

# Chapter 7 Large Datasets

## About this Chapter

---

Sounds can be collected and managed in a variety of ways. Some continuous recordings are made into sequences of large files, while others are made into sequences of many small files. Other recordings are separated into smaller clips of relevant information by using detectors or browsing and making selections by hand and saving the selections into individual files. All of these cases fall into a broad category of having a lot of data to manage, or having a large dataset.

Raven offers many features for working with large datasets including paging, performing batch operations, and handling streams. This chapter includes sections on the following topics:

- Paging
- Batch operations
- Opening file sequences
- Opening multiple sound files in multiple windows
- Streams
- Opening sounds with altered speeds

## Paging

---


*Paged sound windows* enable you to open sounds that are too large to fit in the memory available to Raven. In a paged sound window, Raven loads only a specified amount of the sound, called a *page*, into memory at one time. The portion of the entire signal that is presently in memory is called the *current page*. A paged sound window includes a page navigation panel that enables you to jump easily to any other part of the signal that's not presently in memory. Paged sound windows can be used to display single sound files or file sequences.



Paged sound windows behave like non-paged windows, except that you cannot perform editing operations that alter the audio data (i.e. Cut, Paste, Delete, Filter, or Amplify).

Configuring a new paged sound window

When you first open a sound file or a sound file sequence, Raven displays the Configure New Sound Window dialog box, which gives you the choice of loading the entire sound into memory or creating a paged sound window (Figure 7.1). To open a sound in a new paged sound window, choose **Page Sound**. To specify how much of the sound should be loaded into memory at any given time, enter a value into the **Page size** field. You can choose units of seconds or minutes for page size from the drop-down menu.

 When you plan to browse large sections of data, you may wish to use a page size of twice the length you plan to view at a time. Then, before you begin browsing, zoom to half of the page length. This helps to reduce interruptions from progress bars as you scroll through the data.

Page increment and Step increment specify the amount by which the paged display moves when you click the **Page Forward**, **Page Back**, **Step Forward**, or **Step Back** controls as described in [“Moving the page incrementally: the Page and Step buttons”](#) on page 190. You can specify Page increment and Step increment either as a percent of the page size or as an absolute number of seconds or minutes by choosing the appropriate units from the drop-down menu next to each field.

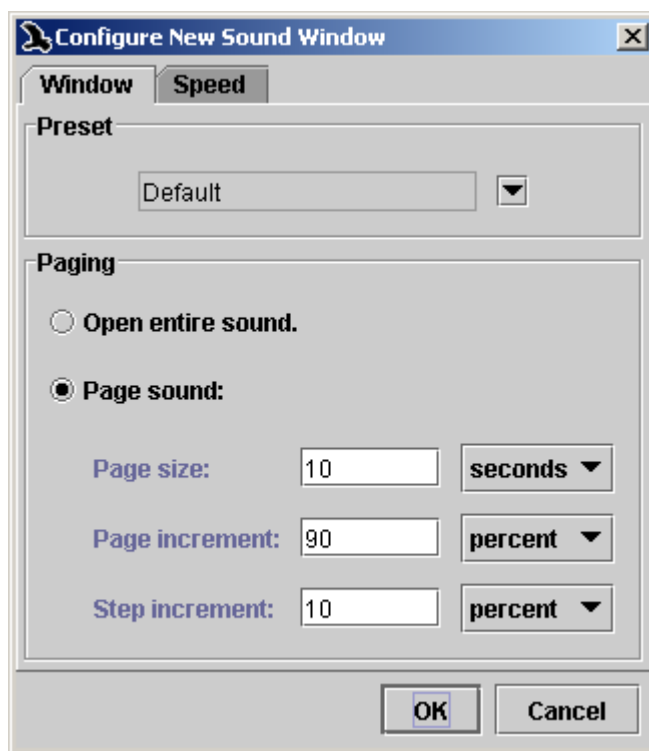
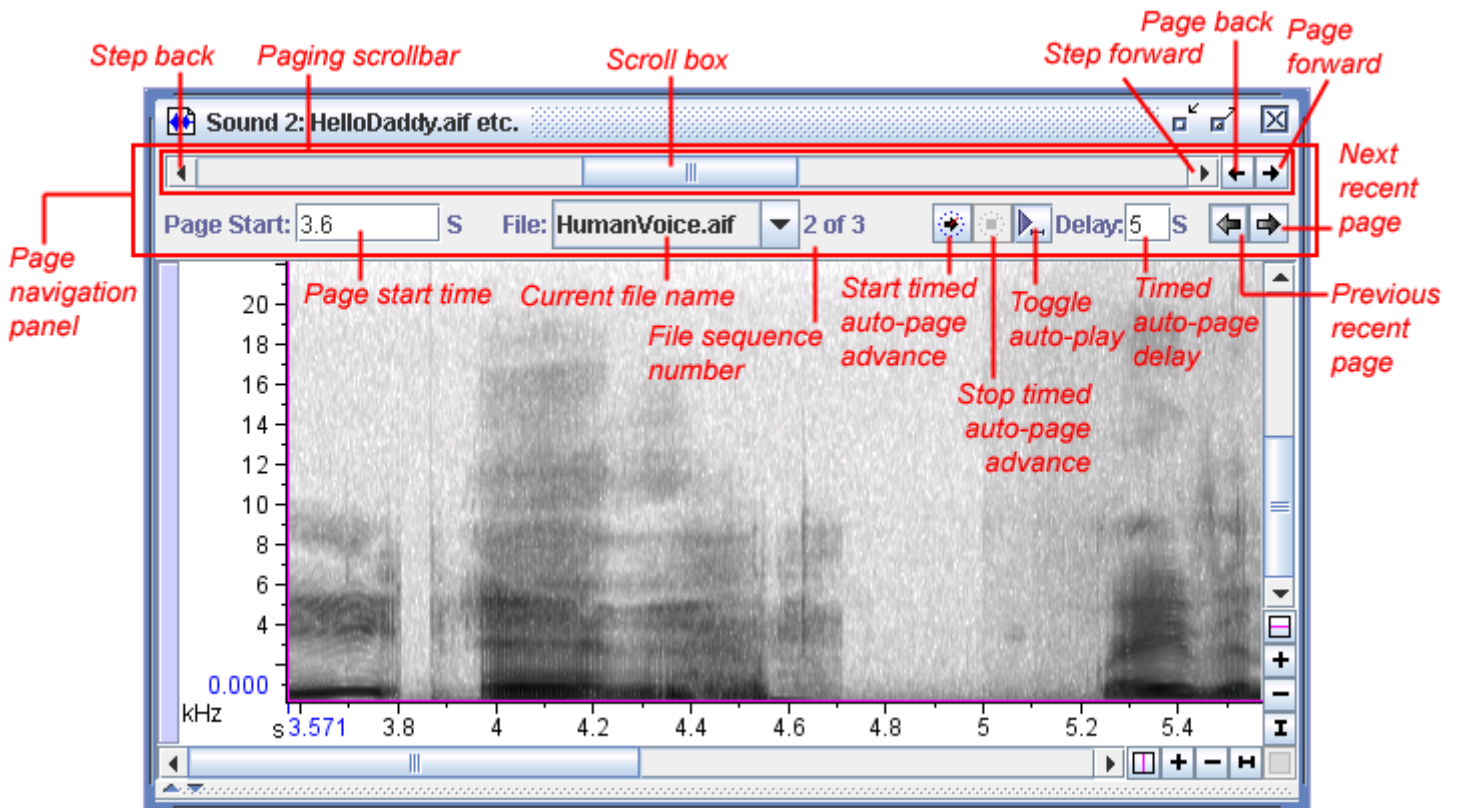


Figure 7.1. The Configure New Sound Window dialog box.

The Configure New Sound Window dialog box also allows you to choose a window preset to be applied to the new window when it is first created, as discussed in “Window preset and paging” in Chapter 1 (page 4).

When you open a single file or a file sequence in a paged sound window, a *page navigation panel* is displayed at the top of the window, above any views that are displayed (Figure 7.2).



**Figure 7.2.** A file sequence of five files opened in a paged sound window, with page size of 30 seconds. The entire file sequence is about four minutes long. For a single file opened in a paged window, the current file name and the file sequence number are not displayed.

### The paging scrollbar

The page navigation panel contains a *paging scrollbar* whose length corresponds to the entire signal, including those portions of the signal that are not presently in memory. The scroll box in the paging scrollbar represents the data in the current page (i.e., the data currently in memory). The length of this scroll box, relative to the length of the entire scrollbar, indicates what proportion of the entire signal is included in the current page. The position of the scroll box within the scrollbar indicates the position (in time) of the current page within the entire signal. The start time of the current page is displayed below the scrollbar. You can move the page to any point within the signal by dragging the scroll box. As you

drag the scroll box, the displayed start time of the page is continually updated to correspond to the position of the scroll box. If the signal is a file sequence, the name and sequence number of the file containing the start of the page are also continually updated as you move the scroll box. When you release the scroll box at a new start time, Raven loads the data for the new page position as needed from the file or file sequence and updates the views displayed in the window. Depending on the page size and the number and type of views shown, there may be a slight delay while Raven updates the views.

Clicking in the scrollbar to the right or left of the scroll box pages forward or back, equivalent to clicking the Page forward or Page back button (see below).



Don't confuse the functions of the two horizontal scrollbars in a paged sound window. The scrollbar in the page navigation panel at the top of a paged sound window refers to the time dimension of the entire signal. The scrollbar and navigation controls at the bottom of the window refer to the horizontal axis of the active view for the current page.

#### *Moving the page incrementally: the Page and Step buttons*

The Page Forward, Page Back, Step Forward, and Step Back buttons (Figure 7.2) move the current page to a new location within the entire signal by large or small increments. The amount that the page moves forward or back through the signal depends on the values that were entered in the Page increment and Step increment fields of the Configure New Sound dialog box (see “Configuring a new paged sound window” on page 188). By default, paging forward or back moves the page by 90% of the page size, resulting in a 10% overlap between successive pages. Stepping forward or back moves the page by 10% of the page size by default, resulting in a 90% overlap between successive pages.

#### *Jumping directly to a particular time within a signal*

To jump to a particular page start time within the entire signal, enter a time value in the Page Start Time field in the page navigation panel (Figure 7.2), and press <Enter>. You can enter a time either in units of seconds, or minutes and seconds, or hours, minutes, and seconds. The latter two formats use colons to separate the units, for example 1 : 02 : 03 to represent 1 hour, 2 minutes, and 3 seconds.

#### *The current file indicator; jumping to a particular file in a sequence*

For file sequences, the page navigation panel displays the name of the file containing the start of the current page. To move the start of the page to the beginning of a particular file, choose the name of that file from the file name drop-down menu.

*Timed auto-page advance*

To facilitate paging through long files, you can choose to automatically page through a file at a specified rate. The buttons to the right of the current file indicator allow you to start and stop auto-paging as necessary. You can also choose the amount of time that is spent on each page before advancing by changing the value in the timed auto-page delay field next to the start and stop buttons. You can choose whether or not to play each page as you advance using the auto-play toggle button. If the toggle button is on, Raven plays each page before waiting the specified amount of time and paging forward.

*Jumping to recently viewed pages*

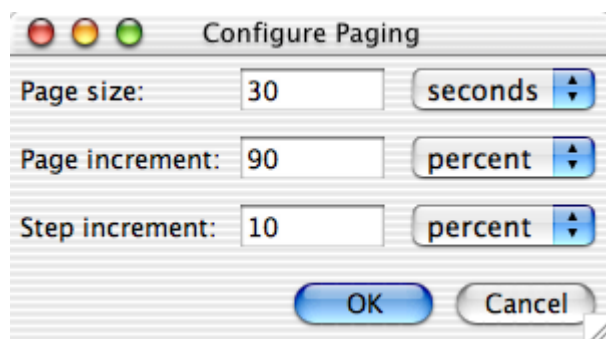
You can also navigate to recently viewed pages by using the buttons at the bottom right of the page navigation panel. This may be useful if you are moving around in a file to browse selections or other events.

**Selections spanning pages** You can create a selection that spans boundaries between pages by shift-clicking. Click once at the start or end of the interval you wish to select, then page forward or back as needed, and shift-click at the other limit of the interval. Once the selection is created, you can copy or save it as you would with any selection in a non-paged window.

**Reconfiguring paging for an existing window** You can change the page size and the page and step increments for a paged sound window at any time by choosing **Configure Paging...** from the contextual menu for the page navigation panel (right-click with a 2-button mouse or **<Control> + click** on a Mac). You can then enter new values for any of the page configuration parameters in the **Configure Paging** dialog box ([Figure 7.3](#)). If you adjust the page size or number of lines while zoomed to show an entire page of sound, Raven will automatically resize the view to show the entirety of the new page size. However, if you are zoomed in to a shorter timespan, the view will not be resized. If you would rather automatically zoom to view an entire page when the page size or number of lines is changed, you can change the corresponding entry in the preferences file from false to true.

```
raven.ui.soundWindow.pageSizeChanged.zoomToAllX=false
```

```
raven.ui.soundWindow.numberOfLinesChanged.zoomToAllX=false
```



**Figure 7.3.** The Configure Paging dialog box, which is used to reconfigure paging for an existing paged sound window.

## Batch Operations

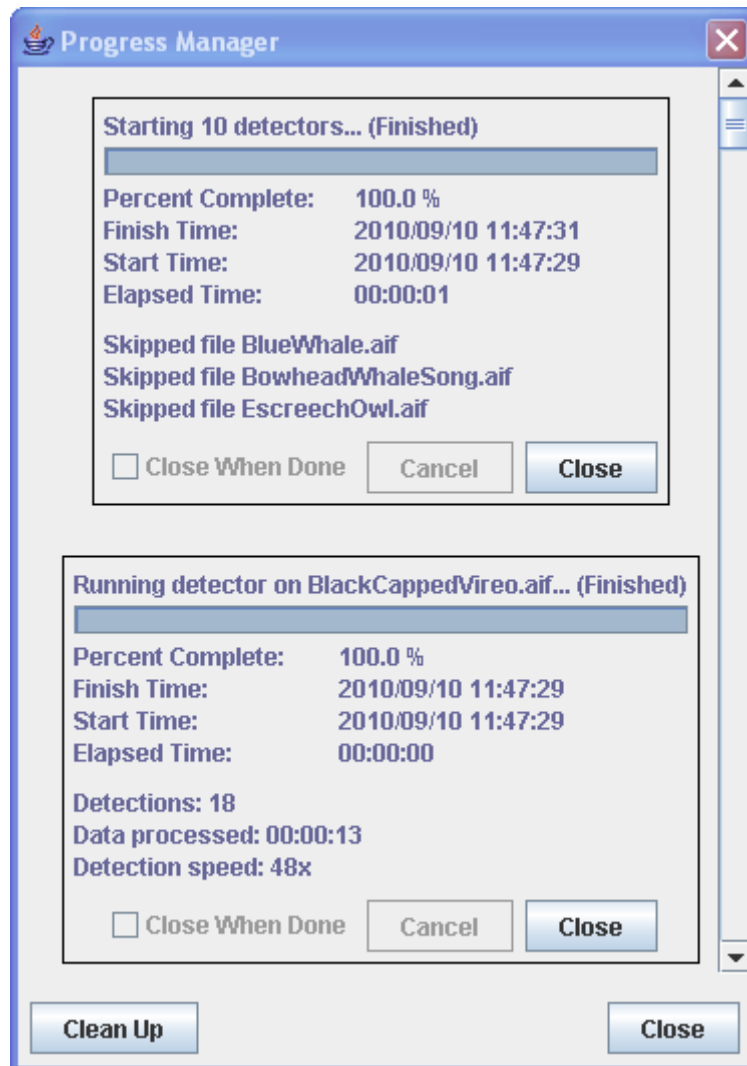
---

Raven has a number of features for working with many files simultaneously. The batch detector allows you to perform simultaneous detection runs on large numbers of files. The batch channel exporter allows you to export a specified subset of channels from all the files in a given folder. The batch correlator performs the same correlation on a large number of files and displays the peak correlation values in a table for easy comparison. Finally, it is also possible to filter or amplify a set of files by using the batch filter and amplify tools. For complete descriptions of these features, see [“Batch channel export” in Chapter 8](#) (page 212), [“Batch correlation” in Chapter 9](#) (page 240), [“Batch Filtering and Amplifying” in Chapter 6](#) (page 150), or [“Batch detection” in Chapter 10](#) (page 262).

## Background Tasks

---

**Progress Manager** When you work with large datasets, some operations can take a long time to run. Raven allows you to run many of these operations in the background while you continue to work. The Progress Manager keeps track of these background tasks, showing you how close they are to completion, how much time has elapsed since they began, and other information specific to the tasks ([Figure 7.4](#)). The Progress Manager will appear automatically when a new background task begins. If you close or hide it to work on another task and then would like to open it again, click on **Window > Progress Manager** or click the Progress Manager button in the toolbar.



**Figure 7.4.** The Progress Manager with two finished task panels visible. Note the Clean Up button in the lower left.

A new panel opens in the Progress Manager for each background task you begin. By default, these panels will remain in the Progress Manager after their tasks are complete, so that you can review information about the tasks. If you would like a panel to close automatically after its task finishes, check the **Close When Done** checkbox in that panel. Once you check that checkbox in one panel, subsequent panels will have the checkbox checked by default. At any point you can close a finished panel by clicking its **Close** button, or clear all finished panels from the Progress Manager by clicking on the **Clean Up** button.

Tasks that appear in the Progress Manager include batch correlation, batch detection, and batch amplification and filtering (see [“Batch Operations” in Chapter 7](#) (page 192)). Other tasks, such as opening large files or generat-



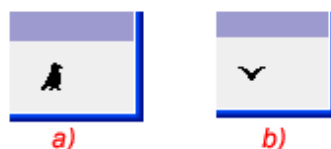
ing spectrograms, do not run in the background. These tasks' progress bars will appear in separate windows.

By default, Raven displays a progress dialog if it estimates that a task will take longer than two seconds. If you'd like Raven to wait a longer or shorter period of time, change the following preference:

```
raven.ui.progressDialog.thresholdTime=2
```

Change the number 2 to the length of time you'd like Raven to wait, in seconds. For more information about how to use preferences, see ["About Raven preferences" in Chapter 11](#) (page 299).

**Background task indicator** The raven icon in the lower right corner of the Raven window indicates whether a background task is running. Most of the time, when no background tasks are running, the raven is perched. However, if a task such as a batch operation or large detector is running in the background, the raven will be in flight (See [Figure 7.5](#)).



**Figure 7.5.** Raven icon shown (a) with no background task running (b) while running a background task.

## Opening file sequences

---

Raven allows you to open multiple sound files as a *file sequence*, provided that the files have the same file format (e.g., AIFF or WAVE), sample rate, and number of channels. When files are opened as a sequence, the audio data in the files are concatenated into a single sound signal that opens in one window. A sound window containing data from a file sequence looks and behaves the same as a single-file sound window with a few small exceptions. When a file sequence is opened, the title bar of the sound window displays the name of the first file followed by "etc". To find the name of the file that is currently displayed, position the mouse cursor over any part of the sound view. That section's file name will be displayed in the mouse measurement field, just below the Raven desktop ([Figure 7.6](#)). Also, the Cut, Delete, and Paste operations cannot be used on a file sequence, as changing the length of the file sequence would result in the file boundaries appearing in the wrong places.

You can open a sound file sequence either by giving Raven the name of a *list file*, or by manually choosing a set of files to be opened as a sequence.



**Using a list file** A list file is a plain text file (filename extension .txt) that lists the names of a series of sound files to be opened together as a sound file sequence. A list file is created automatically by Raven when you record to a file sequence (see “[Recording to a file sequence](#)” in [Chapter 4](#), page 92). You can also create a list file manually in a word processing or text editing program.

To open a file sequence using a list file, choose **File > Open Sound Files...** or type <Ctrl-O> (Windows, Linux) or <Command-O> (Mac OS). In the Open Sound Files dialog box that appears ([Figure 1.2, page 3](#)), choose or type the name of the list file, then click **OK**. In the Configure New Sound Window dialog box, you can then choose either to open the entire file sequence or to open the sequence in a paged sound window. Paged sound windows are discussed in “[Configuring a new paged sound window](#)” on page 188. A new sound window will appear displaying all of the audio data from all of the sound files specified in the list file, or data from the first page of the sequence, in the order that the file names appear in the list file.

#### *Format of a list file*

A list file consists of a series of file name entries, one per line. Each entry can be a simple file name, a relative path name, or an absolute path name. If the entry is a simple file name, Raven looks for the file in the same directory as the list file. A relative path name can start with the name of a subdirectory within the list file’s directory, or it can use the double-dot (. .) notation to indicate the parent directory of the list file’s directory. For example, consider a list file that resides in the directory `c:\SoundData` and contains the following four entries:

```
firstFile.wav
moreData\secondFile.wav
..\otherSounds\thirdFile.wav
c:\frogSounds\fourthFile.wav
```

When opening this file sequence, Raven will look for the following files:

```
c:\SoundData\firstFile.wav  
c:\SoundData\moreData\secondFile.wav  
c:\otherSounds\thirdFile.wav  
c:\frogSounds\fourthFile.wav
```



When specifying directories in list files, the forward slash (/) and back slash (\) characters may be used interchangeably.

On a Mac, consider the list file /Applications/RavenPro1.4/Sounds/MyList.txt that contains the following entries:

```
sound.wav  
../examples/CanyonWren.wav  
birds/Chickadee.wav  
/users/Amanda/Sounds/rugbyGrunt.wav
```

When opening this file sequence, Raven will look for the following files:

```
/Applications/RavenPro1.4/Sounds/sound.wav  
/Applications/RavenPro1.4/examples/CanyonWren.wav  
/Applications/RavenPro1.4/Sounds/birds/Chickadee.wav  
/Applications/RavenPro1.4/Sounds/users/Amanda/Sounds/  
rugbyGrunt.wav
```

**Selecting files in a sequence individually**

You can manually specify a set of sound files to open as a sound file sequence by choosing multiple files from within the Open Sound Files dialog box. There are three ways to choose multiple files to open as a sequence:

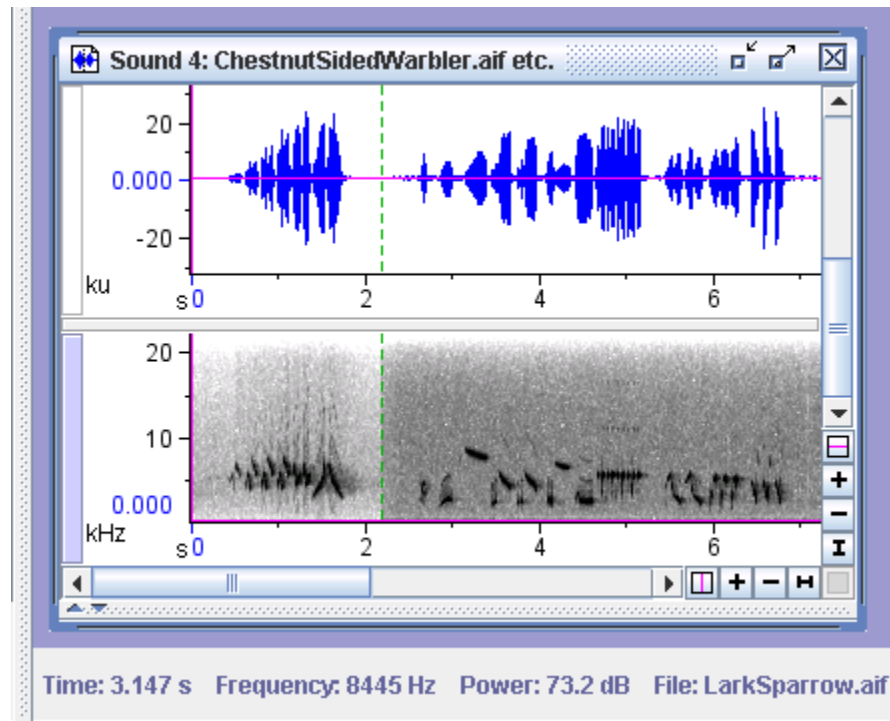
- Click on the name of the first file you want to open in the list of files. Then select additional files to open by clicking on their names while holding down the **Ctrl** key. The names of all of the files you clicked will be highlighted, and will be shown in the **File Name** field. Each file name in the **File Name** field will be enclosed in double quotation marks (Figure 7.7).
- If the names of the files you want to open are listed consecutively (with no intervening file names) in the file list, click on the name of the first file to open, then shift-click on the name of the last file. The names of the two files you clicked and all files between those two will be highlighted, and will be shown in the **File Name** field. Each file name in the **File Name** field will be enclosed in double quotation marks.
- You can type the names of the files that you want to open in the **File Names** field, with each name enclosed in double quotation marks (like "myFile.wav"), and a space between successive quoted file names.

After you click **OK** in the Open Sound Files dialog box, the Configure New Sound Window dialog box appears (Figure 7.8). In the Multiple Files pane

at the bottom of the Window tab, click on Open as file sequence in one window.

After choosing either to open the file sequence in its entirety in a single window or in a paged sound window (see “[Configuring a new paged sound window](#)” on page 188), click OK. A new sound window will appear displaying all of the audio data from all of the sound files specified in the Open Sound Files dialog box, or data from the first page of the file sequence. When Raven opens a sequence of files that were individually specified (as opposed to given in a list file), the order in which data from the different files are displayed is determined by the lexicographic order of the file names.<sup>1</sup>

When data from more than one file in a sequence are visible in a view, Raven displays dashed green *file boundary markers* (Figure 7.6) between the data from each file.



**Figure 7.6.** The files “ChestnutSidedWarbler.aif” and “LarkSparrow.aif” from the Raven Examples folder opened as a file sequence. The dashed green vertical line shows the boundary between the two sound files.

1. In lexicographic ordering, numbers precede letters. For example, if you were to open two files named `abc.wav` and `ab1.wav` as a sequence, the data from `ab1.wav` would be displayed first followed by the data from `abc.wav`.

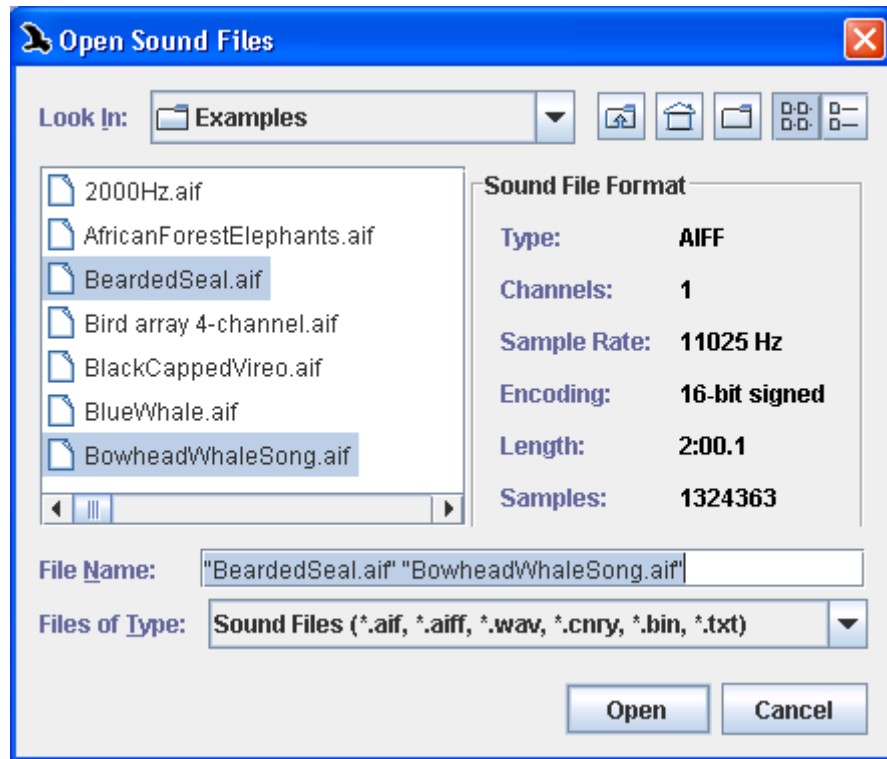
## Opening multiple sound files in separate windows

---

You can open more than one sound file at a time, with each sound appearing in a separate window, by choosing multiple files from within the Open Sound Files dialog. To display the Open Sound Files dialog, choose **File > Open Sound Files....** There are three ways to choose multiple files to open:

- Click on the name of the first file you want to open in the list of files. Then select additional files to open by clicking on their names while holding down the **Ctrl** key. The names of all of the files you clicked will be highlighted, and will be shown in the **File Name** field. Each file name in the **File Name** field will be enclosed in double quotation marks ([Figure 7.7](#)).
- If the names of the files you want to open are listed consecutively (with no intervening file names) in the file list, click on the name of the first file to open, then shift-click on the name of the last file. The names of the two files you clicked and all files between those two will be highlighted, and will be shown in the **File Name** field. Each file name in the **File Name** field will be enclosed in double quotation marks.
- You can type the names of the files that you want to open in the **File Names** field, with each name enclosed in double quotation marks (like `"myFile.wav"`), and a space between successive quoted file names.
- You can click on a `.txt` file that contains a list of file names to open, like those produced by the Raven file sequence recorder. See a description of list files in ["Using a list file"](#) on page 195.

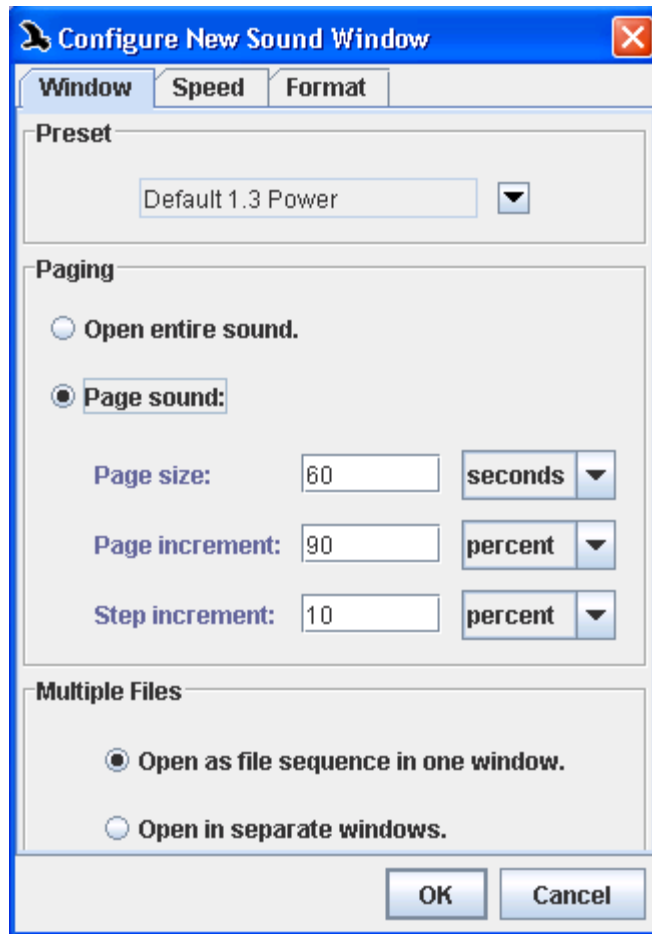
Once you have chosen all of the files you want to open, click **OK**.



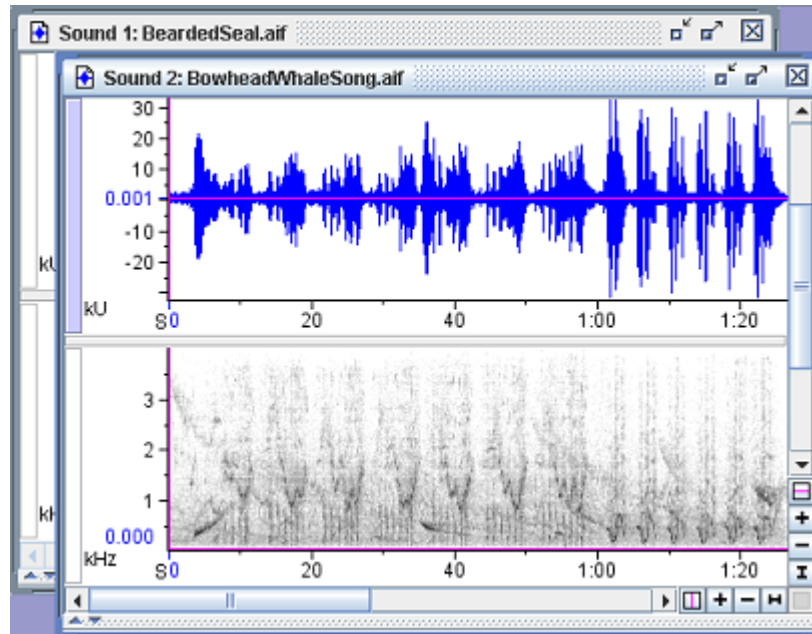
**Figure 7.7.** The Open Sound Files dialog box, with two files selected by control-clicking in the file list.

After you click OK in the Open Sound Files dialog box, the Configure New Sound Window dialog box appears (Figure 7.8). In the Multiple Files pane at the bottom of the Window tab, click on Open in separate windows.

After choosing either to open each file in its entirety in a single window or in a paged sound window (see [“Configuring a new paged sound window”](#) on page 188), click OK. Each of the chosen files will open in a separate window, using the window preset that was specified in the dialog.



**Figure 7.8.** The Configure New Sound Window dialog for opening multiple files.



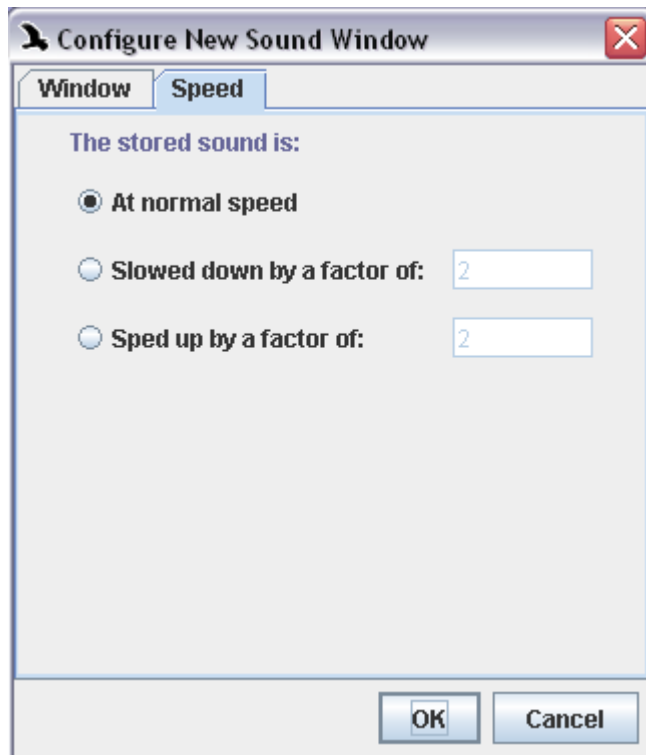
**Figure 7.9.** The sound files selected in [Figure 7.7](#) shown in separate sound windows after selecting “Open in separate windows” in the Configure New Sounds Window dialog.

### Correcting sounds acquired at altered speeds

It is sometimes necessary or useful to acquire sounds from recordings played at speeds slower or faster than the speed at which they were recorded, as discussed in [“Advanced options: acquiring from recordings at altered input speeds”](#) in [Chapter 4](#) (page 99). Sounds that were acquired with the appropriate speed correction in Raven 1.2 and subsequent versions, as discussed in that section, require no further corrections. However, if you have sound files that were acquired at altered input speeds with older versions of Raven or with another program that did not correct for the altered input speed, you can correct the speed when opening the file.

**Correcting speed when opening file** The Speed tab in the Configure New Sound Window dialog box ([Figure 7.10](#)) allows you to specify the factor by which the sound was slowed down or sped up (relative to normal speed) when it was acquired. The Configure New Sound Window dialog box is displayed after you choose one or more files to open in the Open Sound Files dialog box.





**Figure 7.10.** The Speed tab in the Configure New Sound Window dialog box.

If you provide the appropriate speed correction factor when opening the file, then Raven will display time and frequency axes of views, and measurements so that they refer to the original sound at normal speed.



Sounds that are acquired at an altered speed using the appropriate speed correction factor behave in Raven as if they had been acquired at normal speed. Thus, to hear sounds that were originally above or below the human hearing range, you will need to specify an appropriately low or high speed for playback, as described in [“Playback rate” in Chapter 1](#) (page 11).

Raven corrects a file’s speed by altering its sample rate. If you save a file after applying a speed correction, it is saved with the new corrected

sample rate. The next time the file is opened (either by Raven or any other program), its speed will be correct without any adjustment.



In order to save a file with a corrected speed, you must open a single file (not a sequence) in its entirety (not in a paged sound window). For file sequences, you must either apply the correction each time you open the sequence, or else open and save each individual file once with a speed correction. For files that are too large to open in unpagged windows, you must apply the speed correction each time you open the file.