REVOX



OPERATOR'S MANUAL

AI VOICE CONVERSION IN YOUR DAW · V0.11 BETA · BORN SECOND

01 · QUICK START — CAPTURE, THEN PLAY

REVOX converts a recorded vocal into another voice, entirely on your machine. It works like VocAlign or Melodyne, not like a live effect: you **capture** a pass, REVOX converts it in the background, then plays the converted voice back in perfect sync with your session.

INSTALL

1. **macOS** — run `REVOX-<version>.pkg`, then double-click **Setup REVOX.command** (downloads the shared AI models, ~700 MB, one time).
   Windows — run `REVOX-<version>-Setup.exe` and tick the model download on the last page.
2. Drop an RVC voice model (`.pth` or `.onnx`) into **Documents / REVOX / Models**. Keep its `.index` file next to it for the INDEX knob.
3. Rescan plug-ins in your DAW. REVOX lives under **Born Second**.

CONVERT A VOCAL

1. Insert REVOX on a vocal track and click the black name-plate to pick a voice. A fresh `.pth` bakes natively in a few seconds, one time.
2. Press **CAPTURE**, play the section through the plug-in, press **STOP**.
3. Watch the CNV lane — PROCESSING runs behind the scenes, then the converted region plays back sample-aligned with your timeline.
4. Tweak **PITCH / INDEX / PROTCT** — regions re-render automatically, fast: the heavy analysis is cached and only the synthesizer runs again.
5. Drag a converted region out of the screen to export its WAV onto a track.

> **Good to know:** the status bar at the bottom always says what the engine is doing. The engine runs in its own isolated process — a misbehaving model can never take your DAW session down.



02 · THE INTERFACE



1. **Name-plate** — shows the loaded voice. Click it to switch between the models in your Models folder (each shows its type, e.g. **[V2 48K]**).
2. **Folder button** — load a model from anywhere on disk, and enter your license key.
3. **Screen** — ORG lane holds your captures, CNV lane the converted takes. Drag a converted region out to export its WAV.
4. **CAPTURE** — arm, record through the track, stop. RE-CAPTURE adds another region; existing ones stay.
5. **Knob row + readouts** — the five voice controls (next page), each with its value on an LCD chip below.
6. **Status bar + OUT** — what the engine is doing right now, and the output level fader with meter.

> **Right-click anywhere** on the faceplate for 1x / 2x / 3x interface size, version info and license entry.
> **BYPASS** A/Bs against the dry vocal instantly.

03 · CONTROLS

CONTROL	WHAT IT DOES
PITCH	Shifts the converted voice in semitones (±24). Match the model's natural register: +12 is typical going from a low source voice to a high target, 0 for like-to-like. Re-renders are quick — experiment.
FORMNT	Formant tilt (experimental in this beta — subtle; a full formant engine is on the roadmap).
INDEX	Retrieval blend. Pulls each sound toward the target voice's training timbre using the model's `.index` file: higher = more of the target's character, lower = more of your articulation. Needs the `.index` next to the model.
PROTCT	Consonant protection. Below 0.5 it keeps your original consonants and breaths crisp where the index blend tends to smear them. 0.5 = off (RVC convention); 0.2–0.35 is a good starting zone when INDEX is high.
DRY/WET	Blends the converted voice with your original performance. 100% = converted only.
OUT	Output level (−24 to +12 dB) with meter.
CAPTURE	Arms recording; audio passing through the plug-in is captured at timeline position. Press again to stop and start conversion.
BYPASS	Instant A/B against the unprocessed track.

STATUS BAR DECODER

MESSAGE	MEANING
`BAKING V2 48K MODEL (NATIVE)`	One-time conversion of a fresh `.pth` — a few seconds.
`EXTRACTING INDEX (NATIVE)`	One-time read of the model's `.index` for the INDEX knob.
`... + COREML / + DIRECTML (ISOLATED)`	Hardware acceleration active; engine running in its own process.
`DEMO: FIRST 15S CONVERTED`	No license yet — enter your key via the folder button.
`ENGINE CRASHED - RETRYING ON CPU`	The isolated engine hit trouble and restarted safely. Your session is fine.

04 · VOICE MODELS, LICENSE & SPECS

VOICE MODELS

REVOX uses standard community **RVC v2** models — the format trained and shared everywhere. Drop them in **Documents / REVOX / Models**:

.pth — baked natively in seconds on first use. No Python, no setup. 32k / 40k / 48k, f0 models.
.onnx — used directly.
.index — keep it beside its model to power the INDEX knob.

V1 models (the pre-2023 format) are detected and declined. Rare non-standard models fall back to an optional Python helper (see the README).

> **Voice consent:** only convert to voices you own or have the rights to use. No voice models ship with REVOX — the voices, and the responsibility, are yours.

LICENSE

Unlicensed, REVOX converts the first 15 seconds of each capture. Click the **folder button → Enter license key...** and paste the key from your purchase email. One online check, then it's yours offline.

UNDER THE HOOD

FORMATS	VST3 + AU (macOS), VST3 (Windows), 64-bit
ENGINE	RVC architecture, ONNX Runtime, isolated engine process
PITCH	RMVPE neural pitch tracking
ACCEL	CoreML (Apple Silicon GPU/ANE) / DirectML (any DX12 GPU), automatic CPU fallback
PRIVACY	All conversion on-device. Audio never leaves your machine.

TIPS & BETA LIMITS

» First render after choosing a model takes longest (one-time bake + warm-up). Knob tweaks after that are much faster.

» Captures live in **Documents / REVOX / Sessions** and reopen with your DAW project.

» Mono processing (inputs are summed) and roughly real-time render speed on CPU are the current beta limits.

» Found weirdness? Send your DAW name and what the status bar said — screenshots help.

