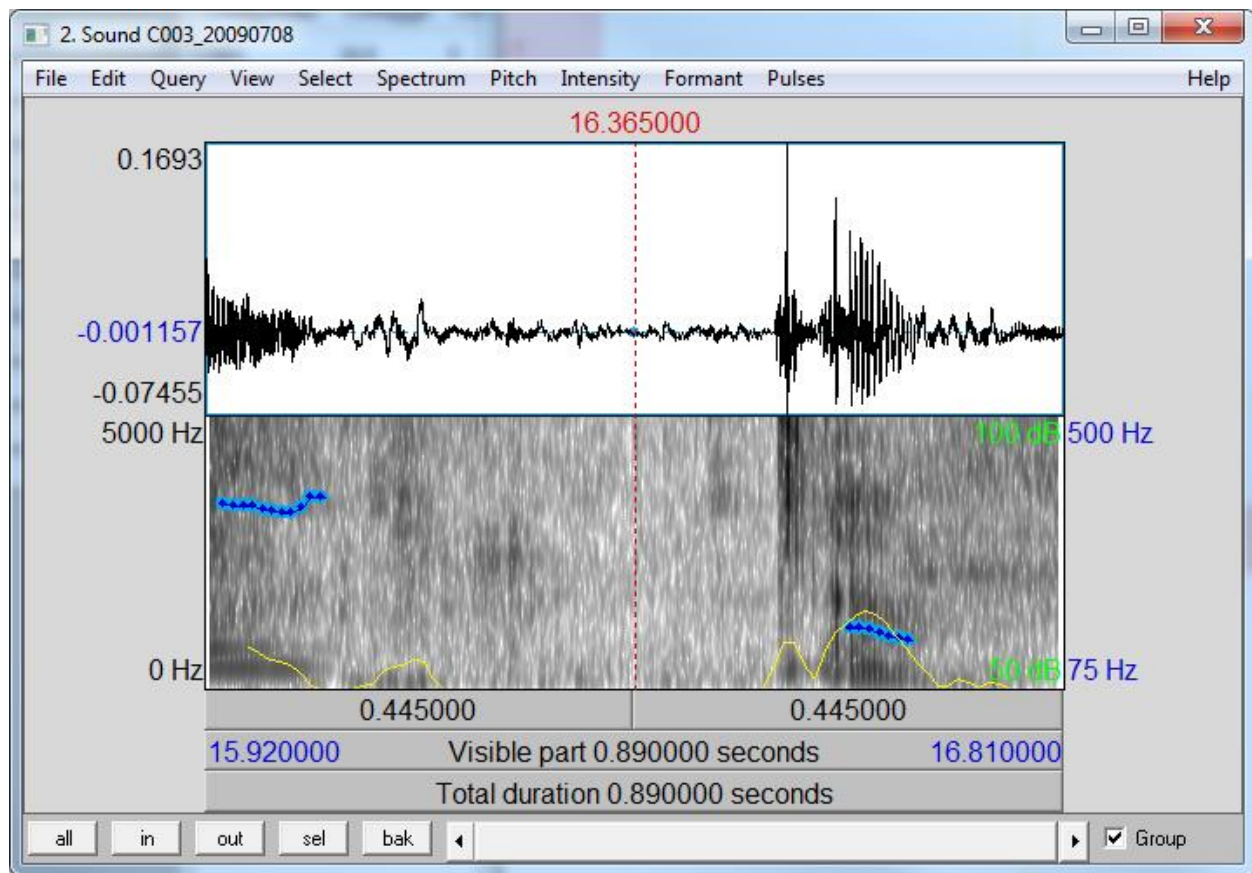
BlockType	Initby	SegmentDuration	SpeakerID	BlockInfoSegment	BlockInfoRTF	BlockInfoRTC	TurnType	FloorType	Marked
XIC	CX	1.05	CXN	BC	0.0	0.0	NT	FI	<input type="checkbox"/>
XIC	CX	2.51	SIL						<input type="checkbox"/>
XIC	CX	0.97	CHN	RC	0.0	0.0	TIFI	FI	<input type="checkbox"/>
XIC	CX	1.7	FUZ						<input type="checkbox"/>
XIC	CX	1.37	OLN						<input type="checkbox"/>
XIC	CX	1.82	FAN	RC	1.0	1.0	TIFR	FI	<input type="checkbox"/>
XIC	CX	1.03	FUZ						<input type="checkbox"/>
XIC	CX	0.81	OLN						<input type="checkbox"/>
XIC	CX	1.71	FUZ						<input type="checkbox"/>
XIC	CX	1.29	FAN	RC	1.0	1.0	NT	FH	<input type="checkbox"/>
XIC	CX	0.86	FUZ						<input type="checkbox"/>
XIC	CX	0.8	CXN	RC	1.0	1.0	NT	FI	<input type="checkbox"/>
XIC	CX	0.89	FUZ						<input type="checkbox"/>
XIC	CX	1.42	SIL						<input type="checkbox"/>
XIC	CX	1.12	CXN	EC	1.0	1.0	NT	FH	<input type="checkbox"/>
Pause	Pause	1.4	SIL						<input type="checkbox"/>
Pause	Pause	1.21	OLN						<input type="checkbox"/>
Pause	Pause	2.88	FUZ						<input type="checkbox"/>
AMF	FA	1.68	FAN	BC	1.0	0.0	NT	FI	<input type="checkbox"/>
AMF	FA	2.0	FUZ						<input type="checkbox"/>
AMF	FA	1.52	FAN	RC	1.0	0.0	NT	FH	<input type="checkbox"/>

If you check off the ones you're interested in, when you're done, you can export the results as a new spreadsheet. The spreadsheet will contain everything your old one did, plus this extra checkbox column (containing "True" in place of the checked boxes and "False" in place of the unchecked ones).

To do this this, click the "Export" button in the top right-hand corner of the window. This will bring up a "save file" dialog box.

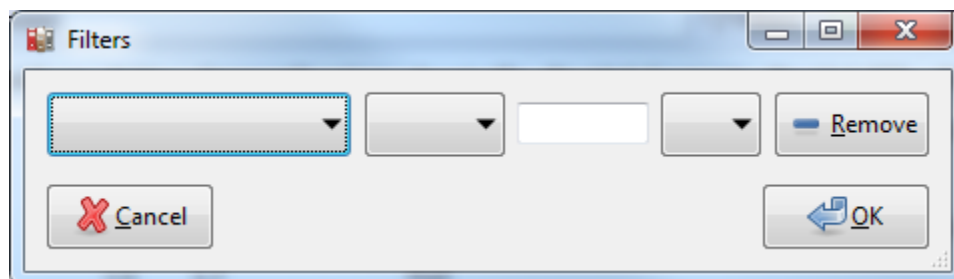
### Opening segments in Praat

You can open a segment in a Praat spectrogram window by selecting a row and clicking the "Praat" button in the toolbar. That'll look something like this:

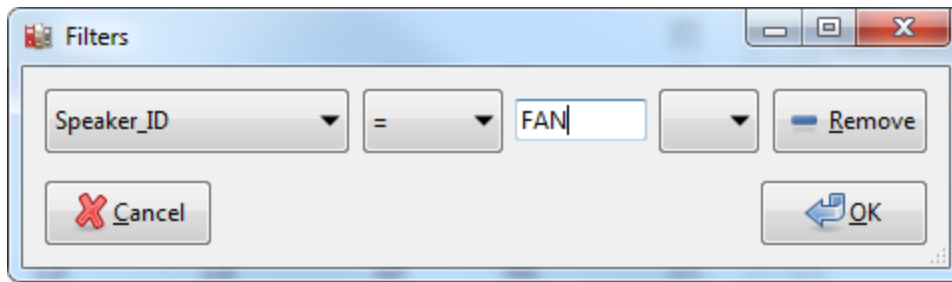


### Filtering the data

You can filter the view so that it only shows the specific rows you're interested in. To do this, click the "Filter" button in the middle of the toolbar. This will bring up a window that looks like this:

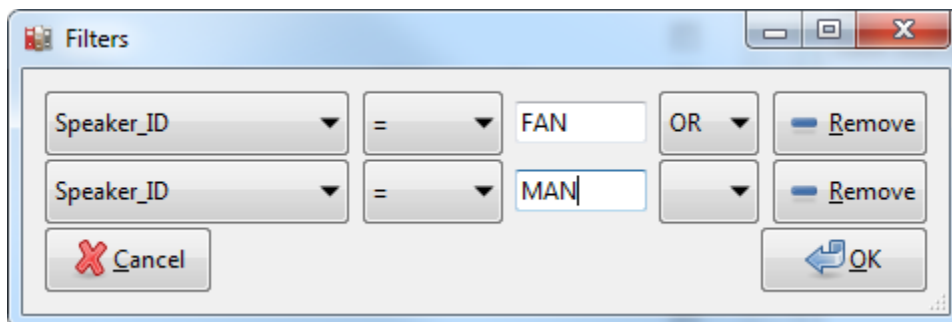


The easiest way to describe how this works is to show you an example. Let's say I wanted to filter the spreadsheet down to only those rows that have a LENA speaker ID of "FAN". I set the dropdowns like this:



I'll describe the dropdown boxes moving from left to right. The first one lists all of the columns in your spreadsheet. The next allows you to pick a search operator like "=" or "<". You can enter a value in the box that indicates what you're looking for.

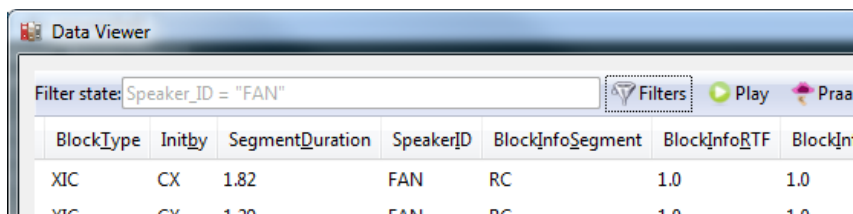
The rightmost dropdown allows you to add logical conjunctions (AND, OR) so that you can add more filtering criteria. For example, if I wanted all the rows that have a LENA speaker ID of "FAN" OR "MAN", I could do this:



Selecting a value (AND or OR) in the rightmost dropdown causes another row to appear. You can add as many criteria like this as you want. To get rid of a row, click the "Remove" button beside it, and then reset the conjunction dropdown of the previous row to blank.

When you're done, click ok. Filtering can take a long time if your spreadsheet is large. If you find things taking too long, you can always use Excel to pre-filter your spreadsheet and then load it into the data viewer.

Once a filter has been applied, you'll see it appear in the "Filter state" box, like this:



To remove all filters, click the "Filter" button and hit all the "Remove" buttons until there are no rows left. Then click ok.

### Sorting rows

There is a small arrow next to the heading name for each column. Clicking this arrow toggles the sorting state. There are three states: unsorted (sideways arrow), sorted ascending (up arrow), and sorted descending (down arrow). Clicking the arrow toggles cycles between them in this order.

Note that for very long files (> 100,000 rows), sorting can take some time. Other spreadsheet programs like Excel use some fancy tricks to make sorting much more efficient (e.g. only sorting rows that will show up in the currently viewable part of the window). However, this takes a long time to code, and is probably a little overkill for a simple script like this. If the waiting is significant enough to get really frustrating, just let me know. Alternatively you can use Excel to pre-sort the file, then save it and open it in the data viewer.

If you have to filter the spreadsheet (see below), one way to speed up sorting is to filter down to spreadsheet first so that it contains less rows. Only these rows will be sorted.