

# Wave Scroller 1.2

## Introduction:

Wave scroller was designed for speedy transcription of audio files and to quickly interface with the recently developed NUaligner<sup>\*</sup> software application. The goal was to design a software application that was simple and easy to use but at the same time allowed users to transcribe quickly without having to take their hands off the keyboard to handle other controls (such as mouse clicking or scrolling).

## Playing an Audio File:

1. Launch the Wave Scroller application.
2. Select File->Import Audio in the menubar. Select an audio file for transcription. Wave Scroller accepts both mono and stereo audio files as well as .wav and .aif audio formats.
3. Once the audio is imported, the screen should look like this:

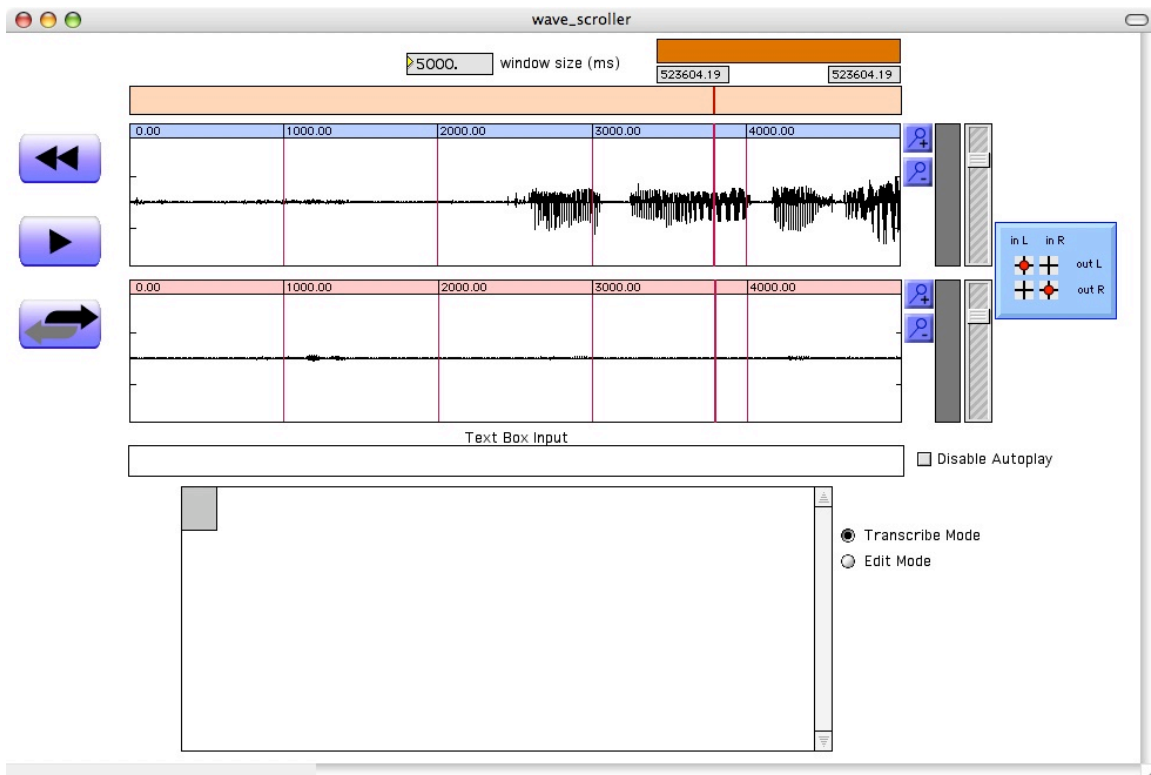


4. Click on the "window size (ms)" number box and type in 5000 and hit enter. This will reduce the audio window size to 5000ms, which is about the ideal window size for most audio files.

---

\* NUaligner is an automatic word aligner that accepts sentence level transcriptions and converts them into word and phoneme aligned transcriptions. NUaligner accepts the text output from Wave Scroller and in turn outputs a praat text grid of the word/phoneme aligned audio files.



- Click and drag on the control bar (light orange bar above the audio window) to control the static playhead position. Move the static playhead about two thirds to the right. Your screen should now look something like this.



- The top half of the audio display represents the left audio channel and the bottom half represents the right audio channel. You can adjust the channel volumes individually using the volume sliders.



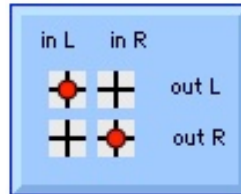
The level meter to the left of the volume slider gives you a visual representation of the volume level.

- Click the rewind button  to set the play position to zero.
- Click the play  button to begin playback of audio. The play button functions as a toggle so clicking on it again causes the audio to pause.
- Use the Zoom buttons to adjust the Y-axis scale.

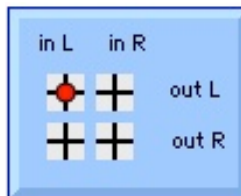


### Mixer Panel

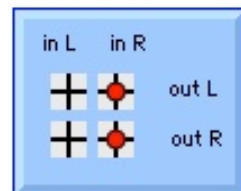
1. The mixer panel allows you to route the left and right channels of audio to any number of combinations over your headphones or loudspeakers.
2. The default setting is always left channel audio->left speaker, right channel audio->right speaker.



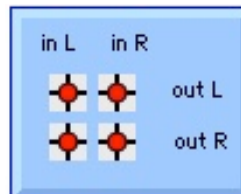
3. The vertical columns represent the Left and Right channels of the current audio file.
4. The horizontal rows represent the Left and Right channels of your output device (i.e. headphones or loudspeakers).
5. The red dots can be clicked on/off and they toggle the appropriate channels and routes.
6. For example, in the following setting, only the left channel audio can be heard through the left speaker:



7. In this second example, only the right channel audio will be heard through both left and right speakers.



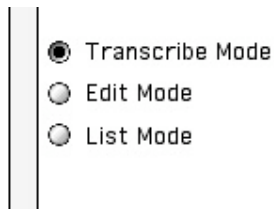
8. Finally, with this setting, both left and right audio channels will be heard through both left and right speakers – essentially a mono setup.



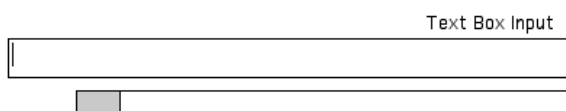
9. Routing combinations and are not restricted to these examples alone.


## Orthographic Transcription:

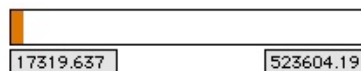
1. Select the Transcribe Mode radio button.



2. Click on the text box input to give it focus. A blinking cursor should appear in the text box indicating that it is ready for text input.



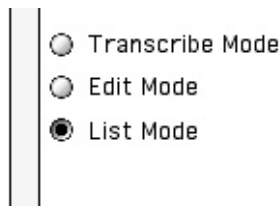
3. With the text box in focus, the enter button doubles as a play/pause toggle. Hit enter to begin audio playback.
4. Hit enter to pause the audio at the end of a sentence.
5. To make small adjustments to the playhead position, hit the tab key to de-focus the text box input and then use the L and R arrow keys to adjust the playhead position. Hit tab to bring the text box input back into focus.
6. Type in the transcription text of the last sentence heard.
7. If you need to listen to the last sentence again, click the  button or use the tab key to de-focus the text box and then hit the 'R' key to activate the replay button.
8. Once you have completed the transcription for the last sentence, hit enter.
9. This will store the transcription text in the transcription table underneath the text box input. Playback begins automatically each time you hit return.
10. Repeat steps 6 to 9 for the rest of the audio file. Ideally each sentence segment should be about 4-6 seconds long.
11. The progress bar on the top right of the screen indicates the current position in ms, the total file length in ms and a visual indication of the current playhead position within the audio file.



12. After completing the sentence transcription of the audio file, select File->Export from the menubar to save the transcription text file.
13. The exported text file can now be used to do word and phonetic alignment with the NUsaligner software application.

### Orthographic Transcription from a list

1. If your audio file contains speech that is read from a fixed script, you can use the “List Mode” to pull in the script from a simple text file.
2. Orthographic transcription from a list can only be done on a single channel (mono) audio file.
3. Select File->Import Audio and import a mono audio file that you wish to transcribe.
4. Set the window size to 5000ms and place the playhead about two-thirds to the right of the orange control bar.
5. Select the List Mode radio button

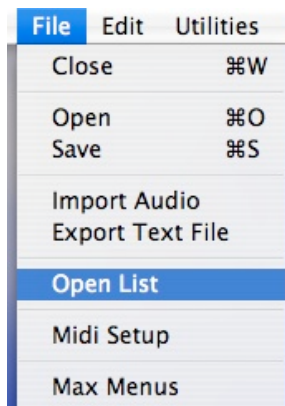


6. Create a text file containing the list of sentences to be transcribed. Use the following format:  
    <index>, "<sentence>";

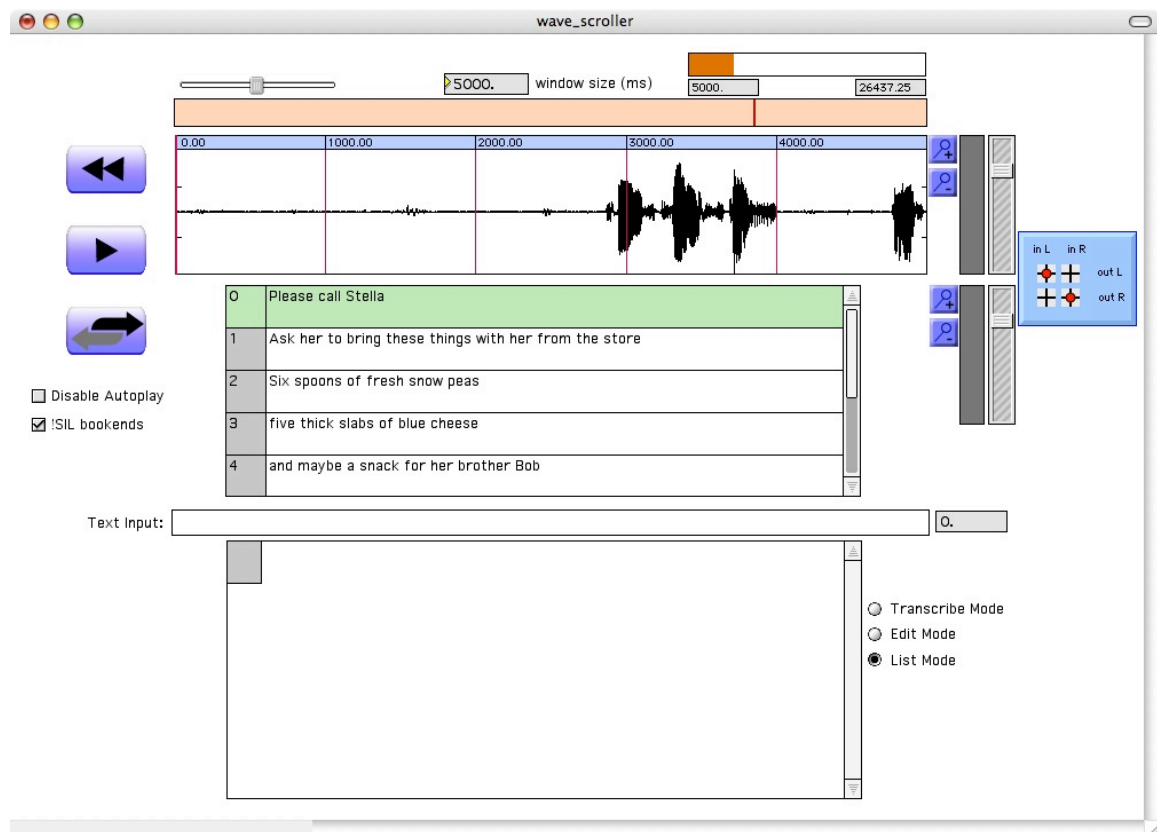
Example:

- 0, "Please call Stella";
- 1, "Ask her to bring these things with her from the store";
- 2, "Six spoons of fresh snow peas";
- 3, "five thick slabs of blue cheese";

7. Select File->Open List from the file menu and open the list of sentences you just created.



8. The screen should now look something like this:



9. Select the Text Input box to give it focus.
10. Hit enter/return to play through the first segment and then hit enter/return again to pause.
11. For example, if the first audio segment was "Please call Stella", then as soon as you hear the appropriate audio, hit the return key to pause the playhead.
12. Next use the up/down arrow keys to scroll through the list. As you scroll through the list, the text appears in the text input box.
13. Select the appropriate text for the audio segment ("Please call Stella") and then hit enter/return. This will store the text and the start/end time of the segment in the transcription table and will also resume audio playback.
14. Repeat steps 9 through 10.

#### Saving a Wave Scroller session

1. When transcribing long audio files, it may be desirable to save a Wave Scroller session.
2. A Wave Scroller session is saved as a .ws file and must be in the same folder as the imported audio file. Wave scroller session files can be saved with any name but must maintain the .ws suffix.

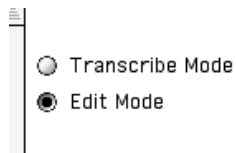
3. If transcribing using the “List Mode”, then the text file containing the list of sentences must also be in the same folder as the .ws file.
4. Check the “disable autoplay” checkbox.
5. Transcribe the last sentence and then hit enter. This stores the text transcription but does not resume audio playback. You are now ready to save your Wave Scroller session.
6. Select File->Save from the file menu. Browse to the folder on your hard drive that contains the imported audio and save the wave scroller session in the same folder.

### Resuming a Wave Scroller session

1. Select File->Open
2. Browse and open the appropriate .ws file.

### Edit Mode:

1. Select the edit mode radio button.



2. Click on any of the rows on the transcription table. This will select the appropriate audio segment in blue.
3. At the same time, the corresponding text will display in text box input, which you can edit. Hit return to store any changes made.
4. Click the replay button to playback the selected audio segment.
5. Click and drag up or down on the number box to the right of the text box input to edit the end-time value of the selected row. You can also type into the number box.

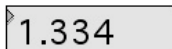


Diagram 1 - Number box

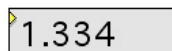


Diagram 2 - Number box selected, small triangle on the left turns yellow. Hint: You can type in values when the number box is selected.

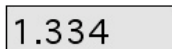


Diagram 3 – Number box without a triangle on the left means it cannot be edited

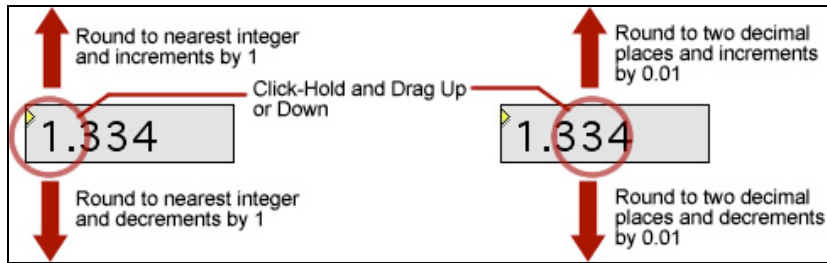


Diagram 4 – Number Box mouse interaction

- Click and drag the green slider left or right to manually scroll left or right through the waveform in edit mode. The slider is disabled in transcribe mode.



### Saving as plain text file

- Select File->Export from the menubar to save the transcription as a plain text file. The saved text file can then be used with NUaligner to create word and/or phoneme aligned Praat text grids

### Additional Features

- Disable Autoplay
  - If the disable autoplay checkbox is checked, the “enter” key no longer doubles as a play/pause button.
- !SIL bookends
  - Check this option to automatically insert “!SIL” at the beginning and the end of each segment as it is typed.
  - !SIL is the default transcription for silences.
- Use Shift-\ (shift-back slash) to automatically insert !SIL in the text input box.

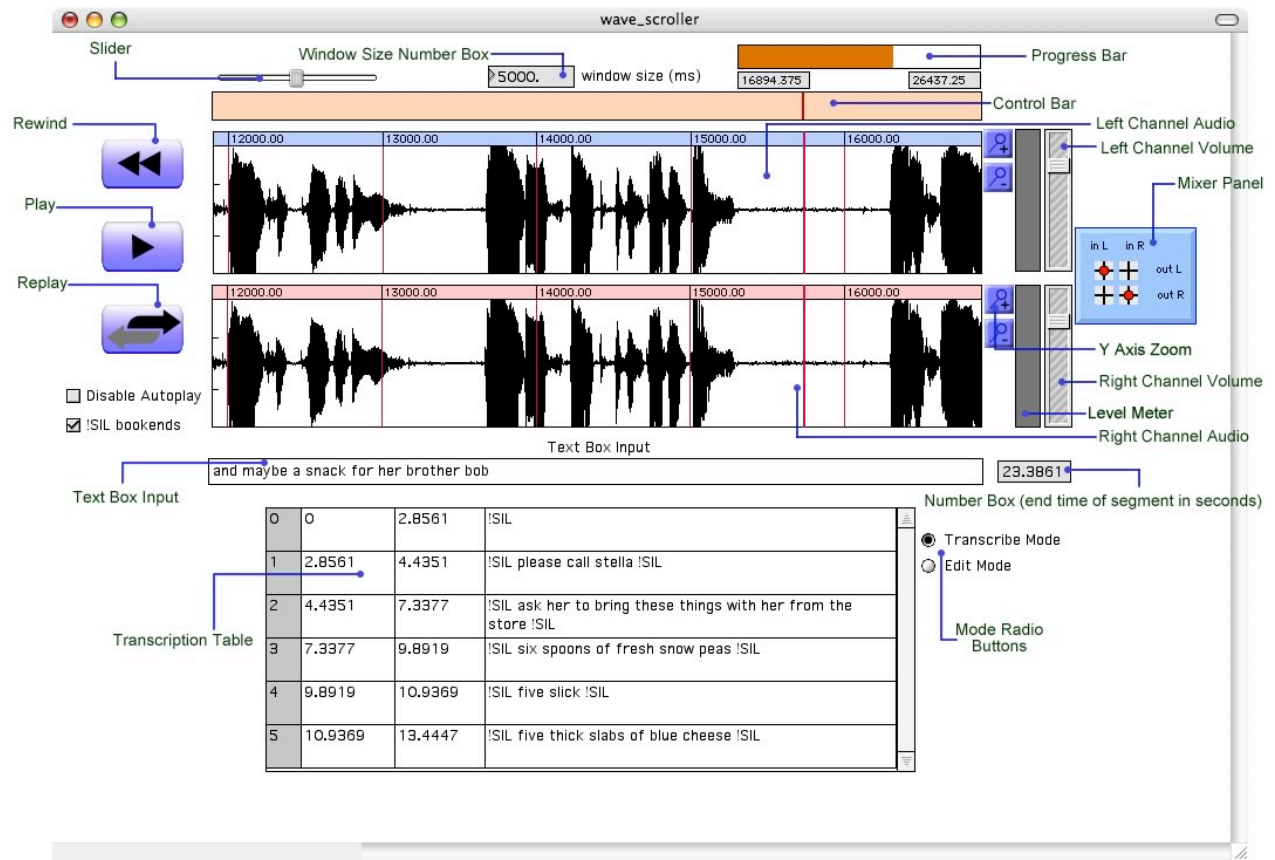
### Known Issues

- The enter button stores the text box input only when there is text in the text input box. If there is no text, the enter button functions only as a play/pause button (except when the “disable autoplay” option is checked). Eventually a pedal interface will be combined into the application to allow the transcribers to access the play and replay buttons with the foot and will free up the enter button.
- There is currently no way to delete any transcription entries that have been stored in the transcription table made.



## Glossary

The following diagram outlines the names of each of the various components that make up Wave Scroller.



## Troubleshooting Audio

If you are not getting any audio from Wave Scroller or the playhead is not moving, it's possible the audio devices need to be reset.

1. Select File->Audio Setup
2. The following window should appear:

The screenshot shows the 'DSP Status' window with the following settings:

- Audio:** Off
- Driver:** ad\_directsound
- Input Device:** Creative SB Live! Series
- Output Device:** Creative SB Live! Series
- Thread Priority:** Medium
- Latency (ms):** 80
- CPU Utilization:** 0. % ☒ Poll ☐ Update
- Function Calls:** 6
- Signals Used:** 3
- Sampling Rate:** 44100 Hz ☐ Override
- Input Channels:** 2
- Output Channels:** 2
- I/O Vector Size:** 512 ☐ Override
- Signal Vector Size:** 512 ☐ Override
- Max Scheduler in Overdrive:** Off ☐ Override
- Scheduler in Audio Interrupt:** Off ☐ Override
- Input Channel 1:** 1 Left Input ☐ Override
- Input Channel 2:** 2 Right Input ☐ Override
- Output Channel 1:** 1 Left Output ☐ Override
- Output Channel 2:** 2 Right Output ☐ Override
- Optimize:** On ☐ Override
- CPU Limit:** 0 % ☐ Over ☐ Override

Buttons at the bottom: Audio Driver Setup, I/O Mappings

3. In the "Driver" menu, select "ad\_directsound" (windows) or core audio (mac).
4. Under input and output devices, select your computers default audio device.  
In most cases, the audio is deactivated due to the fact that no input/output device was selected.
5. In the "Thread Priority" menu, select "Medium"
6. In the "Latency" menu, select "80".
7. In the "Sampling" rate menu, select "44100"
8. For "I/O Vector Size" and "Signal Vector Size", select "512". Another typical problem is that these numbers are mismatched. Both these values must be the same for the audio to function correctly.
9. If the audio still sounds choppy, try increasing the "Latency" values.