ROBOT VOCAL

USER MANUAL

Robot vocals from typed lyrics, every syllable on the beat.



Version 0.1.0 · Made for Ableton Live 12 · Windows, macOS and Linux · lloydstellar.nl

1. WHAT ROBOT VOCAL DOES

Robot Vocal is an instrument plug-in. You type lyrics, and it raps them with a robot voice, every syllable on the beat of your song.

Each line of the lyrics has its own key. A MIDI note on that key plays the line at the tempo of your Live set. Robot Vocal can make the MIDI clip for you: one note for every line, ready to drag into Live.

It comes with 49 voices, 17 flows (ways to place the syllables in the bar) and effects such as vocoders, talkboxes and delays.

2. INSTALLATION

Robot Vocal is a VST3 plug-in on Windows, a VST3 and an Audio Unit on macOS, and a VST3 and an LV2 on Linux. Installing it means putting the plug-in in the folder where your DAW looks for plug-ins.

WINDOWS

1. Close Ableton Live.
2. Unzip the download.
3. Copy the folder `Robot Vocal.vst3` to `C:\Program Files\Common Files\VST3`. Windows asks for permission to copy to this folder: click **Continue**.
4. Start Live. Open **Settings** (Ctrl+,), go to **Plug-Ins** and turn on **Use VST3 Plug-In System Folders**.
5. If Robot Vocal does not show up under **Plug-Ins** in Live's browser, click **Rescan** in the same settings.

> `Robot Vocal.vst3` is a folder, not a single file. Copy the whole folder: the ESPEAK NG voices read their data from inside it.

MACOS

The download works on Apple Silicon and Intel Macs with macOS 11 or later. Robot Vocal is not yet signed by a registered Apple developer, so macOS blocks it until the install script has run. The script takes care of that.

1. Close Ableton Live.
2. Unzip the download. You get a folder with `Robot Vocal.vst3`, `Robot Vocal.component`, `install.command` and `README.txt`.
3. Open **Terminal** (Applications > Utilities > Terminal).
4. Type `bash` and a space, drag `install.command` from the folder into the Terminal window, and press Return.
5. Start Live. Open **Settings** (Cmd+,), go to **Plug-Ins** and turn on **Use VST3 Plug-In System Folders** and, for the Audio Unit, **Use Audio Units**. Click **Rescan** if Robot Vocal does not show up.

The script copies the plug-ins to these folders in your home folder. You can also copy them there yourself:

PLUG-IN	FOLDER
Robot Vocal.vst3	~/Library/Audio/Plug-Ins/VST3
Robot Vocal.component	~/Library/Audio/Plug-Ins/Components

> To open the Library folder in Finder, choose **Go > Go to Folder** (Cmd+Shift+G) and type `~/Library/Audio/Plug-Ins` . If you copy the plug-ins yourself, run this in Terminal afterwards, or macOS will not load them:
> `xattr -dr com.apple.quarantine ~/Library/Audio/Plug-Ins/VST3/"Robot Vocal.vst3" ~/Library/Audio/Plug-Ins/Components/"Robot Vocal.component"`

LINUX

The download works on 64-bit Linux on an Intel or AMD processor, from 2022 or later: Ubuntu 22.04, Linux Mint 21, Debian 12, Fedora 36 or newer, or a rolling distribution such as Arch. ARM computers, such as a Raspberry Pi, cannot run it.

Live does not run on Linux. Robot Vocal works in DAWs that load VST3 plug-ins, such as Bitwig Studio, Reaper, Ardour, Renoise and Tracktion Waveform, or LV2 plug-ins, such as Ardour, Carla and Qtractor. Chapter 6 says what DRAG MIDI CLIP gives you there.

1. Close your DAW.
2. Unpack the download. You get a folder with `Robot Vocal.vst3` , `Robot Vocal.lv2` , `install.sh` and `README.txt` .
3. Open a terminal in that folder and type `bash install.sh` .
4. Start your DAW. If Robot Vocal does not show up, let the DAW scan for new plug-ins.

The script copies the plug-ins to these folders in your home folder. You can also copy them there yourself:

PLUG-IN	FOLDER
Robot Vocal.vst3	~/.vst3
Robot Vocal.lv2	~/.lv2

> Robot Vocal needs ALSA, FreeType, Fontconfig and the X11 libraries. Every Linux desktop has them. On Wayland, its window opens through XWayland, which most desktops start on their own. Both plug-ins are folders: copy them whole, as the ESPEAK NG voices read their data from inside them. In a DAW that loads both formats, such as Ardour, use one of the two.

3. QUICK START

1. Make a MIDI track in Live and drag Robot Vocal onto it from **Plug-Ins** in the browser. The panel opens with four lines of example lyrics.
2. Type your own lyrics on the **WORDS** page: one rap line per line.
3. Choose a **FLOW** and a **VOICE** with the arrows next to their displays.
4. Press **PREVIEW** while Live stands still. You hear your lyrics on an example beat at the tempo of your set.

5. Drag the amber **DRAG MIDI CLIP** key into the arrangement, onto the Robot Vocal track.
6. Press play in Live.

4. HOW THE LINES PLAY

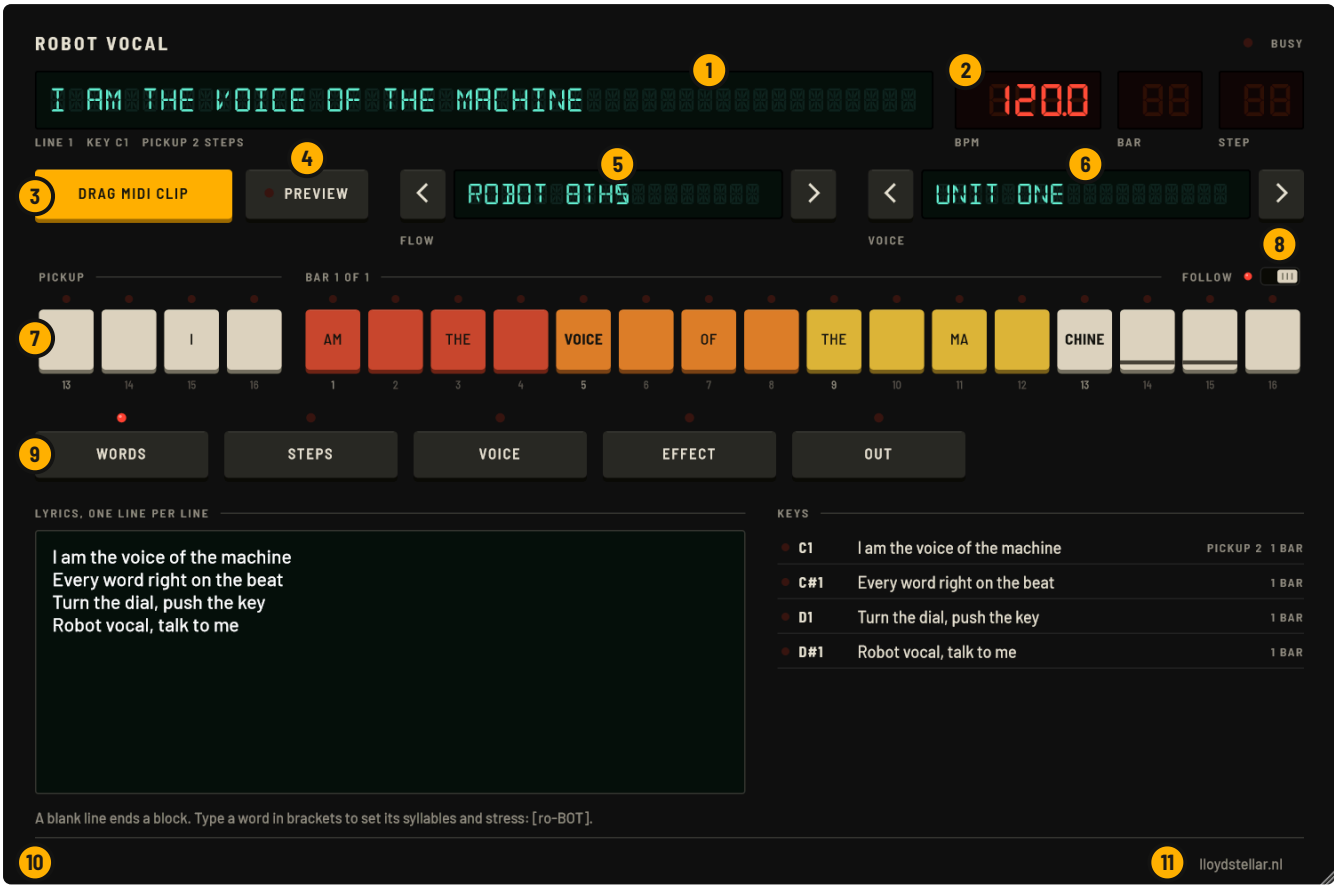
Every line of the lyrics is a key. The first line is C1, the second C#1, the third D1, and so on up the keyboard. You can move the first key on the OUT page.

- **A line starts on a bar line.** A note plays its line from the first bar line after the note starts.
- **Start the note a beat early for the pickup.** Many lines start with a short word before the bar line, such as "I" in "I am the voice of the machine". This is the pickup. If the note starts a beat before the bar line, the pickup sounds there. If the note starts right on the bar line, the line plays without a pickup: the flow fits all its words in the bar.
- **Let go to stop.** When the note stops, the line finishes the syllable it is in.
- **Hold on to go on.** If you hold a note past the end of its line, the next line follows, up to a blank line in the lyrics.

On the step row, the four keys under **PICKUP** are the last four 16th notes of the bar before the line. The keys in the color groups are the 16 steps of the bar.

> Robot Vocal reports a latency of 150 ms to Live. This lets the consonants before a beat sound before it. Live compensates for it automatically.

5. THE PANEL

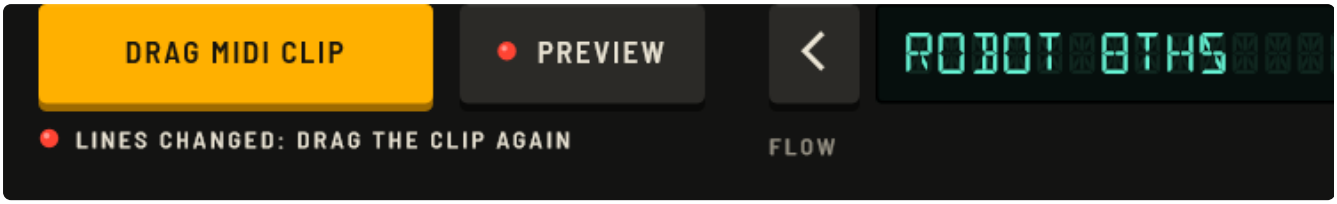


The panel, with the WORDS page open.

1. **Line display**: The line that plays, or the line you edit, with its key and its pickup.
2. **BPM, BAR and STEP**: Live's tempo, and the bar and step of the line that plays.
3. **DRAG MIDI CLIP**: Drag the song into Live as a clip with one note per line.
4. **PREVIEW**: Hear the song on an example beat, without a clip.
5. **FLOW**: How the syllables are placed in the bar.
6. **VOICE**: The voice that raps.
7. **Step row**: The syllables of one bar of the line, with the pickup in front.
8. **FOLLOW**: The step row follows what plays, or stays on the line you edit.
9. **Page keys**: WORDS, STEPS, VOICE, EFFECT and OUT open the pages below them.
10. **Help line**: Point at a control and this line says what it does.
11. **lloydstellar.nl**: The maker's website. Click it to open it in your browser.

To make the window bigger or smaller, drag its bottom right corner, or choose a size on the OUT page.

6. THE MIDI CLIP AND PREVIEW



DRAG MIDI CLIP and PREVIEW. The lamp under the drag key says the clip in Live no longer matches the lyrics.

DRAG MIDI CLIP

Drag this key into the arrangement or into a clip slot. You get a clip with one note for every line. The clip starts with one bar of silence, in which the pickup of the first line comes in. After that, the clip loops the song from the bar line of the first line. The pickup of the first line is also in the last beat of the loop, so the loop runs on without a gap.

To make the song repeat in the arrangement, drag the right edge of the clip further to the right.

When you change the lyrics, the flow or the bars of a line, the lines may move. Then a lamp lights under the key with **LINES CHANGED: DRAG THE CLIP AGAIN**. Drag the clip again and delete the old one.

IN OTHER DAWS

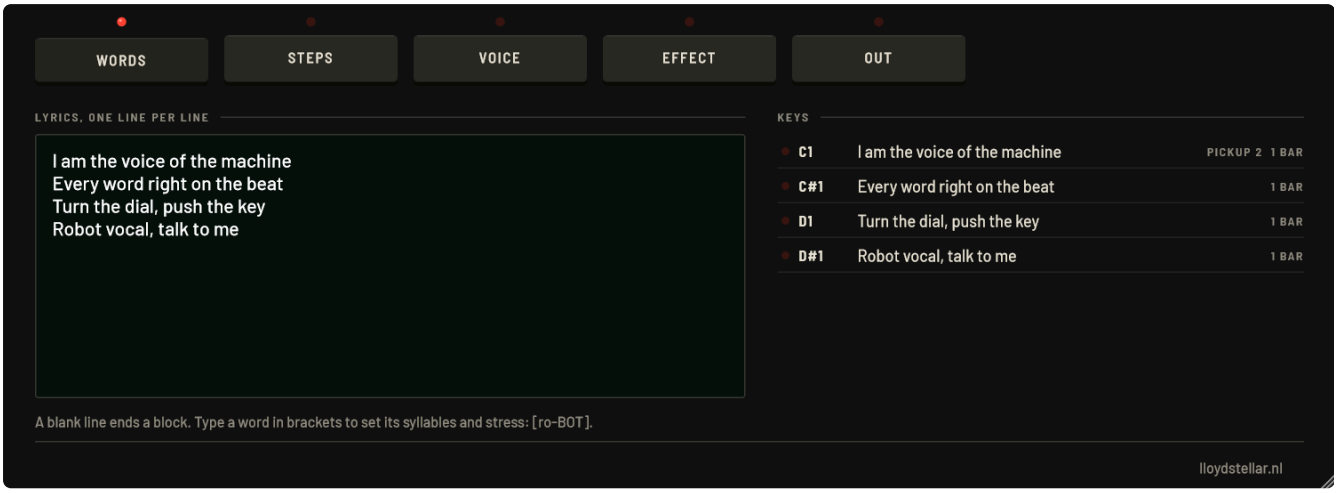
Outside Live, such as in Bitwig Studio, Reaper or any DAW on Linux, DRAG MIDI CLIP gives you a MIDI file with the same notes. Drag it onto the Robot Vocal track. It starts with the bar of silence and the first pickup, then plays the song once. A MIDI file cannot say where a loop starts, so to repeat the song, loop the clip from its second bar. The lamp under the key works the same.

PREVIEW

Press **PREVIEW** to hear the song without dragging a clip. Robot Vocal plays the lines as the clip would, at Live's tempo, round and round, on an example beat that suits the flow. Change the flow or the voice and you hear the result as soon as it is ready.

PREVIEW only works while Live stands still. It stops when you press play in Live. The example beat is never part of an export.

7. THE WORDS PAGE



The WORDS page.

1. Type your lyrics in the field on the left: one rap line per line.
2. Leave a blank line between verses. A blank line ends a block: a held note does not go on past it.
3. If a word is split into syllables the wrong way, type it in brackets with a hyphen between the syllables and the stressed syllable in capitals, for example `[ro-BOT]`.

The **KEYS** list on the right shows the key of every line, and how it is placed: its pickup, its bars, **MOVED** when you moved syllables by hand, the name of its own flow if it has one, and **END** at the end of a block. Click a line to see it on the step row.

8. THE STEPS PAGE



The STEPS page with the syllable "voice" chosen on the step row.

On this page you change one syllable, or one line.

A SYLLABLE

1. Click a syllable on the step row. A ring shows it is chosen, and its name, step and note appear at the top of the page.
2. **EARLIER** and **LATER** move it one step. The other syllables of the line stay where they are. You can also click an empty step: the chosen syllable moves there, as long as it fits between the syllables before and after it.
3. **SHORTER** and **LONGER** set how many steps it sounds. The bar under the steps after a syllable shows how long it sounds. A syllable can sound up to the next syllable. For a longer note, move the next syllable later or give the line more bars.
4. **STUTTER** repeats the start of the syllable 2, 4 or 8 times: ro-ro-ro-robot.
5. Click a key on the keyboard to give the syllable its own note. The dot marks the song's note. **BASE NOTE** puts the syllable back on it.

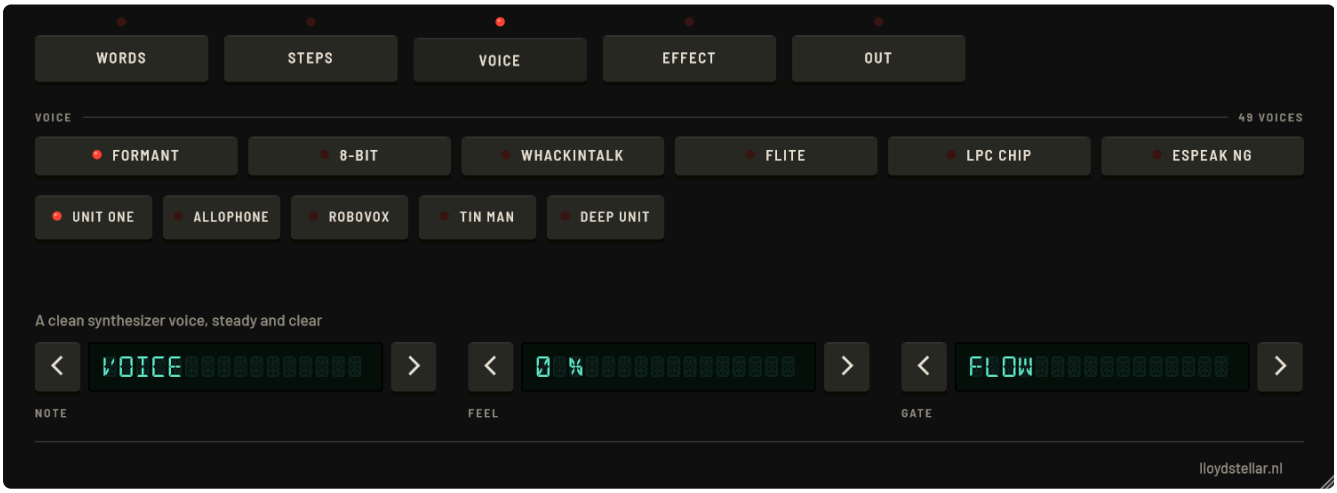
> On its own, a flow lets a syllable sound 2 steps at most, and 3 at the end of a phrase. Only LONGER makes it longer.

A LINE

- **1 BAR** to **4 BARS** spread the line over that many bars. The lit key shows the bars the line has now.
- **LINE FLOW** gives the line a flow of its own. **SONG FLOW** uses the flow on the panel.
- **REFLOW** lets the flow place the line again. Its lamp is lit while syllables of the line are moved by hand.
- **LINE** and **BAR OF THE LINE** choose which line, and which bar of it, the step row shows.

The **FOLLOW** switch above the step row decides what the step row shows while Live plays. On, it follows the line and bar that play. Off, it stays on the line you are editing.

9. THE VOICE PAGE



The VOICE page with the FORMANT bank.

The voices come in six banks, one for each kind of voice machine. Click a bank, then click a voice. The line under the voices tells you what it sounds like.

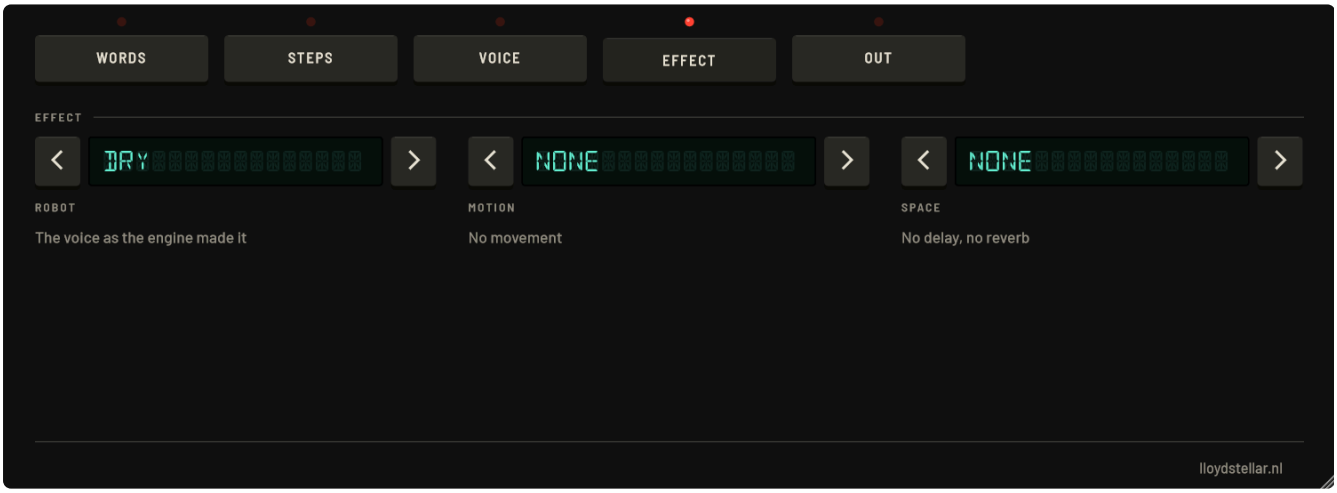


The ESPEAK NG bank.

- **NOTE** is the note the voice raps on. **VOICE** uses the note the voice was made for.
- **FEEL** moves each vowel a little after its step, as a share of the consonants before it. At 0 % the vowel is exactly on the step, where robot voices sound best.
- **GATE** sets how long syllables sound within their steps. **FLOW** uses what the flow does.

The VOICE display on the panel also chooses the voice: use the arrows, or click the display for a list by bank.

10. THE EFFECT PAGE

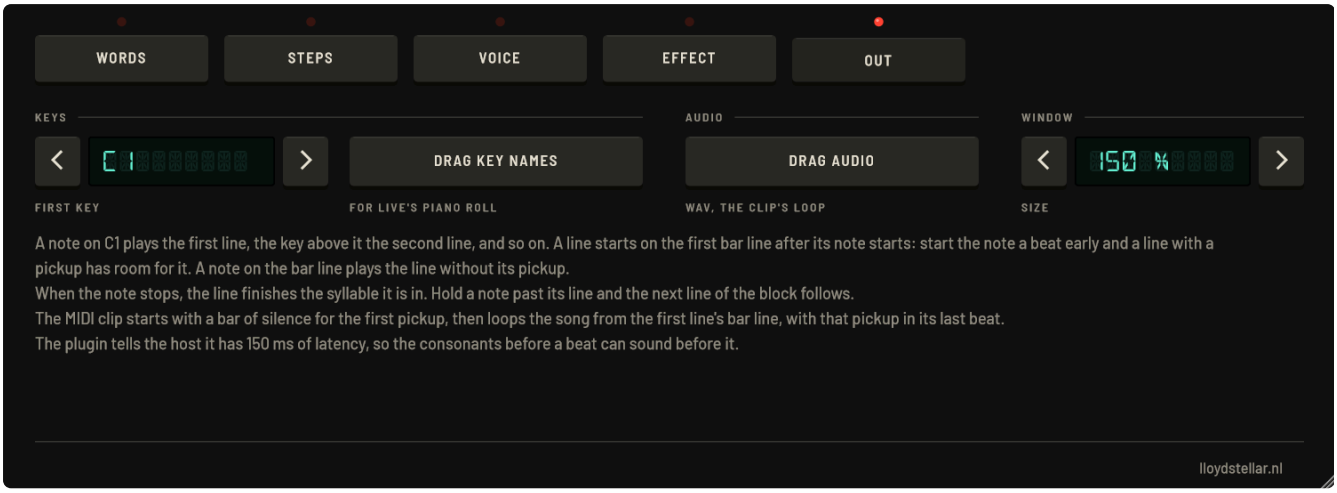


The EFFECT page.

- **ROBOT** is the effect on the voice itself: vocoders, talkboxes, a speech chip and more. The vocoders and talkboxes play on the note set on the VOICE page.
- **MOTION** comes after the robot: ensembles, filter sweeps, beat repeats, panning.
- **SPACE** is the room around the voice: delays on the beat and plate reverbs.

The tables at the end of this manual describe every effect.

11. THE OUT PAGE



The OUT page.

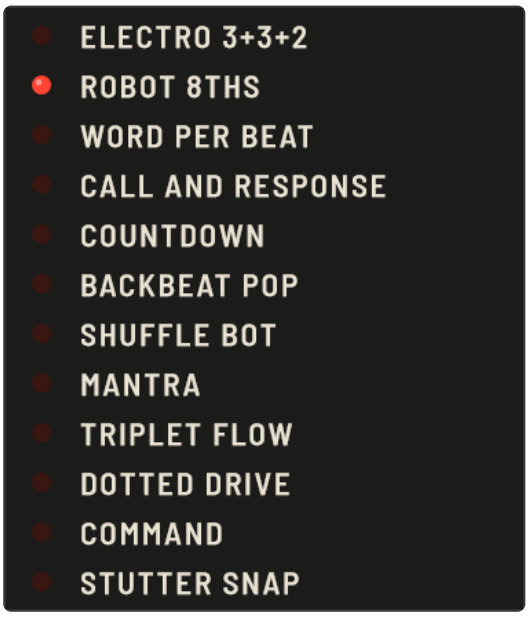
- **FIRST KEY** is the key that plays the first line. The keys above it play the lines after it.
- **DRAG KEY NAMES** makes a MIDI Effect Rack for Live. Drag it onto the MIDI track, to the left of Robot Vocal. Live's piano roll then shows the words of each line on its key. Drag it again after you change the lyrics. Other DAWs cannot open the rack, so there the key is grey.
- **DRAG AUDIO** gives you the song as a WAV file at the tempo of your set. It is the loop of the MIDI clip, starting on the bar line of the first line. Drag it onto an audio track.
- **SIZE** sets the size of the window.

12. CHOOSING FROM A LIST

Click a display, such as the FLOW display, to open its list. The flows come in four groups by tempo. Point at a group to see its flows, then click one. The lit lamp marks the current choice.



The FLOW list, by tempo.



The flows from 120 to 145 BPM.

13. WHEN SOMETHING DOES NOT WORK

PROBLEM	WHAT TO DO
Robot Vocal does not show up in Live	Check that the plug-in is in the right folder (chapter 2). In Live's Settings, Plug-Ins, turn on the system folders and click Rescan. On a Mac, run the install script.
The ESPEAK NG voices are silent	The plug-in folder was not copied whole. Copy the whole Robot Vocal.vst3 or Robot Vocal.lv2 folder again.
Robot Vocal does not load on Linux	It needs a Linux from 2022 or later (glibc 2.35). Type ldd —version in a terminal: the first line ends with the version.
The lines do not fall on their bars	The clip in Live is older than the lyrics. Drag the clip again when the lamp under DRAG MIDI CLIP is lit.
PREVIEW makes no sound	Stop Live first. PREVIEW only plays while Live stands still.
A word sounds wrong	Type it in brackets with its syllables and stress, such as [ro-BOT].
A note sounds too long	Choose the syllable on the STEPS page and press SHORTER.

14. FLOWS

The BPM column is the tempo range a flow suits best, though every flow works at any tempo. Syllables is how many syllables the flow puts in a bar.

UNDER 100 BPM

FLOW	BPM	SYLLABLES	HOW IT GOES	THINK OF
16TH RAPID FIRE	80-100	12-16	A syllable on almost every 16th note.	Blazing Arrow
BOOM BAP LAID BACK	84-100	8-13	Dense, with swing, a little behind the beat.	90s boom bap
PICKUP FLOW	85-100	8-12	Lines that start on the last 16ths of the bar before.	99 Problems

100-120 BPM

FLOW	BPM	SYLLABLES	HOW IT GOES	THINK OF
OLD-SCHOOL COUPLET	95-115	6-9	Early rap: even lines with the rhyme on beat 4.	Kurtis Blow

120-145 BPM

FLOW	BPM	SYLLABLES	HOW IT GOES	THINK OF
ELECTRO 3+3+2	105-140	6-10	Syllables in groups of 3, 3 and 2 16th notes: the electro bounce.	Egyptian Lover, Cybotron
ROBOT 8THS	110-140	4-8	One syllable on every 8th note, flat and even.	The Robots, Pocket Calculator
WORD PER BEAT	115-140	4-8	Every word starts on a beat. Good for lists and commands.	Numbers, Technologic
CALL AND RESPONSE	115-140	4-7	A line in one bar, its answer in the next.	Planet Rock
COUNTDOWN	128-136	3-5	One word per beat, for counts and lists.	Numbers, Expo 2000, Computer World
BACKBEAT POP	128-138	4-6	Words around beats 2 and 4, so the kick has beats 1 and 3.	Planet Rock
SHUFFLE BOT	128-140	5-8	Swung 8ths on a triplet grid.	Pack Jam, Egypt Egypt
MANTRA	128-142	5-8	A short hook every half bar, each with a pickup.	It's More Fun to Compute, Musique Non Stop
TRIPLET FLOW	120-150	9-12	Three syllables per beat.	Versace, Bad and Boujee
DOTTED DRIVE	130-142	5-6	Syllables a dotted 8th apart (3+3+3+3+2+2).	Aux 88, Anthony Rother
COMMAND	130-142	3-6	Short commands on beats 1 and 2, then a long rest.	Jam On It, Destroy Him My Robots

FLOW	BPM	SYLLABLES	HOW IT GOES	THINK OF
STUTTER SNAP	130–150	6–8	Short snaps on each beat.	The Scotch snap: short, then long

145+ BPM

FLOW	BPM	SYLLABLES	HOW IT GOES	THINK OF
HALF-TIME TRAP	130–160	6–10	A half-time feel over two bars.	Trap, UK drill

15. VOICES

BANK	VOICE	DESCRIPTION
FORMANT	**UNIT ONE**	A clean synthesizer voice, steady and clear
FORMANT	**ALLOPHONE**	Jumps from sound to sound like a 1970s speech chip
FORMANT	**ROBOVOX**	A sung monotone with long vowels
FORMANT	**TIN MAN**	A small, bright robot
FORMANT	**DEEP UNIT**	A big, low machine
8-BIT	**HOME MICRO**	Three tones restarted on every pulse, 4-bit: the software voice of the early home computers
8-BIT	**STUFFED**	A low mouth and a blocked nose
8-BIT	**TINY BOT**	A very small robot
8-BIT	**VISITOR**	A long throat and a tiny mouth, from far away
8-BIT	**CRISP MICRO**	The same three tones in 8 bits and twice the steps: cleaner, still a machine
WHACKINTALK	**DESKTOP**	Low, level and a little nasal, like a desktop computer of 1984
WHACKINTALK	**MONOTONE**	A machine: dead flat, no flutter, a pressed buzz
WHACKINTALK	**GRAVEL**	Deep, with some gravel in it
WHACKINTALK	**KID**	A short vocal tract, young and bright
WHACKINTALK	**ALTO**	A higher register and a breathier source
WHACKINTALK	**TREBLE**	Young, high and bright
WHACKINTALK	**WHEEZE**	Low, wheezy and a little unwell
WHACKINTALK	**ALIEN**	A tiny vocal tract and a hard buzz
WHACKINTALK	**HUSH**	Breath through the formants, no voice at all
WHACKINTALK	**UNDERWATER**	Muffled, as if under water
WHACKINTALK	**BLEAT**	A sheep: nearly all of it is pitch flutter
WHACKINTALK	**UNHINGED**	Jumps all over the place, with a bad wobble
WHACKINTALK	**GIGGLE**	Very high and far too pleased, jumping up
WHACKINTALK	**ARPEGGIO**	Sings a rising arpeggio over your notes
WHACKINTALK	**FANFARE**	Sings a bright major fanfare, with vibrato
WHACKINTALK	**MOUNTAIN KING**	Sings Grieg's Mountain King figure, dark, with vibrato
WHACKINTALK	**DIRGE**	Sings a funeral march in minor, low and resigned
WHACKINTALK	**ORGAN**	Sings a solemn line that steps down and back
FLITE	**KAL**	The diphone voice of 1990s speech software, 8 kHz
LPC CHIP	**SPELLER**	A talking spelling toy of 1978: ten poles at 8 kHz, the chirp, 8 bits

BANK	VOICE	DESCRIPTION
ESPEAK NG	**ROBOSOFT**	A buzzing sawtooth formant voice
ESPEAK NG	**ROBOSOFT 2**	Robosoft, rounder
ESPEAK NG	**ROBOSOFT 3**	Robosoft, higher and thinner
ESPEAK NG	**ROBOSOFT 4**	Robosoft, rough
ESPEAK NG	**ROBOSOFT 5**	Robosoft, deep
ESPEAK NG	**ROBOSOFT 6**	Robosoft, nasal
ESPEAK NG	**ROBOSOFT 7**	Robosoft, bright
ESPEAK NG	**ROBOSOFT 8**	Robosoft, hard
ESPEAK NG	**UNIROBOT**	A flat machine voice
ESPEAK NG	**ANIKA ROBOT**	A female robot
ESPEAK NG	**KLATT 1**	The Klatt synthesizer, impulse source
ESPEAK NG	**KLATT 2**	Klatt with a natural glottal source
ESPEAK NG	**KLATT 3**	Klatt with a sampled source
ESPEAK NG	**KLATT 4**	Klatt with a sampled source, brighter
ESPEAK NG	**KLATT 5**	Klatt with a sawtooth source
ESPEAK NG	**WHISPER**	Only breath, no voice
ESPEAK NG	**CROAK**	A rough, broken voice
ESPEAK NG	**STANDARD M**	eSpeak's plain male voice
ESPEAK NG	**STANDARD F**	eSpeak's plain female voice

16. EFFECTS

ROBOT

ROBOT	DESCRIPTION
DRY	The voice as the engine made it
BRAZEN HEAD I	Twenty-band vocoder on a raw sawtooth, some bare voice, coarse steps
BRAZEN HEAD II	The vocoder on a choir of detuned minor triads
BRAZEN HEAD III	The vocoder on a sub bass an octave and two below
FLAT TWENTY	Wider bands, clearer words
SIXTEEN CHANNEL	Sixteen bands from 50 Hz to 5 kHz on a pulse wave
SUPERSAW CHOIR	Seven detuned saws, wide and bright
MONOTONE UNIT	Any voice on one note, clear and level
MOUTH TUBE	A talkbox on a sawtooth synth
TALK TUBE	Talkbox with the hiss of the voice for clear consonants
SPELLING CHIP	Ten-pole LPC at 8 kHz, the way a 1978 talking toy spoke
EXTERMINATOR	Ring modulation at 30 Hz, lifted mids, driven
CENTURION	A pulse-wave vocoder through a slow phaser, dark
PHASE ROBOT	A phaser that sweeps once a bar
JET	A flanger sweeping over two bars
TIN CAN	Telephone band, a metal ring and coarse steps
TIGHT DELAY	A very short ringing echo and a slap one 16th later
STUTTER	A square-wave vocoder cut in 16ths on the beat
BASS DROID	A vocoder on a square an octave down, driven
WHISPER UNIT	A vocoder on noise: words without pitch
EIGHT BIT	Eight bits at 11 kHz, nothing more

MOTION

MOTION	DESCRIPTION
NONE	No movement
ENSEMBLE	Three slowly drifting copies, as the choir ensembles of 1979
FILTER SWEEP	A resonant filter that opens and closes every two bars
FILTER WAH	The filter sweep once a bar, narrower
BEAT REPEAT	The last beat of every two bars stutters in 16ths
BEAT REPEAT 8	A roll over the last half bar of every four bars

MOTION	DESCRIPTION
TAPE STOP	The last word of every line slows to a standstill
WIDE CHORUS	Two drifting copies, one left and one right. Stereo
AUTO PAN	The voice moves from side to side once a bar. Stereo
AUTO PAN 1/2	From side to side every two beats. Stereo

SPACE

SPACE	DESCRIPTION
NONE	No delay, no reverb
DELAY 3/16	Echoes a dotted 8th apart, the classic electro bounce
DELAY 1/4	Echoes on every beat
DELAY 1/8	Quick echoes on the 8ths, tucked behind the voice
DUB DELAY	Long, dark repeats on the beat that slowly fall apart
ECHO THROW	Only the last word of every line echoes, as in rap records
SMALL PLATE	A short, bright plate
BIG PLATE	A long plate for held words
GATED PLATE	A big plate cut off an 8th after the voice, 1980s style
DELAY AND PLATE	Dotted 8th echoes in a small plate
REVERSE SWELL	Backwards reverb that grows into every line
PING-PONG 3/16	Dotted 8th echoes from left to right. Stereo
PING-PONG 1/8	8th echoes bouncing between the sides. Stereo
STEREO PLATE	A wide plate around the voice. Stereo