VEIRA

VOCAL REMOVER

VEIRA Vocal Remover

User Manual

Take the vocal out of a finished song, keep only the vocal, or move it up or down in the mix. An AI model splits the song into vocal and music on your own computer, and two knobs set how loud each comes back.



Contents

Welcome 3
Getting started 4
What you need 4
Install with VEIRA Manager 4
Activation and the licence 5
Formats and where the files go 5
The panel at a glance 6
Using the controls 7
Quick start: a karaoke track or an a cappella in four moves 8
How vocal separation works 9
Why a mixed song is hard to take apart 9
Listening by pitch and time 9
The AI model 10
What happens inside VEIRA Vocal Remover 10
Why it needs more than a second of latency 11
Singing and speech 12
Controls 13
LEVELS: the VOCAL and MUSIC knobs 13
SEPARATION: STRENGTH and the VOCAL RANGE 16
OUTPUT 19
Displays 20
The SEPARATION display 20
The status panel 21
Meter bridge (IN | OUT) 22
Latency and your DAW 24
Playback, mixing and bouncing 24
Recording and live monitoring 24
Pro Tools 24
Other hosts and setups 24
Presets 25
The preset browser 25
The preset menu 25
Saving your own presets 25
Presets in your DAW 26
The factory folders 26
Recipes 28
A karaoke track 28
An a cappella for a remix 28
Turn a vocal up or down in a finished mix 28
A practice track for your instrument 28
Lower the music under a voice-over 29
Effects: telephone, radio, spotlight 29
Specifications 30
Troubleshooting and questions 31

Welcome

Thank you for choosing **VEIRA Vocal Remover**. It works on a finished, mixed stereo song, the kind of file you can't open up into its tracks, and splits it into two parts: the **vocal** and **everything else** (the music). Two knobs, VOCAL and MUSIC, then set how loud each part comes back.

Pull VOCAL all the way down and you have the song without its singer: a karaoke track, an instrumental for a DJ set, a play-along for practice. Pull MUSIC down instead and you have the voice on its own, an *a cappella* for a remix or for study. Anywhere in between, you rebalance a song that was mixed without you: a vocal a few dB louder or quieter, a band dropped under a voice.

The separation is done by an **AI model**, a neural network trained on a large collection of songs to hear what is voice and what isn't. It runs inside the plugin, on your computer, in real time. Nothing is uploaded and the plugin never goes online.

What makes it different

- **Runs locally, in any DAW.** It's an ordinary insert plugin: put it on any track, play, and the separation happens there and then. No upload, no waiting for a web service, no internet.
- **Levels, not a switch.** You don't just choose *remove* or *keep*: each part has its own knob from Off to +6 dB, so the same plugin makes karaoke tracks, a cappellas and subtle vocal rides.
- **Untouched when you don't use it.** With both knobs at 0 dB, the output is the input bit for bit, only delayed. Everything the AI does goes into the difference you ask for.
- **A limit on what counts as voice.** The VOCAL RANGE (LOW and HIGH) keeps the kick, the bass and the cymbals' air out of the separation entirely, so they stay as they were whatever the AI thinks.
- **You see what it hears.** The **SEPARATION** display draws the song's last eight seconds or so as a scrolling picture of its pitches: the vocal in amber, the music in blue.
- **A clean, legible face.** Large knobs with the value in their middle, text you can read at a glance, and a window you can scale from 75 % to 200 %.
- **45 measured presets**, sorted by job: karaoke, a cappella, vocal balance, practice, remix and sampling, speech and creative effects. Each was measured on test songs whose parts are known.

How this manual is organised. *Getting started* covers installing and activating. *The panel at a glance* names every part of the window. *How vocal separation works* explains, without maths, how a finished song can be split at all, what the AI does and why the plugin needs more than a second of latency; read it before you judge the results. *Controls* explains every knob and button with drawings of what it does, *Displays* teaches you to read the SEPARATION display, the level bars and the status panel, and *Latency and your DAW*, *Presets*, *Recipes*, *Specifications* and *Troubleshooting* follow.

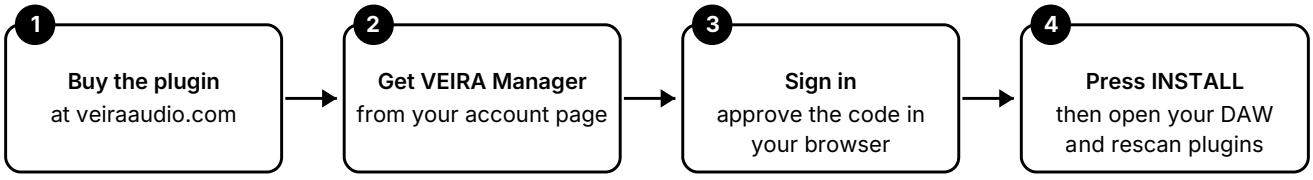
Getting started

What you need

macOS	macOS 11 Big Sur or later. Native on Apple silicon and Intel Macs (universal).
Windows	Windows 10 or 11, 64-bit.
Plugin formats	VST3 (macOS and Windows), Audio Units (macOS), and a standalone app. Pro Tools (AAX) versions arrive once they are signed for Pro Tools.
Hosts	Any DAW that loads VST3 or Audio Units: Logic Pro, Cubase, Nuendo, Ableton Live, Studio One, Reaper, FL Studio, Bitwig and others.
Account	A VEIRA account at veiraaudio.com, and the free **VEIRA Manager** app.
Internet	Needed to install and to refresh the licence. The plugin itself never goes online.

Install with VEIRA Manager

VEIRA plugins are not downloaded from the website one by one. You install them with **VEIRA Manager**, the free VEIRA app that installs, updates and activates everything you own. Its own manual covers it in full; the short version is:



One account works on up to three computers. Updates are free for life.

FIGURE 1 From purchase to first sound in four steps.

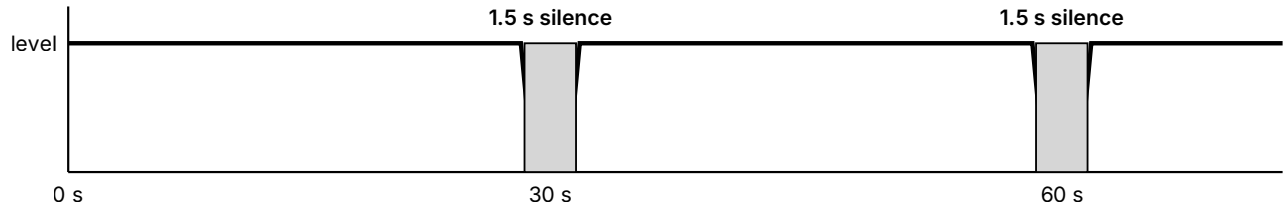
1. **Buy** VEIRA Vocal Remover at **veiraaudio.com**. It's a one-time purchase: you own it for good, with free updates.
2. **Download VEIRA Manager** from your account page (or veiraaudio.com/manager) and install it.
3. **Sign in.** VEIRA Manager shows a short code and opens your browser; confirm the code there. Your password is never typed into the app.
4. Find VEIRA Vocal Remover in the list and press INSTALL. When it says SHOW, it's installed. Open your DAW and let it rescan its plugins.

After installing, VEIRA Vocal Remover appears in your DAW under the manufacturer **VEIRA** (in Logic Pro: Audio Units ▸ VEIRA ▸ VEIRA Vocal Remover).

Activation and the licence

VEIRA Manager keeps a licence on your computer that the plugins check. It is refreshed every time VEIRA Manager opens and stays valid for 30 days offline, so you can work without internet as long as you open VEIRA Manager now and then.

Without a valid licence, VEIRA Vocal Remover still loads and works, but **its sound fades out for 1.5 seconds every 30 seconds**, and a strip along the bottom of its window says why (for example, *NOT ACTIVATED ON THIS COMPUTER. SIGN IN WITH VEIRA MANAGER.*). Clicking the strip opens your account page.



Without a licence: short silences every 30 s, faded in and out so nothing clicks.

FIGURE 2 What an unlicensed copy does. With a valid licence the audio is never touched.

- **Activating while your DAW is open** takes effect within about two minutes. You don't need to reload the plugin.
- **"Not activated on this computer"**: open VEIRA Manager and sign in.
- **"Expired"**: open VEIRA Manager while online; it renews the licence at once.
- **"Not included"**: the signed-in account doesn't own VEIRA Vocal Remover. Check that you're signed in with the account you bought it with.

Formats and where the files go

VEIRA Manager puts every format in the standard place, so you never need to copy files by hand:

Format	macOS	Windows
VST3	`~/Library/Audio/Plug-Ins/VST3` (just for me, the default) or `/Library/Audio/Plug-Ins/VST3` (all users)	`C:\Program Files\Common Files\VST3` (all users, the default) or `%LOCALAPPDATA%\Programs\Common\VST3`
Audio Units	`~/Library/Audio/Plug-Ins/Components` or `/Library/Audio/Plug-Ins/Components`	(Audio Units are macOS only)
Standalone	`~/Applications/VEIRA` or `/Applications/VEIRA`	`C:\Program Files\VEIRA` or `%LOCALAPPDATA%\Programs\VEIRA`

You choose *just for me* or *all users* in VEIRA Manager's install settings.

The standalone app runs VEIRA Vocal Remover on its own, with your audio interface's input and output, which is handy for trying presets without a DAW.

The panel at a glance



FIGURE 3 VEIRA Vocal Remover with the factory preset *Default: Init* loaded: the vocal removed. In the SEPARATION display the vocal shows as the wavy amber lines, the music as the steady blue ones.

	Part	What it does
1	Preset browser	Steps through and saves presets. See *Presets*.
2	SEPARATION display	The last eight seconds or so of the song as the AI hears it: vocal in amber, music in blue.
3	Legend	Names the two colours; a part you turn down is drawn dimmed.
4	VOCAL RANGE grips	The HI and LO limits of the vocal range, which you can drag on the display.
5	VOCAL	The level of the voice the AI finds, with its level bar.
6	MUSIC	The level of everything else, with its level bar.
7	REMOVE VOCAL, VOCAL ONLY, ORIGINAL	One click sets both knobs for the three everyday jobs.
8	STRENGTH	How firmly the AI decides what is voice.
9	VOCAL RANGE LOW, HIGH	Outside this range everything counts as music, untouched.
10	OUTPUT	The final level.

	Part	What it does
11	Status panel	AI LOAD, LATENCY and the LATE lamp.
12	Meter bridge	Input and output level and loudness.

The window has three areas: the **SEPARATION display** across the top, the controls underneath in three groups captioned LEVELS, SEPARATION and OUTPUT, and the **meter bridge** down the right.

Using the controls

Every VEIRA plugin is played like a piece of hardware. The same few parts appear on every panel:

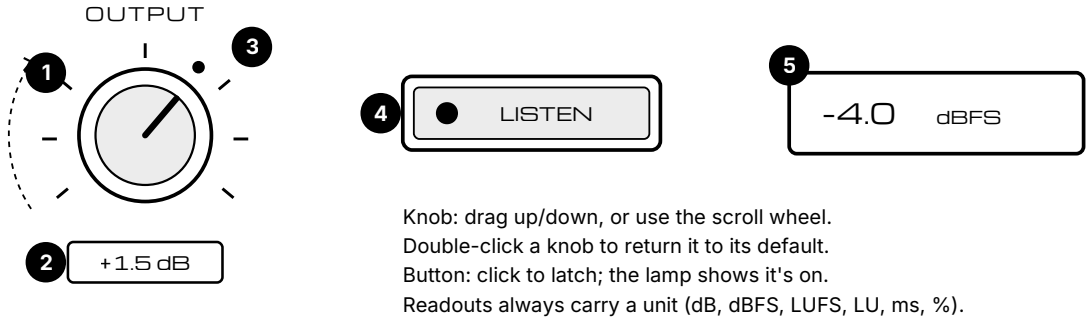


FIGURE 4 The parts you'll find on every VEIRA panel (the labels are examples).

- 1 **Knob.** Drag up to turn it clockwise, down to turn it back, or use your mouse wheel. **Double-click** returns it to its default. The engraved pointer and the small dot on the scale show where it's set.
- 2 **Value window.** The backlit window under each knob shows its exact value with its unit.
- 3 **Silkscreen label.** The control's name, as it appears in your DAW's automation list.
- 4 **Push button.** Click to switch on; the lamp lights and the cap sits lower in its bezel. Click again to switch off.
- 5 **Readout.** Live numbers update ten times a second so you can read them while the meters move. A value over 0 dBFS shows as a warning.

Tooltips: hover over any control or display for a short explanation. On displays, the tooltip depends on where you point.

Automation: every knob and button is a host parameter, so you can automate it in your DAW and control it from a MIDI controller or control surface that your DAW maps.

Units: VEIRA never shows a bare number. **dBFS** is level against digital full scale (0 dBFS is the loudest a digital signal can be); **dB** is a change in gain; **LUFS** is loudness as streaming services and broadcasters measure it; **LU** is a difference in loudness (1 LU = 1 dB).

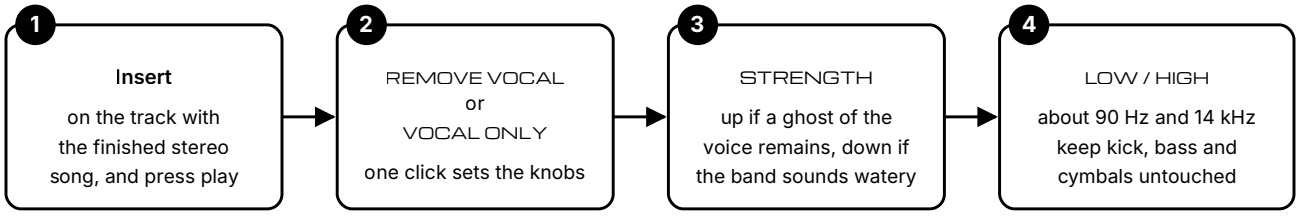
On VEIRA Vocal Remover's panel the parts look a little different from the drawing above:

- **Knobs** show their value **in their middle** (Off at the bottom of VOCAL, MUSIC, LOW and HIGH), and a lit ring (blue on screen; in print, the lighter arc) shows how far each is turned. Drag, use the wheel,

or double-click (or Alt/Option-click) for the default: VOCAL returns to Off, MUSIC to 0 dB. VOCAL and MUSIC replace the faders of earlier versions, and each has a slim **level bar** to its right.

- **Type a value:** right-click a knob (Control-click on a Mac). A field opens over the knob with the current value; type the new one and press Return. The number is enough: **-6** for VOCAL, **60** for STRENGTH; **off** switches VOCAL, MUSIC, LOW or HIGH off, and for LOW and HIGH, **12k**, **12 kHz** and **12000** all mean 12 kHz. Press Escape or click elsewhere to leave it unchanged.
- **Buttons** have no lamps: the button whose setting the knobs are at is filled (orange on screen; in print, the solid, lighter button), the others are dark with an outline.
- **Resize the window:** drag its bottom-right corner (the small diagonal lines). It scales from 75 % to 200 % of its normal size and keeps its proportions, so the display and the text stay sharp. Your DAW saves the size with the session.

Quick start: a karaoke track or an a cappella in four moves



The sound arrives about 1.2 s after you press play the first time; your DAW lines it up with the other tracks.

FIGURE 5 The quickest way to a good result. Or load a preset from the Karaoke or A Cappella folder: they already set steps 2 to 4.

How vocal separation works

If you have used other vocal removers, some of this will be familiar; if not, these pages explain what the plugin can and can't do, and why.

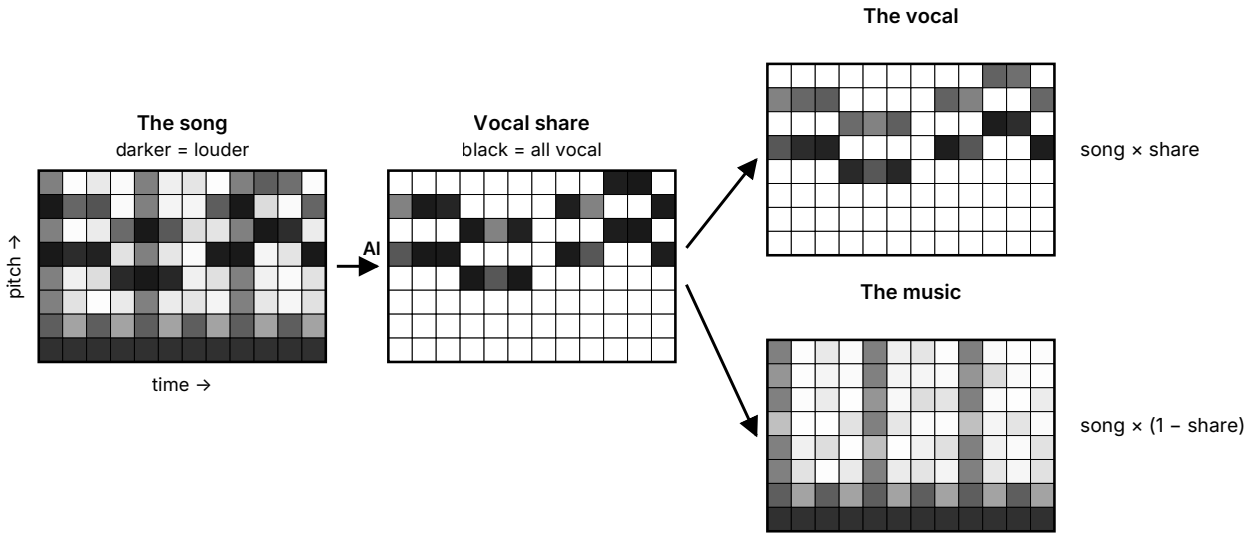
Why a mixed song is hard to take apart

When a song is mixed, every track is added together into one stereo file. The voice and the guitar that play the same note at the same moment become one sound wave. There is no switch inside the file that says which part of the wave was the singer: it has to be worked out by listening, the way your ear does.

The old trick for removing a vocal was **centre cancellation**: subtract the left channel from the right. Whatever sits exactly in the middle of the stereo picture, usually the lead vocal, cancels out. So do the kick, the bass and the snare, which are mixed in the middle too, while the vocal's stereo reverb stays. And the result is mono. EQ is no better: a voice shares its frequencies with almost every instrument, so any filter that takes out the voice takes out the band with it.

Listening by pitch and time

The way out is to look at the song not as one wave but as a picture of its pitches over time, a **spectrogram**: time runs left to right, pitch from bottom to top, and the louder a pitch is at a moment, the darker (on screen, the brighter) its square. In that picture a voice and a band look very different. A sung note is a stack of wavy lines that glide and wobble with the singer's pitch and vibrato, with vowels and consonants shaping them; drums are short vertical stripes; a bass is a steady line at the bottom.



Each square is one pitch at one moment. The two parts always add back up to the song.

FIGURE 6 How a separation works. The AI looks at the picture of the song and gives every pitch at every moment a vocal share, from none (white) to all of it (black). The song times the share is the vocal; the song times what's left is the music.

The AI model looks at that picture and decides, for every pitch at every moment, **what share of it is voice**. Where only the voice plays, the share is close to 100 %; where only the band plays, close to 0 %; where both play the same pitch, somewhere in between. Multiply the song by the share and you have the vocal; multiply it by the rest and you have the music. Because the two shares always add up to 100 %, the two parts always add back up to the song.

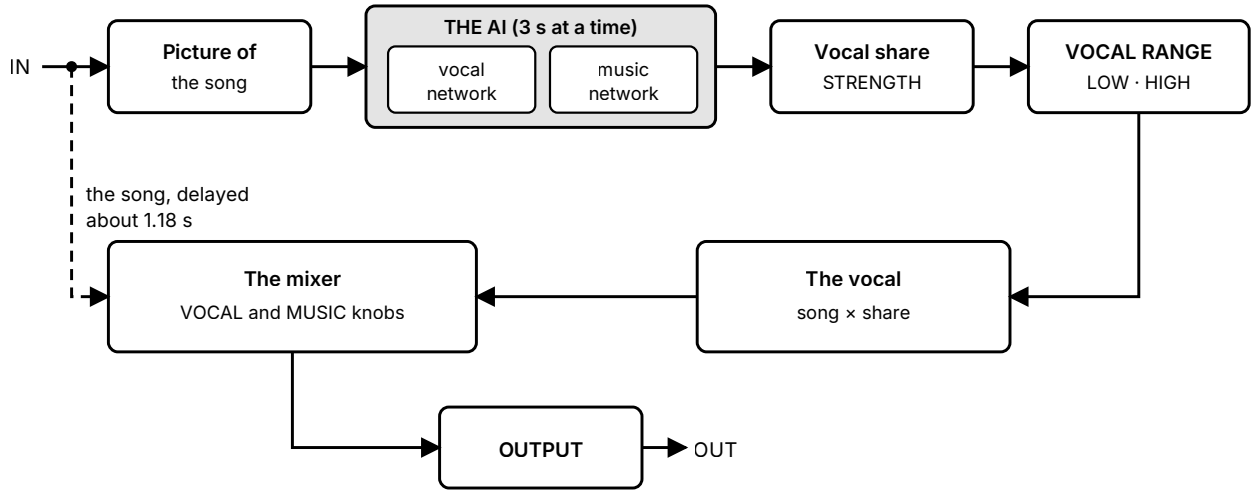
That last point matters. The AI never invents sound: it only divides what is in the song. Where it divides wrongly, a little of the voice stays in the music (a faint **ghost** of the singer), or a little of the music goes with the voice (**bleed**), and where the share changes quickly the result can sound **watery** or swirly. These artefacts are the limits of every vocal remover, and STRENGTH and the VOCAL RANGE are there to steer them.

The AI model

VEIRA Vocal Remover uses **Spleeter**, a separation model published by the research team of the music service Deezer under the open-source MIT licence. It is two neural networks of the U-Net kind, one trained to find the vocal and one to find the accompaniment, each of which looks at three seconds of the song at a time. VEIRA's own code runs them inside the plugin, on your computer's processor, without any AI framework or internet connection. Its licence is installed with the plugin.

Spleeter was trained on **sung music**. It is at its best on a lead vocal over a band. It works on spoken voices too, but much less cleanly (see *Singing and speech* below), and it hears a choir, a rap or a doubled backing vocal as voice like any other.

What happens inside VEIRA Vocal Remover



Output = MUSIC × the song + (VOCAL − MUSIC) × the vocal. With both knobs at 0 dB the song passes untouched.

FIGURE 7 Signal flow. The AI works on a picture of the song and gives each pitch a vocal share; STRENGTH and the VOCAL RANGE adjust it. The two knobs then mix the separated vocal with the original song, which waits for it.

Step by step:

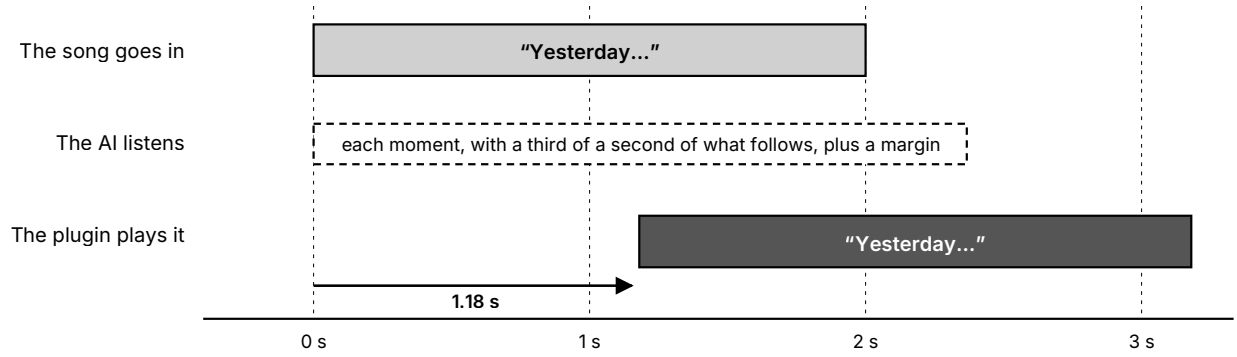
1. **The picture.** The song is cut into short overlapping slices and turned into a picture of its pitches (a spectrogram), as described above. The AI hears the song at 44.1 kHz, the rate it was trained at; at other

sample rates the plugin converts a copy of the song for the AI only, while the original audio stays at its own rate.

2. **The AI** runs both networks over the last three seconds of the song, a third of a second at a time, on a separate processing thread, so your DAW's audio never waits on it.
3. **The vocal share** comes from the two networks' answers together. [STRENGTH] weighs one against the other: higher counts more as voice.
4. **The VOCAL RANGE** sets the share to zero below [LOW] and above [HIGH]: there, everything is music.
5. **The vocal** is the song times the share.
6. **The mixer.** The [VOCAL] and [MUSIC] knobs mix the separated vocal with the original song. The music isn't a separate signal: it is *the song minus the vocal*. That is why [ORIGINAL] (both at 0 dB) returns the song exactly, and why every move of the knobs only changes the part the AI found. Level changes glide over 20 ms, so they never click.
7. [OUTPUT] sets the final level.

Why it needs more than a second of latency

To decide whether a sound is a voice, the AI needs to hear it in context: the note before, the note after, where the line goes. The plugin therefore **listens ahead**. Every sound waits until the AI has heard about a third of a second of what follows it, and the plugin keeps a further safety margin so that the AI, running on its own thread, always has time to finish. In all the song is held back by about **1.18 seconds** (52,224 samples at 44.1 kHz, 56,925 at 48 kHz).



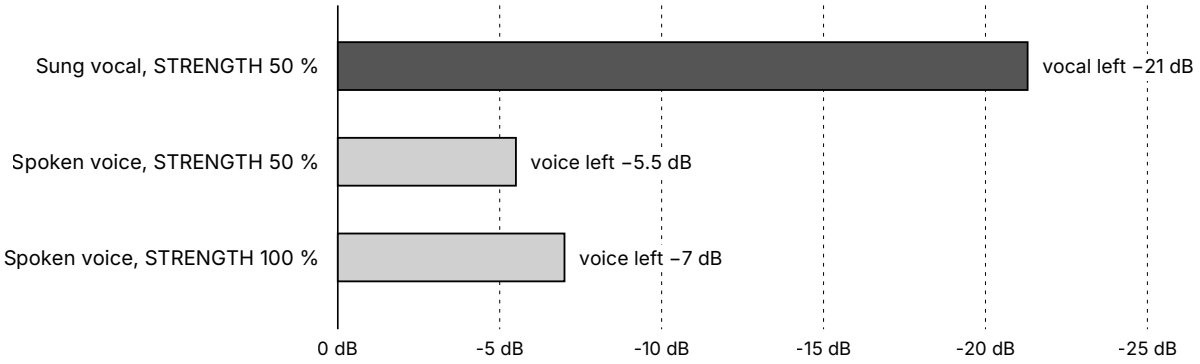
Your DAW knows the delay and reads the track 1.18 s early, so on playback it lines up with every other track.

FIGURE 8 The latency. The plugin reports it to your DAW, which compensates automatically on playback and when you bounce.

The plugin tells your DAW about the delay, and the DAW **compensates** automatically: it plays the track that much earlier, so on playback and in a bounce everything stays in time, as if there were no delay at all. What compensation can't fix is a **live** signal: if you sing or play *through* the plugin while recording, you hear yourself 1.18 seconds late. VEIRA Vocal Remover is made for playback, mixing and bouncing. The chapter *Latency and your DAW* has the details, including Pro Tools.

Singing and speech

The model was trained on singing. On a **spoken** voice, such as a voice-over or an interview with music under it, it separates far less: in our tests a spoken voice was turned down only 5 to 7 dB with VOCAL off, against about 21 dB for a sung one, and when isolated it kept only part of the voice. The *Speech* presets are built for this: they duck a voice or a music bed rather than removing it, and hold the voice up with VOCAL.



REMOVE VOCAL on the test songs: how far down the voice ends up. Longer bar = more of it removed.

FIGURE 9 Singing against speech, measured on the test songs (the same band under a female pop, a male rock and a spoken voice). With speech, expect a quieter voice rather than none.

Controls

LEVELS: the VOCAL and MUSIC knobs

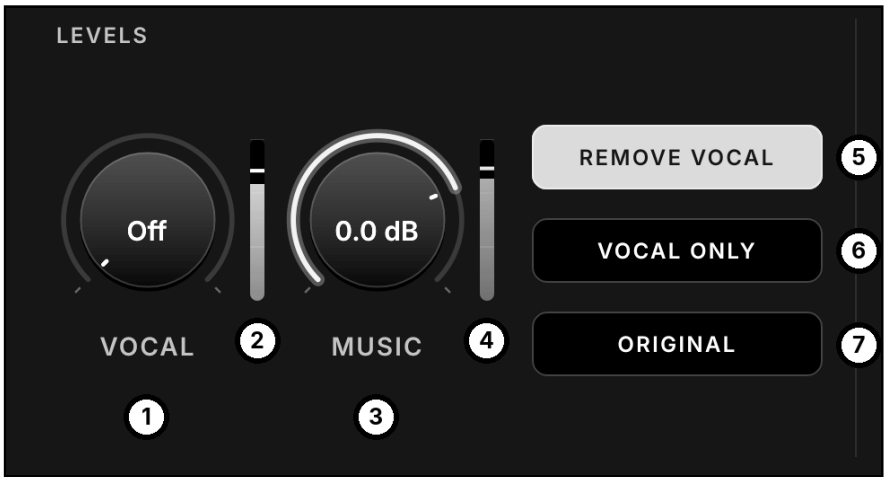


FIGURE 10 1 VOCAL (here Off) with 2 its level bar, 3 MUSIC (here 0.0 dB) with 4 its level bar, and the buttons 5 REMOVE VOCAL (filled: the knobs are at its setting), 6 VOCAL ONLY and 7 ORIGINAL.

Control	Range	Default	What it does
VOCAL	Off, then −60 to +6 dB	Off	The level of the voice the AI finds.
MUSIC	Off, then −60 to +6 dB	0.0 dB	The level of everything else.

Both knobs move in 0.1 dB steps; fully anticlockwise is Off (silence). 0 dB sits at three quarters of the way round, as on a console fader: the travel is spread out around 0 dB, where you make fine moves, and squeezed towards the bottom. To set an exact level, right-click the knob and type it.

VOCAL sets the level of the separated voice. Off removes it: the song without its singer. 0 dB leaves it where it was. In between turns it down; above 0 dB (up to +6 dB) turns it up.

MUSIC sets the level of everything that isn't voice. Off leaves the voice alone: an a cappella. 0 dB leaves the music where it was.

Both at 0 dB, the song passes untouched, bit for bit, just delayed by the latency. Every other setting changes only the part the AI found. You can think of it like this:

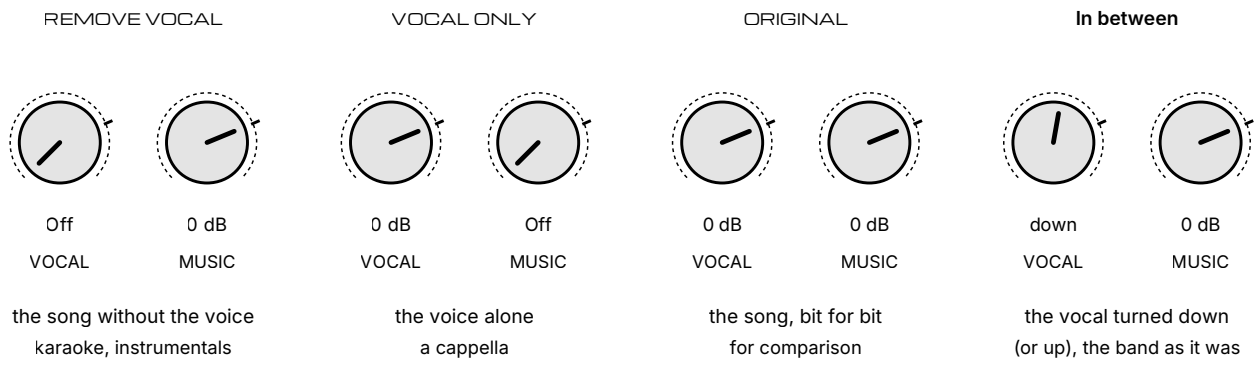


FIGURE 11 The three buttons are three settings of the two knobs. Every other setting lies in between.

The three buttons 5 6 7 set both knobs in one click:

Button	VOCAL	MUSIC	Use
REMOVE VOCAL	Off	0 dB	The music without the voice: karaoke, instrumentals, practice.
VOCAL ONLY	0 dB	Off	The voice without the music: an a cappella.
ORIGINAL	0 dB	0 dB	The song as it was, to compare.

A button is filled (orange on screen) when the knobs are at its setting, whether you clicked it or turned the knobs there yourself. The buttons aren't parameters of their own: they move the two knobs, so your DAW records a click as two moves in its automation.

> **Tip:** ORIGINAL is the fair way to compare: the same delay, the same level, the song exactly as it went in. Click it, then click back to hear what the separation does.

The knob is not the vocal's exact level

Turning VOCAL down 6 dB does **not** turn the singer down exactly 6 dB. About **0.8 dB** of a typical vocal is never separated: the AI hears that part of the voice as music, so it stays at the MUSIC knob's level whatever VOCAL does. For small moves this hardly matters; for big ones it does:

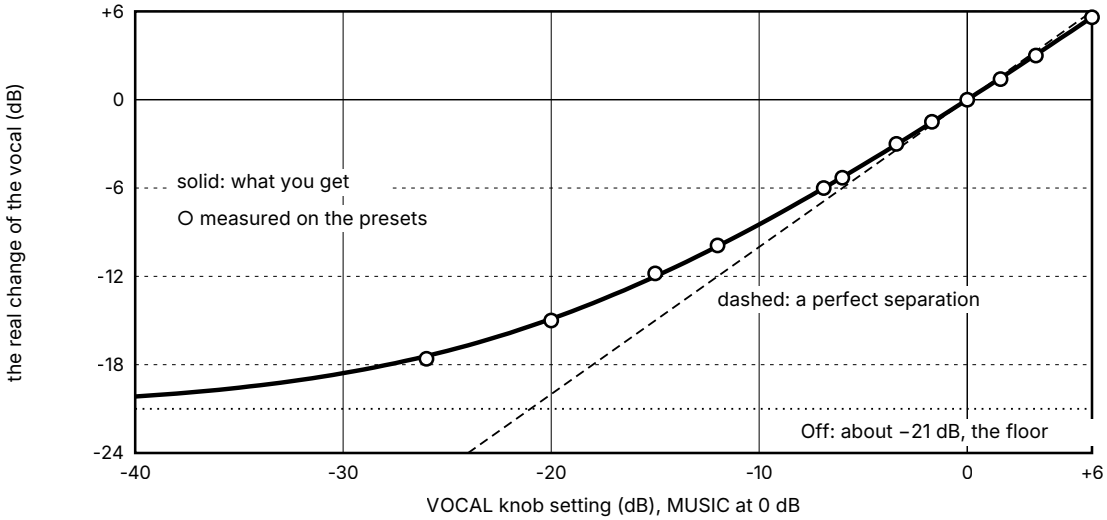


FIGURE 12 The VOCAL knob against the real change of the sung vocal in the output (with MUSIC at 0 dB, measured on the test songs). Near 0 dB the two agree within a few tenths of a dB; lower down the real change falls behind; [Off] leaves the vocal about 21 dB down.

[VOCAL] knob	The vocal really changes by about
+6 dB	+5.6 dB
+3.3 dB	+3.0 dB
−3.4 dB	−3.0 dB
−6 dB	−5.3 dB
−12 dB	−10 dB
−20 dB	−15 dB
[Off]	−21 dB (the floor)

So the lowest 40 dB of the knob's travel only cover the last few dB of real change, and **removal bottoms out about 21 dB down**: that's a voice you'd barely notice under a full band, but in a quiet passage or a sparse arrangement you may still hear a faint ghost of it. [STRENGTH] lowers the floor a little (see below). The *Vocal Balance* presets set the knob past their target, so their measured change matches their name.

Big upward moves have a different catch: **there is no peak protection.** Turning the vocal up on a loud, mastered song raises its peaks: in our tests [VOCAL] at +6 dB took a master that peaked at −0.9 dBFS to +2.5 dBFS. Watch the OUT peak on the meter bridge and lower [OUTPUT] until it stays under 0 dBFS (the Vocal Up presets' notes say how much).

The level bars

To the right of each knob, a slim bar ❷ ❹ shows **how much of that part there is right now, before the knob**, from −60 to 0 dBFS, with faint marks at 0, −20 and −40: amber for the vocal, blue for the music, red near the top (above −3 dBFS). In print, the two bars look alike; the one beside [VOCAL] is the vocal's. A thin

line holds the recent peak for a moment and a half. Because it's measured before the knob, the VOCAL bar keeps showing the voice the AI finds even while VOCAL is at Off: it's a quick way to see where the singing starts and stops.

SEPARATION: STRENGTH and the VOCAL RANGE

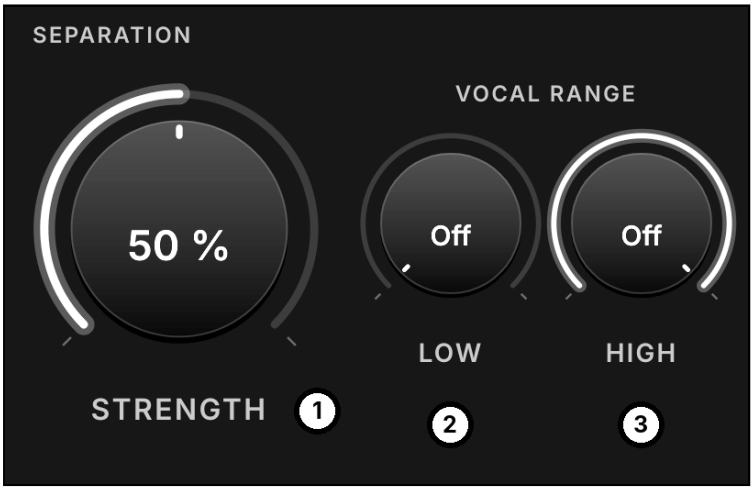


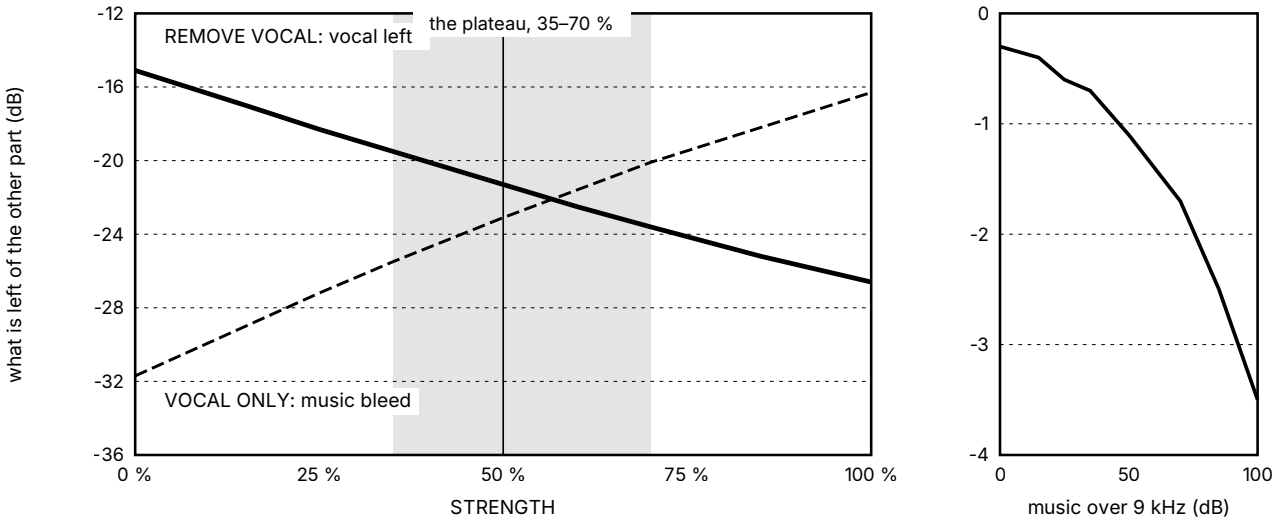
FIGURE 13 1 STRENGTH (here 50 %), and the VOCAL RANGE: 2 LOW and 3 HIGH (both here Off).

Control	Range	Default	What it does
STRENGTH	0 to 100 %	50 %	How firmly the AI decides what is voice.
VOCAL RANGE LOW	20 Hz (Off) to 1 kHz	Off	Below this, nothing counts as vocal.
VOCAL RANGE HIGH	1 kHz to 20 kHz (Off from 19.5 kHz)	Off	Above this, nothing counts as vocal.

STRENGTH

STRENGTH sets how much of the sound counts as voice. At **50 %**, the default, the AI's two networks are weighed as they were trained: the model's own balance. Higher counts more as voice; lower counts less.

- **Removing a vocal** (REMOVE VOCAL): higher STRENGTH takes more of the voice out, including its reverb tails, breaths and doubled backing vocals, but takes more of the music with it, mostly the cymbals' air, and can sound watery. Lower is gentler and more natural, and leaves a faint ghost of the voice.
- **Keeping only the vocal** (VOCAL ONLY): it works the other way round. Higher lets more of the band bleed into the vocal; lower gives a cleaner but thinner, drier voice (it loses the quiet tails and breaths first).



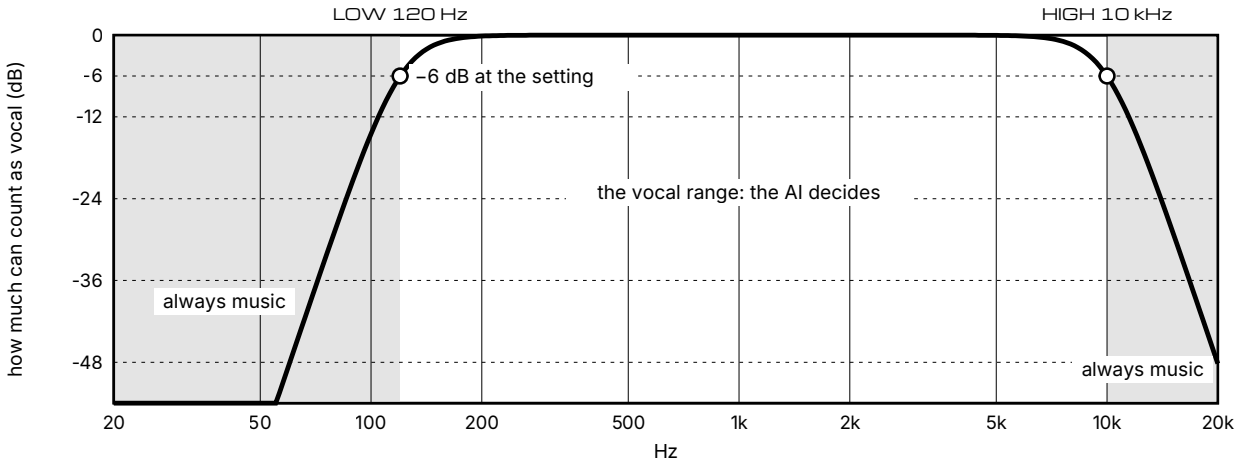
Sung test songs, VOCAL RANGE off. Right: what removing the vocal costs the cymbals and air.

FIGURE 14 STRENGTH measured on the test songs. Removing a vocal, higher STRENGTH leaves less of it (solid). Isolating it, higher STRENGTH lets more music in (dashed). From 35 to 70 % (grey) the overall result hardly changes; past that, removal mostly costs the top of the music (right).

From **35 to 70 %** the overall quality hardly changes: it's a plateau, so in that range STRENGTH is a matter of taste. Around the middle, each 10 % is worth roughly 1 dB more (or less) removal. Outside the plateau the result gets worse overall: the music over 9 kHz loses 1.1 dB at 50 % but 3.5 dB at 100 %. If you need a high STRENGTH to get rid of a stubborn vocal, pair it with HIGH at 10 to 12 kHz to keep the cymbals (see below). The factory presets stay between 30 and 85 %.

The VOCAL RANGE: LOW and HIGH

The VOCAL RANGE is a frequency window. **Inside it**, the AI decides what is voice. **Outside it**, everything counts as music and is left exactly as it was, whatever the AI thinks. A sung voice has little below about 90 Hz, and above 12 kHz little but breath and air, so limiting the range costs the voice very little and protects the music where it matters most.



Beyond each limit the share falls 48 dB per octave: an octave below LOW, nothing is taken as voice.

FIGURE 15 The VOCAL RANGE with LOW at 120 Hz and HIGH at 10 kHz. At each setting the share that can count as vocal is halved (−6 dB); an octave beyond, almost nothing (−48 dB). In the grey areas everything stays music.

- [LOW] keeps the kick and the bass completely clean when you remove a vocal, and keeps them out of an a cappella. Up to about **90 Hz** it costs nothing. At 120 Hz a male voice starts to leave about 1 dB more of itself behind, at 150 Hz about 2 dB more; high voices don't mind 150 Hz. From 200 Hz up, it eats into every voice. In an a cappella of a male singer, [LOW] at 150 Hz cut the bass bleed from −35 to −48 dB in our tests. At its lowest setting the knob reads [Off]; type **off** to get there at once.
- [HIGH] keeps the cymbals, hi-hats and air in the music. At **10 to 14 kHz** it keeps the air almost untouched for very little extra vocal left (0.3 to 0.6 dB). In an a cappella it cuts the cymbal wash steeply; at 12 kHz and above, the voice's "s" sounds stay. From 19.5 kHz up the knob reads [Off].

Above about 11 kHz the model itself doesn't analyse the song; the plugin carries on the decision it made just below. That's why the air follows [STRENGTH] so closely, and why [HIGH] is the precise tool up there.

Both limits are also the [HI] and [LO] grips on the right edge of the SEPARATION display: drag them there, or double-click one to switch it off (see *Displays*). Creative presets use narrow ranges on purpose: 400 Hz to 3 kHz gives a telephone voice, 300 Hz to 3 kHz with [VOCAL] off a hollow one.

OUTPUT

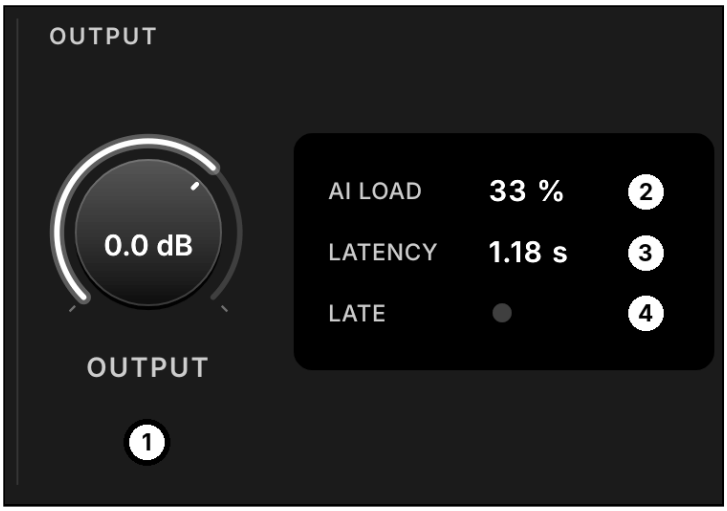


FIGURE 16 1 OUTPUT, and the status panel: 2 AI LOAD, 3 LATENCY and 4 the LATE lamp (see *Displays*).

Control	Range	Default	What it does
OUTPUT	−24 to +12 dB	0.0 dB	The final level, after VOCAL and MUSIC.

OUTPUT sets the level leaving the plugin. Removing the vocal from a song makes it quieter (typically 3 to 4 dB), but its peaks, which usually come from the drums, hardly move: on a mastered song, making the level up by more than about 0.4 dB puts the instrumental's peaks above the original's. Turning the vocal up does the opposite. Watch the OUT peak on the meter bridge when you raise OUTPUT or VOCAL.

Displays

The SEPARATION display

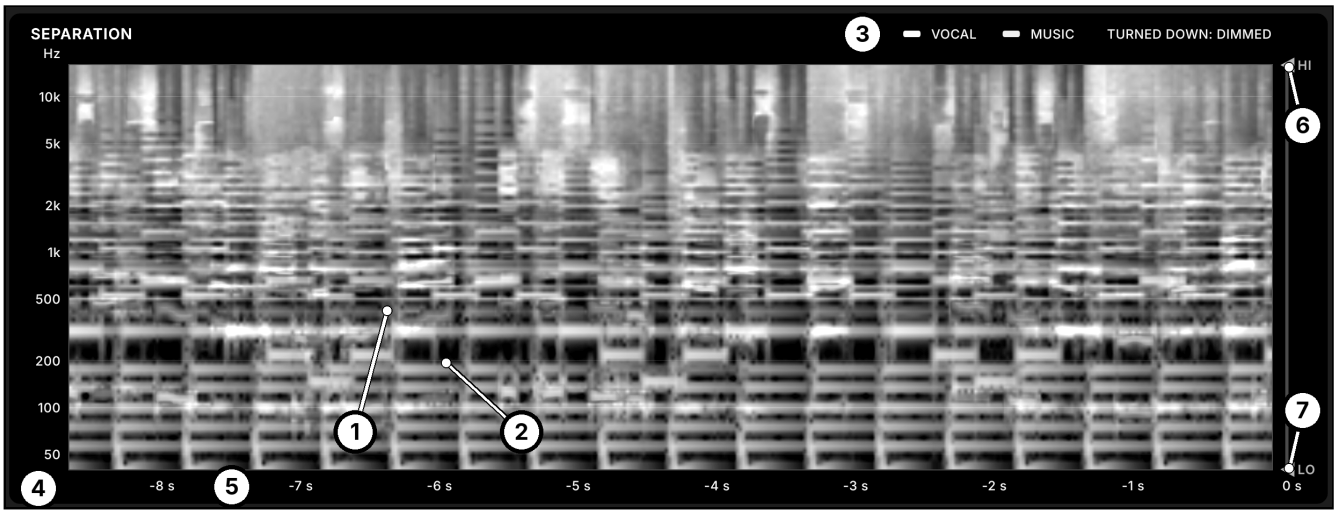


FIGURE 17 1 The vocal (amber on screen: the wavy lines), 2 the music (blue on screen: the steady lines), 3 the legend, 4 the pitch scale in Hz, 5 the time scale, and the VOCAL RANGE grips 6 HI and 7 LO (both here off, at the ends of the rail).

The SEPARATION display is a scrolling picture of the song as the plugin hears it: a spectrogram, as described in *How vocal separation works*.

- **Time** runs from left to right 5 : the newest sound enters at the right edge (0 s) and scrolls left, with a mark every second, so you see the last eight and a half seconds or so.
- **Pitch** runs from bottom to top 4 , from 40 Hz to 16 kHz on a musical scale (each octave the same height).
- **Colour** shows what the AI decided. The **vocal** 1 is drawn in **amber**, and its loudest notes glow towards cream; the **music** 2 in **blue**. Without colour, tell them apart by their shape: a sung line glides and wobbles with the melody and vibrato, while instruments draw straight horizontal lines (held notes, bass) and short vertical stripes (drums). In greyscale the amber reads slightly lighter than the blue.
- **Brightness** shows level: the louder a pitch, the brighter its point.

Your knob settings show too. A part you turn down is drawn **dimmer**, as the legend 3 reminds you (TURNED DOWN: DIMMED): with REMOVE VOCAL the vocal shows as faint amber, with VOCAL ONLY the blue fades. And the area outside the VOCAL RANGE is shaded darker, so you can see at a glance what the AI is allowed to touch.

The VOCAL RANGE grips. The rail on the right edge of the display holds two small triangles: HI 6 for HIGH and LO 7 for LOW. When a limit is off, its grip rests grey at the end of the rail (HI at the top, LO at the bottom). When it's on, the grip sits at its frequency, lit (amber on screen; in print, lighter than a grip that is off), and a short line marks the limit on the edge of the plot.

- **Drag** a grip up or down to set its limit; the knob below follows.
- Drag HI to the very top or LO to the very bottom to switch it off.

- **Double-click** a grip to switch that limit off.

Hover anywhere in the picture for a read-out of that moment and pitch: how long ago (for example *−2.4 s*), the frequency, and the level the AI found there for each part (*VOCAL −18 dB MUSIC −34 dB*). A thin line marks the moment you point at.

What to look for:

- **Before you remove anything,** play the song with [ORIGINAL] and look: where is the amber? If whole chords or a lead guitar show amber, the AI hears them as voice, and they'll go with the vocal; a [LOW] or [HIGH] limit, or a lower [STRENGTH], can keep them.
- **Set [LOW] by eye:** move the [LO] grip up until it sits just under the lowest amber line of the voice, and not above it.
- **Ghosts:** with [REMOVE VOCAL], faint amber where the voice was is the part being removed, drawn dimmed; it's not left in the sound. Judge what's left by ear.

The status panel

The status panel ❷ to ❹ in the *OUTPUT* figure shows how the AI is keeping up:

- [AI LOAD] is how much of its available time the separation is using. Under 100 % it keeps up; the figure turns orange on screen above 80 %. It depends on your computer: on a recent processor it is usually well under 100 %, and Apple silicon Macs need much less than Intel or AMD processors. When the song is silent, the AI doesn't run at all.
- [LATENCY] is the plugin's delay at your sample rate, in seconds: about 1.18 s. Your DAW compensates for it (see *Latency and your DAW*).
- **The [LATE] lamp** lights (red on screen) if the computer ever couldn't finish the AI's work in time, and fades over a couple of seconds. The plugin then keeps its last decision for those moments instead of glitching, so the vocal may leak through briefly. An **offline bounce** never runs late: the plugin then simply takes the time it needs.

If the [LATE] lamp flickers often, see *Troubleshooting*.

If the AI model can't be found or loaded (for example after a damaged installation), the status panel says [NO MODEL] followed by the reason, and the song passes through as music: nothing is separated. Reinstall the plugin with VEIRA Manager.

Meter bridge (IN | OUT)

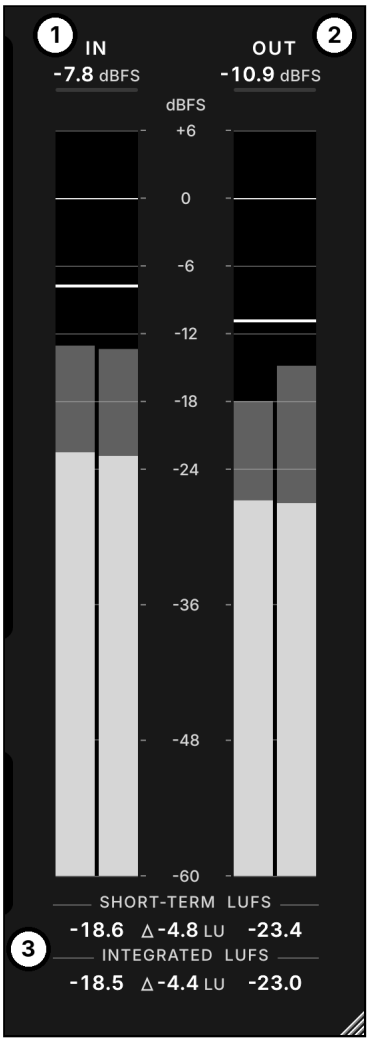


FIGURE 18 The meter bridge: 1 input and 2 output level, 3 short-term and integrated loudness with the OUT − IN difference.

VEIRA Vocal Remover's meter bridge is the standard one, measured on the plugin's input and output. The Δ readings tell you how much quieter the song got without its vocal (with REMOVE VOCAL, typically 3 to 4 LU). On a mono track, both bars of each meter show the same mono signal.

Reading the IN and OUT meters

The meter bridge measures what goes into VEIRA Vocal Remover and what comes out, in level and in loudness. The drawing shows the input and output columns; any extra column on your plugin's bridge is described above.

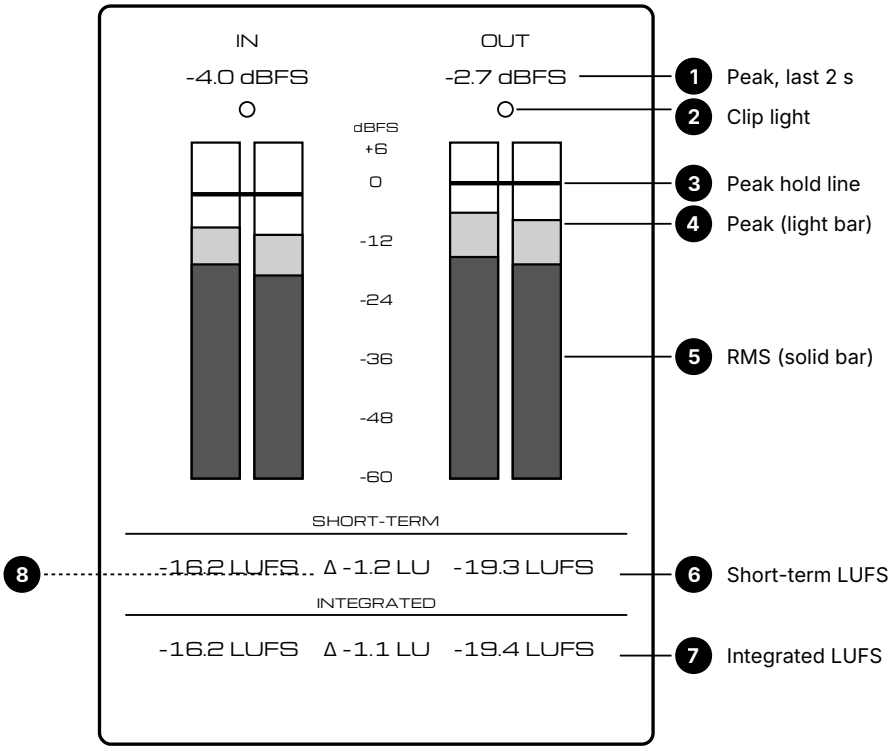


FIGURE 19 The meter bridge, left and right channel bars for input and output, with loudness below.

- 1 **Peak readout:** the highest sample peak of the last two seconds, in dBFS. It turns to a warning above 0 dBFS.
- 2 **Clip light:** lights on any sample at or over full scale. It stays lit until you click the meter.
- 3 **Hold line:** marks the recent peak on the bar.
- 4 **Peak bar** (the lighter part): the instantaneous sample peak of each channel.
- 5 **RMS bar** (the solid part): the average level over 300 ms, closer to how loud it sounds.
- 6 **SHORT-TERM:** loudness over the last 3 seconds, in LUFS, for the input (left) and the output (right).
- 7 **INTEGRATED:** loudness of everything since the measurement started, gated as the ITU-R BS.1770-4 and EBU R 128 standards require.
- 8 **Δ (delta):** OUT minus IN, in LU. It tells you how much louder (+) or quieter (−) VEIRA Vocal Remover made the signal, so you can match levels fairly when you compare with bypass. It reads [--] while either side is too quiet to measure (below −70 LUFS).

The scale: on screen the bars are coloured like a studio console: green below −12 dBFS, yellow from −12 to −3, orange from −3 to 0 and red above 0 dBFS.

Clicks: click a level meter to reset its hold and clip light. Click the loudness readouts to restart the integrated measurement, for example at the start of a song.

Latency and your DAW

Playback, mixing and bouncing

VEIRA Vocal Remover reports its latency of about 1.18 seconds to your DAW, which **compensates** for it: on playback and in a bounce, the track lines up with every other track, sample for sample. You don't need to do anything, and in most DAWs plugin delay compensation is on by default.

- **Bypass stays in time.** When you bypass the plugin, it still delays the song by the same amount, so switching bypass on and off never makes the track jump.
- **Offline bounces** (export, bounce in place, render, freeze) are always complete: the AI takes as long as it needs, and the LATE lamp never lights.

Recording and live monitoring

Delay compensation can't make a live sound arrive early. If you record a singer or an instrument **while listening through** VEIRA Vocal Remover (on the input or on a track you're recording to), they'll hear themselves 1.18 seconds late. For singing along to a karaoke track, or playing along to a practice track:

1. Put VEIRA Vocal Remover on the **song's track**, not on the track you record.
2. Monitor your microphone or instrument on its own track, with no VEIRA Vocal Remover on it.
3. Your DAW plays the song in time, and you hear yourself without delay. Even better, bounce the karaoke track once and record against the bounced file.

Pro Tools

Pro Tools versions arrive once they are signed for Pro Tools (see *Getting started*). When they do:

- Set **Setup ▸ Playback Engine ▸ Delay Compensation Engine** to **Maximum**.
- Pro Tools can compensate only a limited delay per track, and that limit may be shorter than 1.18 seconds; we are still confirming it. If the track plays out of time with the rest of the session, don't use the plugin live on that track: **render** the result instead, offline, and work with the rendered audio.

Other hosts and setups

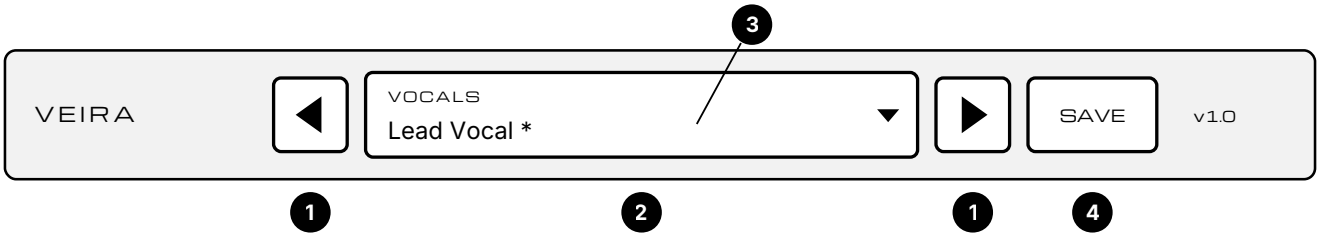
If a host can't compensate a delay this long, or you use the standalone app, the sound simply arrives 1.18 seconds late. Bounce or export the result to a new file and use that in your session: an offline render is exact.

Presets

VEIRA Vocal Remover comes with **45 factory presets** in 8 folders. Each one is a starting point made for a real job. Load one close to what you need, then adjust the first control its note mentions.

The preset browser

The preset browser sits in the header, at the top of the window.



The header: logo and plugin name, then the preset browser, then the version.

FIGURE 20 The preset browser.

- 1 ◀ ▶ step to the previous or next preset, through every folder in order.
- 2 **The display** shows the folder (small, above) and the preset's name. Click it to open the preset menu. Hover over it to read the preset's note: what it's for and what to adjust first.
- 3 **An asterisk (*)** after the name means you've changed something since the preset loaded.
- 4 SAVE stores your current settings as a user preset.

The preset menu

Clicking the display opens a menu with:

- one submenu per **factory folder**, in the order shown below;
- **User Presets**, with your own presets in their folders;
- **Save...**, **Delete** (only for user presets: factory presets can't be deleted), **Rescan** (re-reads the user folder, if you copied presets in by hand) and **Show User Preset Folder** (opens it in Finder or Explorer).

Saving your own presets

1. Set VEIRA Vocal Remover the way you like it.
2. Press SAVE. A dialog opens inside the plugin window.
3. Type a name, choose a folder or type a new folder name, and confirm.

Each user preset is one small .veirapreset file:

	Folder
macOS	~/Library/Audio/Presets/VEIRA/VEIRA Vocal Remover/<Folder>/
Windows	%APPDATA%\VEIRA\VEIRA Vocal Remover\Presets\<Folder>\

Copy these files to move your presets to another computer, or to back them up. Presets keep loading in future versions: any control added later simply starts at its default.

Presets in your DAW

The factory presets also appear in your DAW's own preset or program menu, named *Folder: Name* (for example *Vocals: Lead Vocal*). Whatever preset is selected is saved with your session and comes back when you reopen it, including the asterisk if you'd changed it.

The **Default ▸ Init** preset equals a fresh instance of the plugin: every control at its default.

The factory folders

Folder	Presets
Default	Init
Karaoke	Karaoke, Karaoke Gentle, Karaoke Strong, Karaoke Keep Cymbals, Karaoke Keep Low End, Karaoke Deep Voice
A Cappella	A Cappella, A Cappella Clean, A Cappella Full, A Cappella High Voice, A Cappella No Cymbals, AI Vocal Raw
Vocal Balance	Vocal Up 6 dB, Vocal Up 3 dB, Vocal Up 1.5 dB, Vocal Down 1.5 dB, Vocal Down 3 dB, Vocal Down 6 dB, Vocal Down 10 dB
Practice	Guide Vocal, Faint Guide Vocal, Learn The Melody, Bass Play Along, Drum Play Along, Guitar Play Along
Remix & Sampling	Clean Instrumental, Instrumental Max Removal, Acapella For Remix, Acapella Tight, Vocal Chop Source, Acapella With Ghost Beat, Instrumental Keep Air
Speech	Music Down 3 dB, Music Down 6 dB, Music Down 12 dB, Voice Over Faint Bed, Dialogue Extract, Voice Under Music
Creative FX	Telephone Vocal, Muffled Vocal, Whisper Layer, Hollow Vocal, Vocal Spotlight, Distant Radio Mix

The folders are sorted by job, not by instrument:

- **Karaoke**: the voice out, at different strengths, each keeping a different part of the band untouched (the cymbals, the low end, a deep voice's chest). Measured: the vocal 19 to 23 dB down, the music within 0.6 dB.
- **A Cappella**: the voice alone at its own level, from the fullest (*A Cappella Full*) to the least bleed (*A Cappella Clean*). *AI Vocal Raw* is exactly what the AI hears as vocal, with no limits: a reference for setting STRENGTH and the VOCAL RANGE on a new song. Measured: the band's bleed 20 to 27 dB down.
- **Vocal Balance**: the vocal up 1.5, 3 or 6 dB, or down 1.5, 3, 6 or 10 dB, with the band untouched. Each measured within 0.1 dB of its name (*Vocal Up 6 dB*: +5.6 dB).
- **Practice**: guide vocals for singers (the original voice quiet under you), *Learn The Melody*, and play-alongs for bass, drums and guitar that keep the part you play along to natural.
- **Remix & Sampling**: instrumentals and acapellas for DJs and producers, with the VOCAL RANGE set for the job.

- **Speech**: a music bed under a voice-over turned down 3, 6 or 12 dB, a voice over a faint bed, dialogue pulled out of music, and a narration pushed back under it. Speech separates less cleanly than singing: check these on your own material.
- **Creative FX**: telephone, muffled, whispered and hollow vocals, a breakdown spotlight and a radio in another room.

Levels and peaks. The separation itself doesn't depend on the song's level, so the presets work the same on quiet and loud material. What depends on level is the peak. The presets' OUTPUT never makes up more than the part they keep lost, so on a mastered song they stay under its original peak, **except the three Vocal Up presets**: turning a vocal up raises a master's peaks. Their notes say how much to lower OUTPUT on a loud master (about 0.7, 1.5 and 3.5 dB).

Recipes

Step-by-step starting points. In each, compare with ORIGINAL now and then, and judge by ear: the SEPARATION display shows what the AI decides, not how it sounds.

A karaoke track

1. Insert VEIRA Vocal Remover on the song's track and load **Karaoke ▸ Karaoke**.
2. Play the loudest chorus. If a ghost of the voice remains, raise STRENGTH towards 70 % or load **Karaoke ▸ Karaoke Strong**. If the band sounds thin or watery, lower it, or load **Karaoke ▸ Karaoke Gentle**.
3. If the cymbals lose their sparkle, load **Karaoke ▸ Karaoke Keep Cymbals** or lower HIGH to 8–10 kHz. For a deep male lead, **Karaoke ▸ Karaoke Deep Voice**.
4. Bounce the track: an offline bounce is always complete. Sing over the bounced file (see *Latency and your DAW*).

An a cappella for a remix

1. Load **Remix & Sampling ▸ Acapella For Remix**, or **A Cappella ▸ A Cappella** for listening.
2. Play a busy section. If the old beat bleeds through, lower STRENGTH (or load **Acapella Tight** for the least bleed, at the cost of a thinner, drier voice).
3. Raise LOW until the kick and bass disappear from the vocal, but not into the voice: watch the LO grip against the lowest amber line. 150 Hz suits most female voices.
4. To line the vocal up with a new arrangement, **Acapella With Ghost Beat** keeps the original track 24 dB under it; switch MUSIC off once it's in place.
5. Bounce the result to a new file.

Turn a vocal up or down in a finished mix

1. Load **Vocal Balance ▸ Vocal Down 3 dB** (or whichever is closest to what you want).
2. Fine-tune with VOCAL. Remember that the knob moves a little further than the real change (see *The knob is not the vocal's exact level*).
3. Turning a vocal **up**: watch the OUT peak on the meter bridge and lower OUTPUT until it stays under 0 dBFS.
4. Compare with ORIGINAL at matched loudness: the Δ INTEGRATED reading should be close to 0 LU for a small move.

A practice track for your instrument

1. Load the play-along for your instrument: **Practice ▸ Bass Play Along**, **Drum Play Along** or **Guitar Play Along**. Each removes the voice and leaves the frequency range of your instrument untouched.
2. If you want the singer as a guide, load **Practice ▸ Guide Vocal** instead (the voice about 12 dB down) and lower VOCAL as you get confident.
3. To learn a melody or a lyric by ear, load **Practice ▸ Learn The Melody**: the voice forward and the band back. Lower MUSIC further to hear only the line.

Lower the music under a voice-over

1. Load **Speech ▸ Music Down 6 dB** on the track with the voice-over and its music already mixed.
2. Set the bed with MUSIC. If the voice dips with the bed, raise VOCAL a little: a spoken voice is only partly separated, so part of it moves with the music.
3. For the voice nearly alone, **Voice Over Faint Bed** or **Dialogue Extract**; expect some bed to remain and a thinner voice.
4. Listen to the whole piece: speech separates less evenly than singing.

Effects: telephone, radio, spotlight

1. **Creative FX ▸ Telephone Vocal** gives the voice alone through a 400 Hz to 3 kHz band, for intros and breakdowns. Move LOW and HIGH for a different colour.
2. **Creative FX ▸ Vocal Spotlight** drops the band 12 dB under a lifted voice; automate MUSIC back up for the drop.
3. **Creative FX ▸ Muffled Vocal** removes the top half of the voice only, so it sounds as if it's behind a wall while the band stays bright.

Specifications

Type	AI vocal separation: splits a mixed song into vocal and music and mixes them back at your levels
Model	Spleeter 2-stem by Deezer Research (MIT licence), run by the plugin's own code on your computer; no internet
Channels	Stereo and mono (a mono track is analysed as the same signal on both sides)
Latency	About 1.18 s: 52,224 samples at 44.1 kHz, 56,925 at 48 kHz; constant, reported to the host for delay compensation
Sample rates	Any; the AI hears the song at 44.1 kHz (other rates are converted for the analysis only), and ORIGINAL is exact at every rate
Transparency	VOCAL and MUSIC at 0 dB, and bypass: the output is the input delayed by the latency, bit for bit; level changes glide over 20 ms
Measured separation	Test songs, STRENGTH 50 %: a sung vocal removed to about −21 dB, the band's bleed in an a cappella about −23 dB; speech 5–7 dB
VOCAL	Off, −60 to +6 dB in 0.1 dB steps (default Off); 0 dB at three quarters of the travel
MUSIC	Off, −60 to +6 dB in 0.1 dB steps (default 0.0 dB)
STRENGTH	0 to 100 % in 1 % steps (default 50 %)
VOCAL RANGE LOW	20 Hz (Off) to 1 kHz (default Off); −6 dB at the setting, 48 dB per octave below
VOCAL RANGE HIGH	1 kHz to 20 kHz, Off from 19.5 kHz (default Off); −6 dB at the setting, 48 dB per octave above
OUTPUT	−24 to +12 dB in 0.1 dB steps (default 0.0 dB)
Displays	SEPARATION (scrolling spectrogram, about 8.7 s, 40 Hz–16 kHz, vocal and music coloured, VOCAL RANGE grips); level bar beside VOCAL and MUSIC (−60 to 0 dBFS); status panel (AI LOAD, LATENCY, LATE)
Metering	IN and OUT meters (sample peak, RMS); short-term and integrated loudness per ITU-R BS.1770-4 / EBU R 128
Memory	About 80 MB for the model, shared by every instance
Presets	45 factory presets in 8 folders, plus your own
Window	1520 × 860 points; resizable from 75 % to 200 %, saved with the session
Formats	VST3, Audio Units, standalone; AAX once signed
Plugin ID	Manufacturer Veir, plugin Vvoc

Troubleshooting and questions

The sound drops out for a moment every 30 seconds. The plugin isn't activated on this computer. Open VEIRA Manager and sign in; the strip at the bottom of the plugin's window says exactly what's wrong. See *Getting started ▸ Activation and the licence*.

I don't see VEIRA Vocal Remover in my DAW. Make sure VEIRA Manager shows SHOW (installed) for it, then rescan plugins in your DAW. In Logic Pro, open the Plug-in Manager and choose *Reset & Rescan Selection* if needed. Look under the manufacturer **VEIRA**.

The track is out of time with the rest of the session. Check that your DAW's plugin delay compensation is on. In Pro Tools, set the Delay Compensation Engine to Maximum; if it still plays out of time, the delay is longer than Pro Tools can compensate on that track: render the result offline and use the rendered audio (see *Latency and your DAW*).

I hear myself late when I record. You're monitoring through the plugin. Put it on the song's track only and monitor your input on its own track, or bounce the karaoke track first.

I can still hear the singer faintly. Removal bottoms out about 21 dB down, so in a quiet, sparse passage a ghost can remain. Raise STRENGTH (70–85 %, or load *Karaoke Strong* or *Instrumental Max Removal*). Reverb on the vocal, doubled vocals and harmonies are the hardest parts; a higher STRENGTH takes more of them.

The music sounds watery, swirly or dull after removing the vocal. STRENGTH is too high for this song: lower it towards 50 % or below. If the cymbals have lost their air, lower HIGH to 10–14 kHz; if the bass or kick changed, raise LOW to 90–120 Hz.

An instrument disappeared with the vocal. The AI heard it as voice: a lead guitar, a saxophone, a synth lead or a choir pad can sound voice-like. Look at the SEPARATION display: if it shows amber, lower STRENGTH, or limit the range with LOW and HIGH if the instrument sits outside the voice's range.

It hardly removes the voice on my podcast or video. The model was trained on singing; a spoken voice separates only about 5 to 7 dB. Use the *Speech* presets, which turn a voice down or a bed down rather than removing either.

I turned VOCAL down 6 dB but it sounds like less. About 0.8 dB of the vocal is never separated, so the real change is a little smaller than the knob's (−6 on the knob is about −5.3 dB). Turn it further; the *Vocal Balance* presets already do.

The output clips when I turn the vocal up. There's no peak protection. Lower OUTPUT until the OUT peak on the meter bridge stays under 0 dBFS.

The LATE lamp lights, and I hear the vocal leak for a moment. Your computer didn't finish the AI's work in time. Watch AI LOAD: if it's near or over 100 %, use fewer instances of VEIRA Vocal Remover at once, close other heavy plugins, or bounce the track offline (a bounce never runs late) and work with the bounced file.

Does it upload my audio or need the internet? No. The AI model is installed with the plugin and runs on your computer. Like every VEIRA plugin, it only needs the internet through VEIRA Manager, to install and to refresh the licence.

Does it work on a mono file, or on stems? Yes on mono: it's analysed as the same signal on both sides. On a single stem it works too, but there's little to gain: it's made for finished, mixed songs.

The window is too small (or too big) on my screen. Drag its bottom-right corner. It scales from 75 % to 200 % and keeps its proportions; the size is saved with your session.

How do I type an exact value? Right-click the knob (Control-click on a Mac), type the number (or **off**) and press Return.

Do my sessions keep my settings? Yes. Every knob, the selected preset and the window size are saved with your session, and every control can be automated.

Contact: support@veiraaudio.com. Mention your computer (Mac or Windows and its version), your DAW and version, and what you see.



VEIRA Vocal Remover · User Manual · Version 0.2.1

Support: **support@veiraaudio.com** · Web: **veiraaudio.com**

© 2026 VEIRA. Proudly made in the Dominican Republic. VST is a trademark of Steinberg Media Technologies GmbH. Audio Units, Logic Pro and macOS are trademarks of Apple Inc. Pro Tools and AAX are trademarks of Avid Technology, Inc. All other trademarks belong to their owners and are used only to describe compatibility.