Welcome to the Soundskrit Cricket Kit

Figure 1: Soundskrit Cricket Kit

The Cricket kit is an easy-to-use evaluation platform designed to demonstrate Soundskrit's beamforming, speech isolation, and noise reduction algorithms, optimized for speech-to-text applications in smart glasses. The kit includes a glasses prototype equipped with two orthogonal Soundskrit dipole microphones and an omnidirectional microphone, Soundskrit's PARDI audio interface board, and a Windows-based GUI.

The Cricket uses the microphone signals to isolate the wearer's voice from the voice of a person standing in front of them. These two isolated speech streams can then be fed into a speech-to-text engine, translation software, voice call application, or other signal processing chains. This demonstration uses the Whisper speech-to-text engine from OpenAI for transcription, leveraging the high-quality audio separation provided by Soundskrit.

What's In the Box	
Cricket with audio interface, mounted on glasses	Glasses with attached microphones
Soundskrit PARDI audio interface	USB audio interface to connect microphones to PC
Microphone cable	Cable to connect the Cricket PCB to the audio interface
USB A to USB C Cable	Cable to connect the Soundskrit audio interface to your computer

Table of Contents

System Overview

The Cricket glasses prototype connects to Soundskrit's custom PARDI audio interface, which in turn connects via USB to a Windows PC. On the PC, the Soundskrit GUI processes the incoming microphone signals to isolate the two distinct voices (wearer and front) and reduce background noise. The two processed audio output streams from the GUI can be recorded using software (e.g., Audacity), sent to a speech-to-text engine, or used in a voice call application.

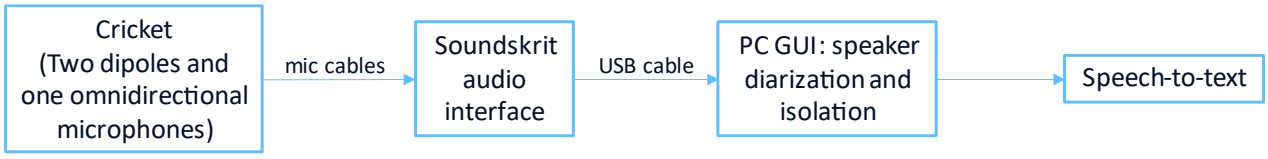


Figure 2: System overview of the demo kit.

Quick Start Guide

Installation

To get started, connect the prototype to your PC using the included USB-C to USB-A cable.

Launch the installer file ‘Installer_Cricket_Soundskrit.msi’ which can be found at Demo kits and samples - Soundskrit and follow the installation procedure. You may have to approve installation from an unknown publisher.

To start the software, find ‘Soundskrit GUI Cricket’ in the start menu or use the desktop shortcut generated by the installer.

Initial Configuration

In the GUI, select Options > Audio/MIDI Settings... from the top-left menu:

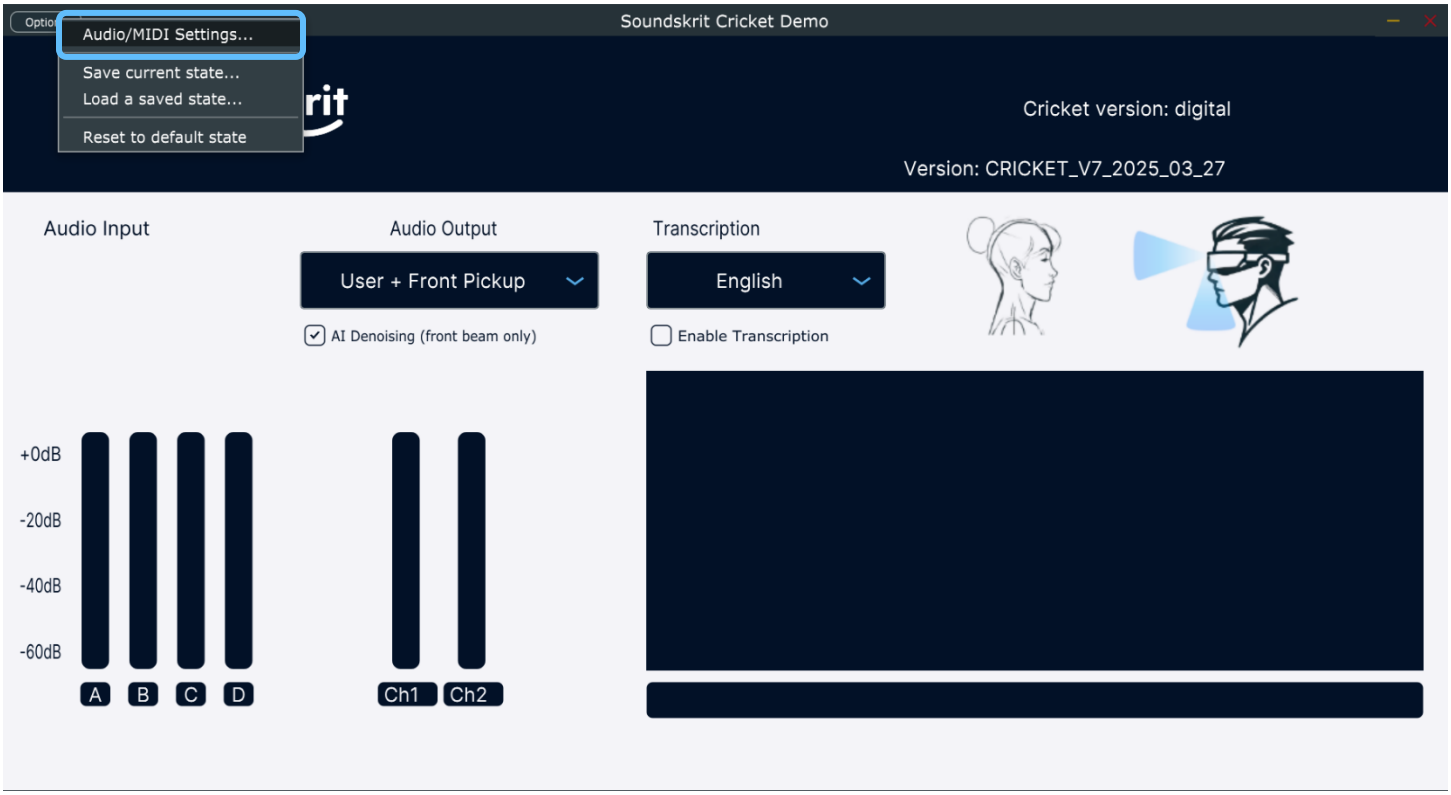


Figure 3: Accessing Audio/MIDI Settings via the Options Menu

In the options menu, ensure that the following options are selected:

Option	Selection
Feedback loop	Mute audio input should not be selected
Audio device Type	Windows Audio
Output	Speakers (…Soundskrit Raw)
Input	Microphone (…Soundskrit Raw)
Active Input Channels	Input channel 1 + 2 and Input channel 3 + 4
Sample Rate	48000 Hz
Audio buffer size	480 samples (10.0) ms
Active MIDI Inputs	Leave empty

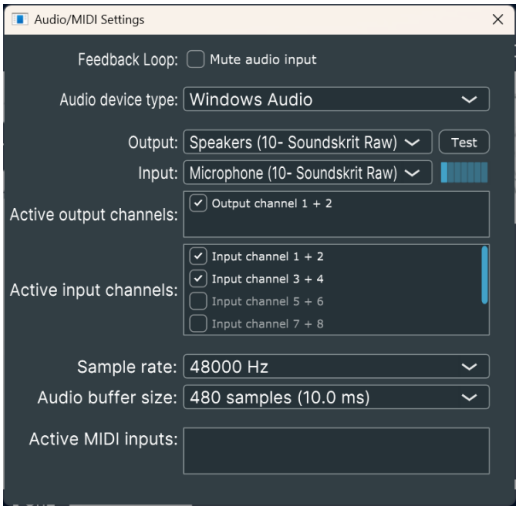
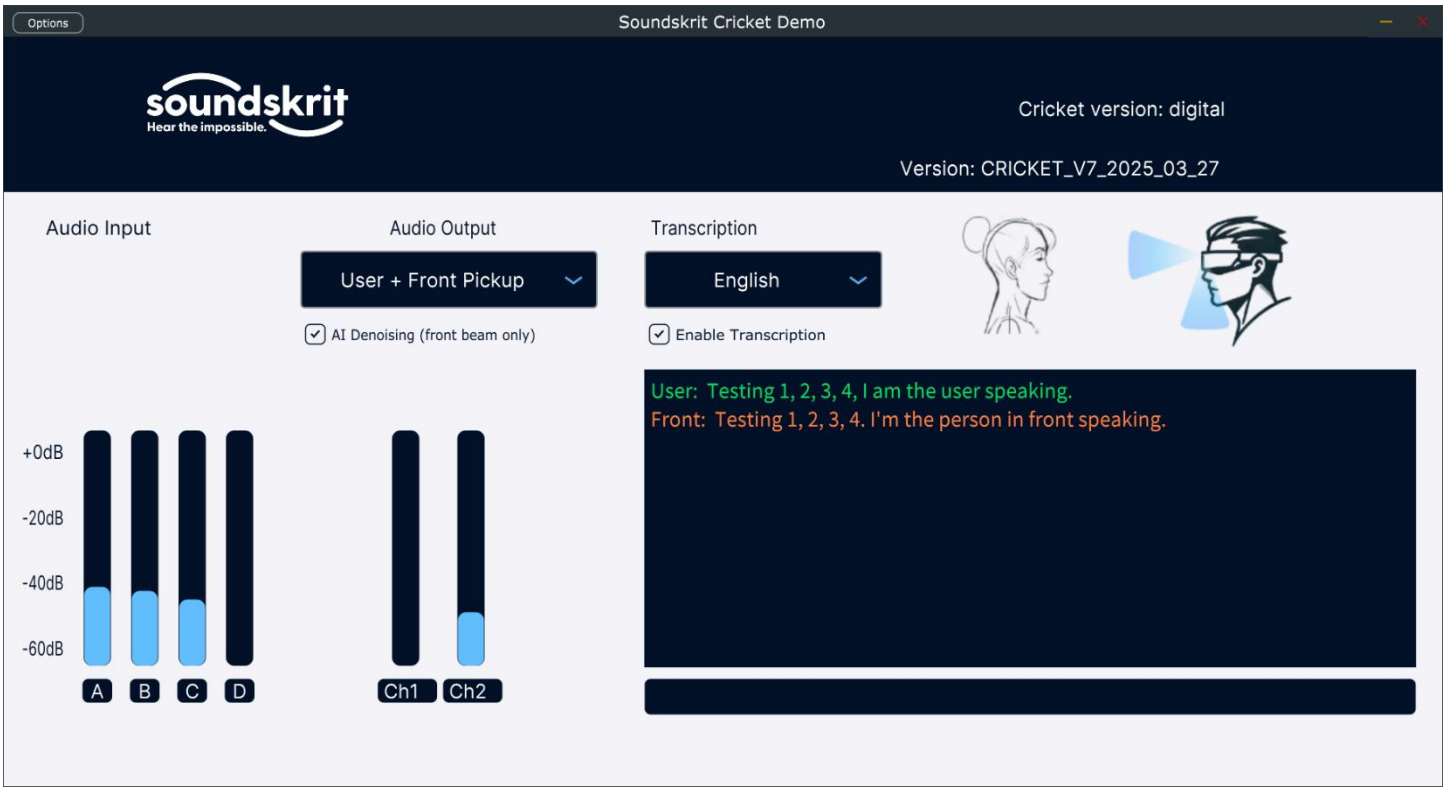


Figure 4: Properly configured options menu

Using the Transcription

Click on “Enable Transcription.” This will begin transcribing the speech of the user (wearer) and the person in front.



Soundskrit processes the microphone signals and separates the distinct audio sources.

Understanding the Cricket Kit

Hardware

The Soundskrit Cricket consists of glasses with a PCB clipped to the left side of the glasses. The PCB holds two Soundskrit dipole microphones, one facing forward (Y) and one facing downward (Z), and an omnidirectional microphone (W).

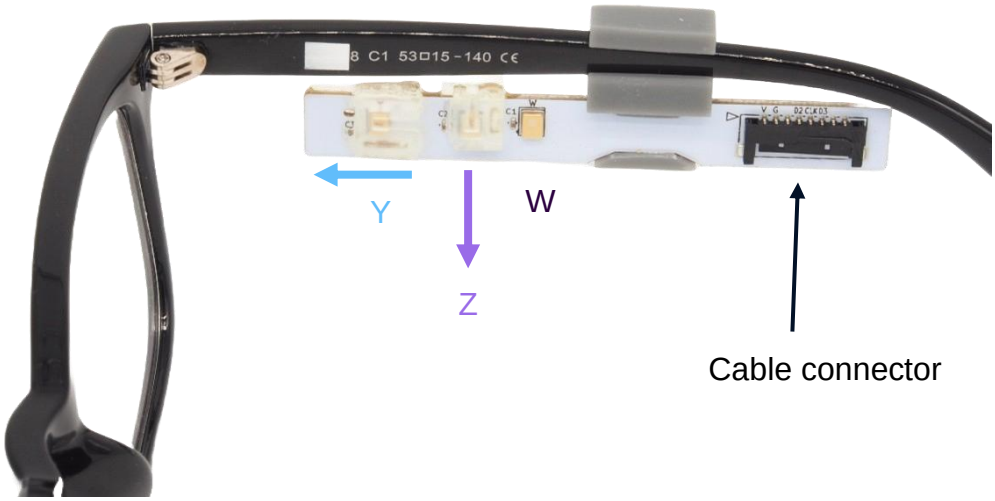


Figure 5: The PCB holds a forward-facing dipole (Y), a downward-facing dipole (Z), and an omnidirectional microphone (W). The connector for the cable to the PARDI audio interface is located at the back. (Note: Although PCBs may be present on both sides, only the left side is active for this demo).

These three microphones are used to create beamformers directed toward the mouth of the wearer, and one facing forward to pick up speech from someone standing in front of the wearer.

Figure 6: Cricket glasses with microphone PCBs on each side. Only the left side is used. The arrows illustrate the directional beams processed by the GUI for capturing the user's voice (towards the mouth) and front speech (forwards).

The PCB is connected with a Molex connector to the Soundskrit PARDI audio interface.

Figure 7: Connection of the Soundskrit multichannel audio interface.

The audio interface connects via a USB-C cable to a PC. The interface acts as a standard USB audio device and sends the raw microphone signals to the CPU to be processed by the GUI.

Cricket GUI In-Depth

The main GUI layout contains two drop downs and level meters. In the bottom left in the Audio Input section, there are levels for each of the microphones built into the prototype (There are four bars where the last one is unused). The blue levels should move if you clap or make noise. In the GUI Audio Input section: Level A corresponds to the forward dipole (Z), Level B to the downward dipole (Y), Level C to the omnidirectional microphone (W), and Level D is unused.

The Audio Output section has two level bars labeled Ch1 and Ch2. Ch1 is the user voice and Ch2 is the front voice. Speaking from each direction will make these bars move. There is a checkbox to denoise the front pickup for better transcription in noisy environments.

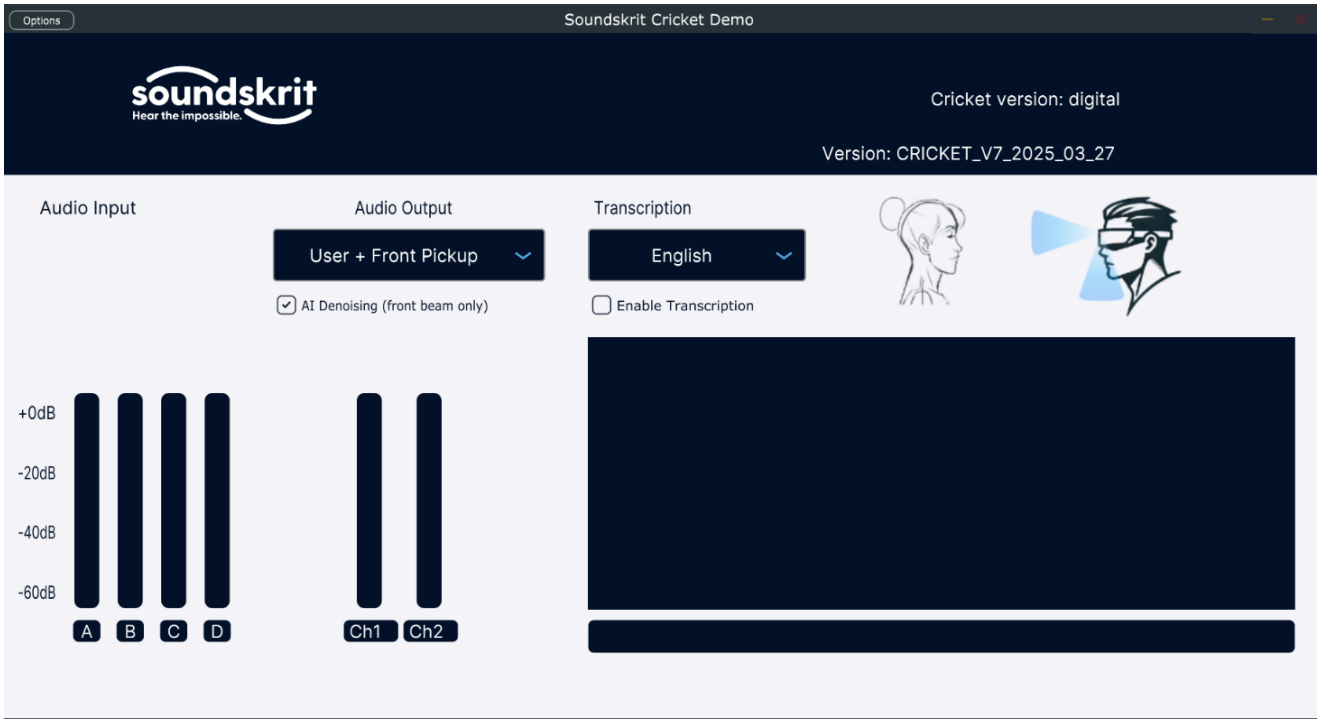


Figure 8: Soundskrit Cricket GUI showing input levels (Z, Y, W) and processed output levels (Ch1, Ch2)

The transcription section provides a drop-down menu to select the language being spoken. The illustration lights up to show the source of sound. Transcription uses the Whisper model by OpenAI. This is a completely offline transcription model. Soundskrit uses its directional microphones and beamforming technology to properly separate the front and the user.

Processing Modes

The Audio Output drop-down menu offers three processing modes: User + Front Pickup, User Pickup, and Front Pickup. These processing modes can be fed to an end application such as a translation device or processed even further.

- User + Front Pickup: This is the default mode it separates the user's voice and the conversation partner's voices.
- User: The user mode feeds the user's voice to both channels of the output. This mode can be used when only the wearer's voice is required. The high level of directionality allows speech from the wearer to be very clear even in high-noise environments.
- Front: The front pickup feeds the voice to both channels of the output. This mode can be used when we only want the front user. The high level of directionality and noise reduction allows speech from the front to be very clear even in high-noise environments.

Recording Audio with the Cricket

The processed audio can be captured for analysis or for other uses. There are many software options which can be used, for simplicity, we recommend using Audacity. Audacity is a trusted, free to use, multiplatform suite of tools for recording and working with audio files.

Once you have installed Audacity, we need to configure the software for use with Soundskrit's audio interface. Configure the settings as listed below:

Audio Host	Windows WASAPI
Input[1]	*Speakers (…Soundskrit Raw) (loopback)* – Ensure the loopback option is selected
Output	Your listening device
Channels	2 (Stereo) Recording Channels

[1]Audacity typically provides two options for audio input devices: a direct input and a 'loopback' input. Selecting the 'Loopback' option is crucial here; it allows Audacity to capture the processed audio stream *output by the Soundskrit GUI*, rather than capturing the raw microphone signals directly.

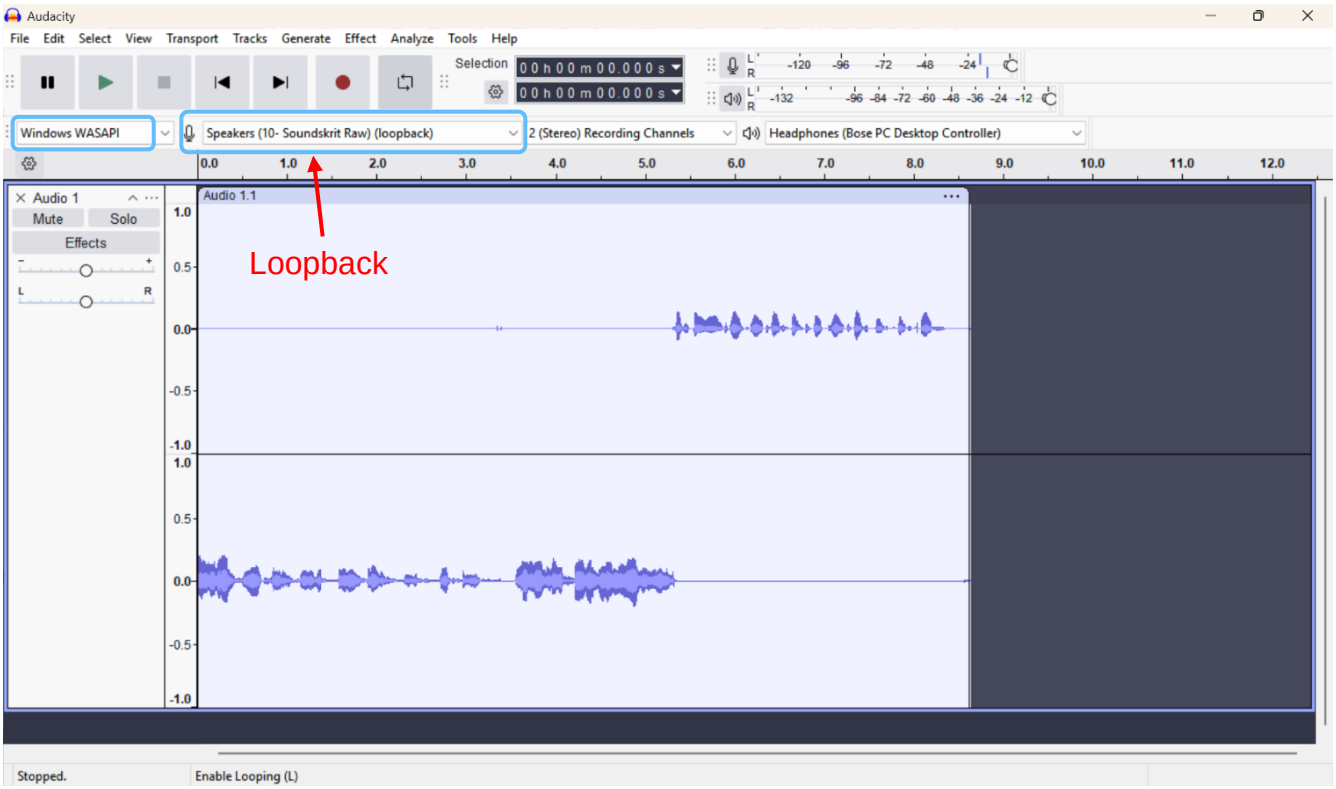


Figure 9: Properly Configured Audacity

To take a recording, press the *record* button. Press the *stop* button to stop.

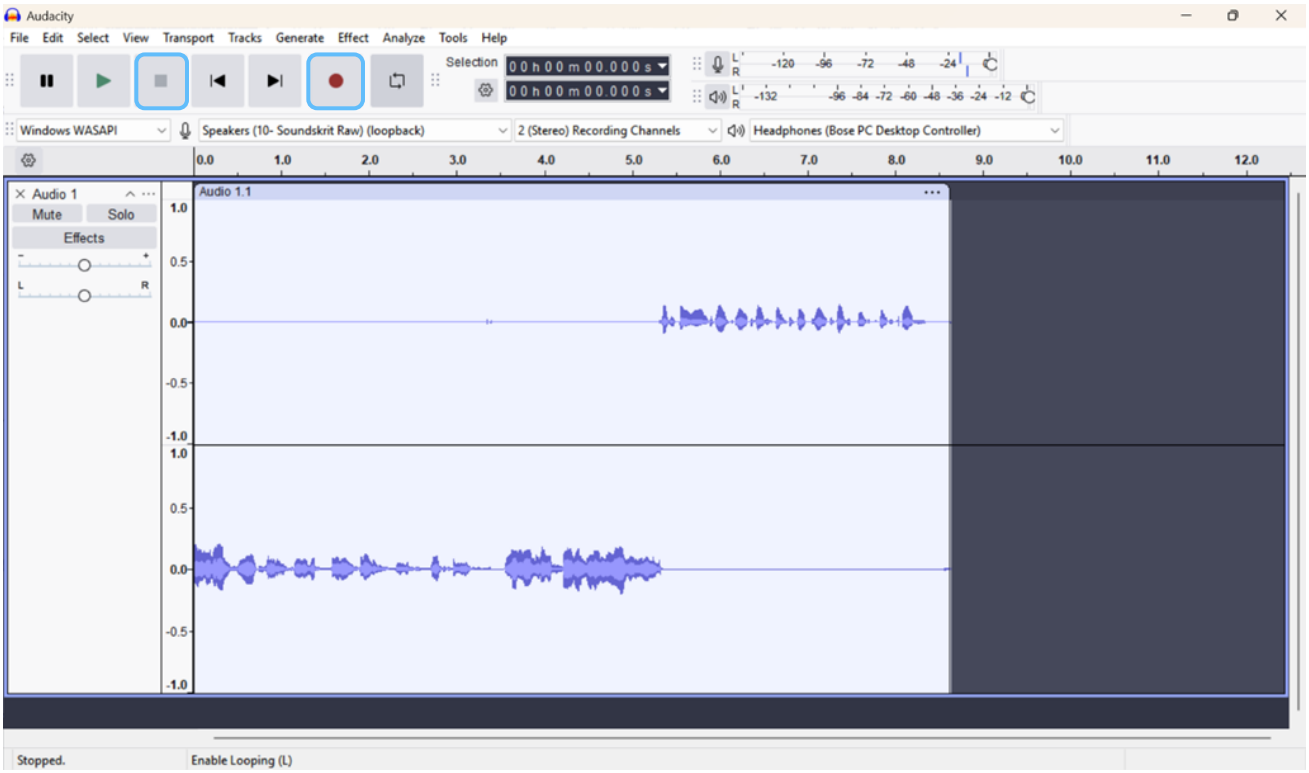


Figure 10: Recording in Audacity.

To listen to only one of the two signals, we need to split the stereo track to listen to them independently. To do this, right-click the track control panel (on the left side of the waveform) and select Split Stereo to Mono.

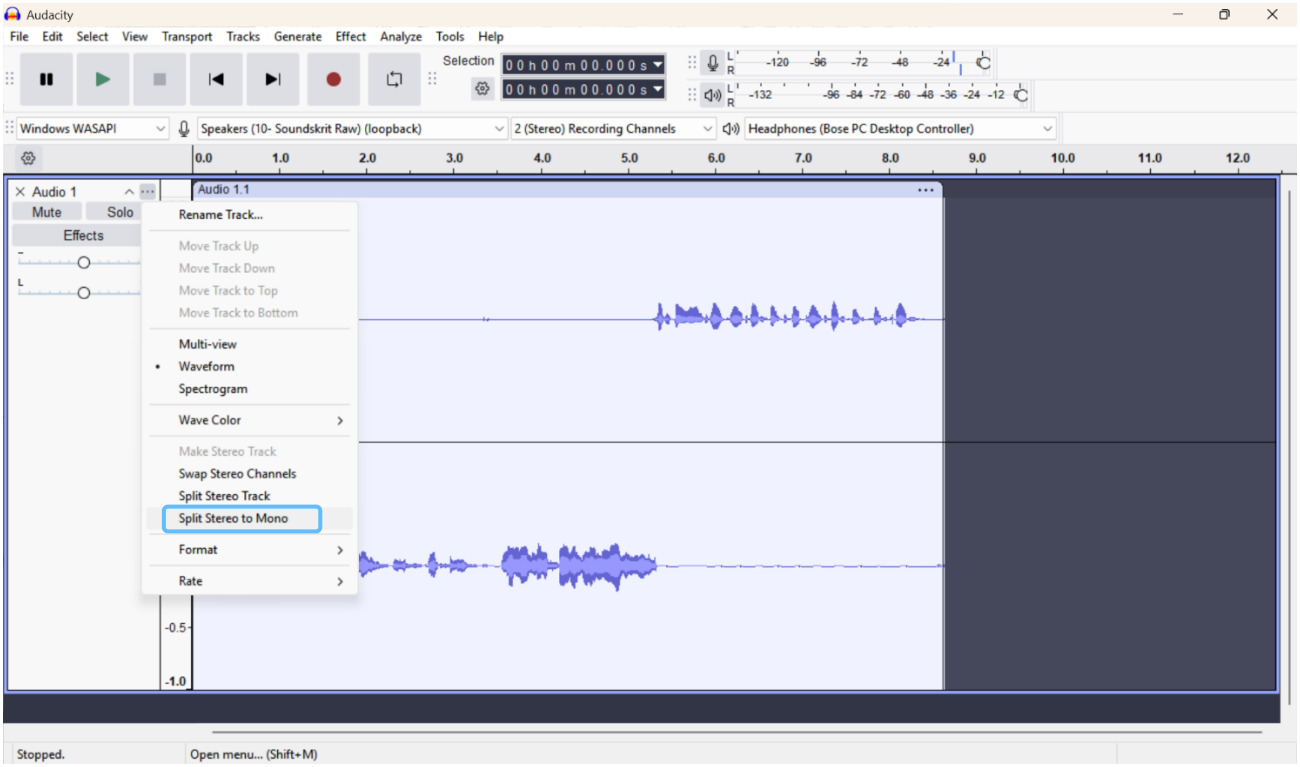


Figure 11: Splitting stereo to mono.

Once you have split the stereo track, you can listen to either track by selecting *solo* and pressing the space bar or play button. In Audacity, the top track corresponds to the User voice (Ch1 from the Gui) in our software and the bottom track to Front voice capture (Ch2 from the GUI).

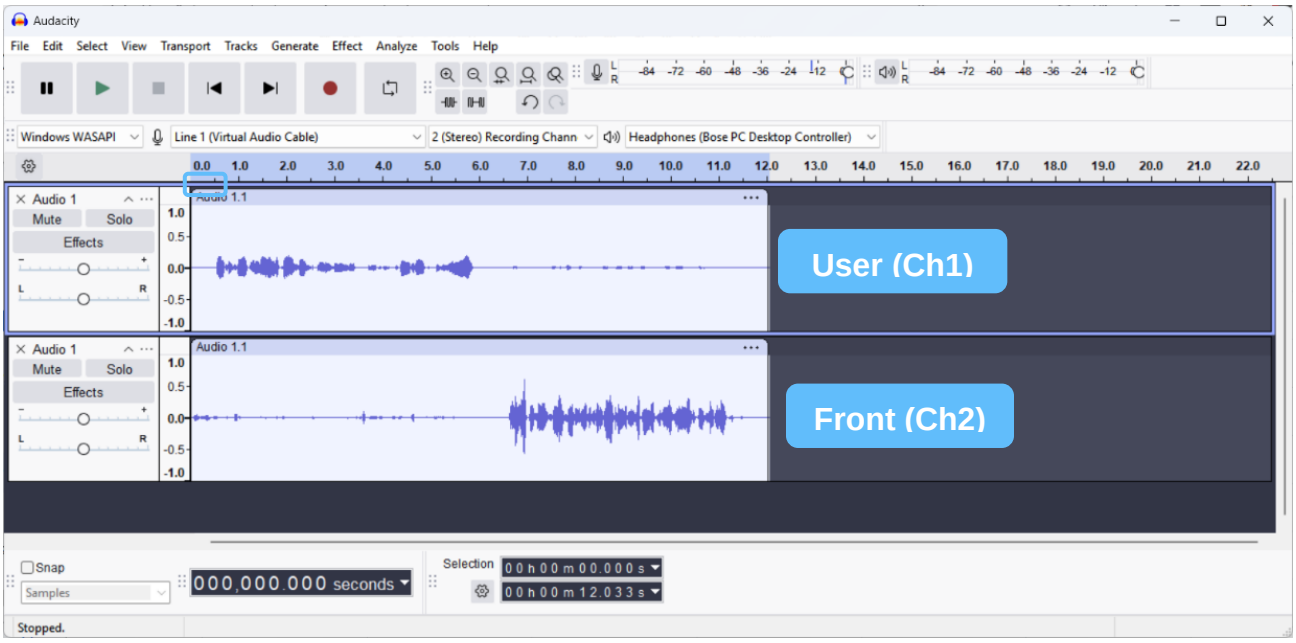


Figure 12: Playing back individual tracks.

Additional Support

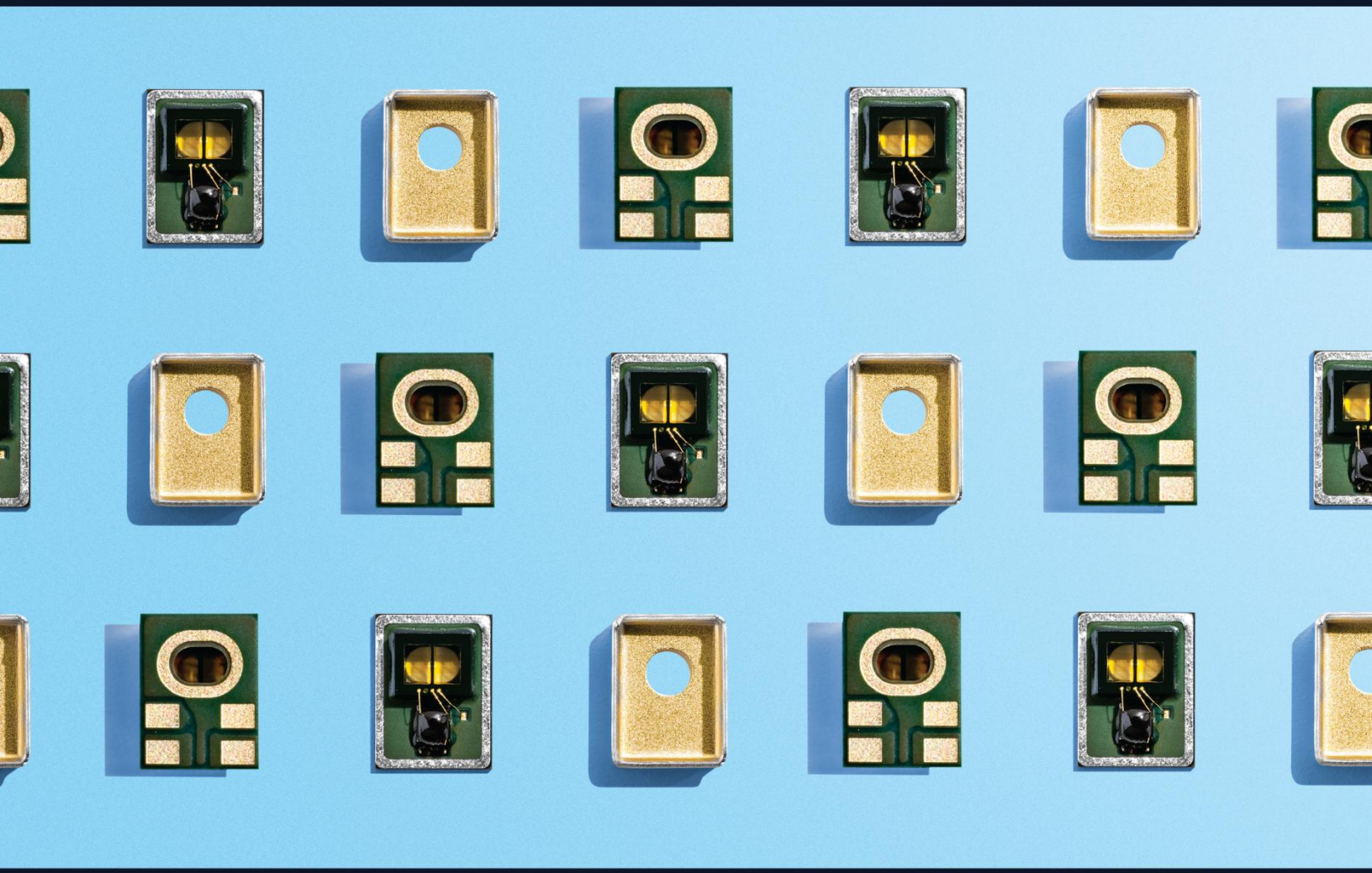
For further information on Soundskrit's products, visit our website at http://www.soundskrit.ca where you can find more application notes, datasheets, and purchasing information. If you have any questions or need technical support, please reach out to applications@soundskrit.ca.

Revision History

Revision Label	Revision Date	Sections Revised
-	January 2025	Initial release
A	April 2025	Update features, screenshots, restructuring
B	April 2025	Updated photos

Soundskrit developed the first high-performance directional MEMS microphone on the market, leveraging years of research in bio-inspired MEMS based on how spiders and other insects in nature hear. In combination with Soundskrit's in-house audio processing algorithms, directional microphones can be used to capture and isolate any sound in an environment with a fraction of the size, power, and computation of traditional omnidirectional-based microphone arrays.

Soundskrit was founded in 2019 and is headquartered in Montreal, Quebec with an R&D facility in Ann Arbor, Michigan.



soundskrit
Hear the impossible.