

AI Voice & Audio with ElevenLabs — Workbook

This workbook turns the course into reps. Each section matches a course module and gives you exercises to run inside ElevenLabs, worksheets to capture your decisions, and checklists to keep your spending, quality, and ethics in line. Work through it with ElevenLabs open in another tab — the goal is a finished, on-brand voiceover and a reusable voice library you keep for every future project.

Getting Started with ElevenLabs

Set up your account, learn the character-credit math, and complete your first full generation loop.

Exercise: Run and Compare Two First Voiceovers

In Eleven Multilingual v2 with a default voice like Rachel or Adam, generate the course starter script twice. Listen to each twice before judging. Note how the two non-identical reads differ — this proves why you must budget for multiple takes.

- Generate: Welcome to the channel. Today we are going to break down three simple habits that will completely change how you manage your time. Let us get started.

- Which of your two takes is better, and write one sentence on exactly why it won (naturalness, pacing, pronunciation, or emotion)?

- How many characters did this script use, and what is your remaining monthly credit balance?

Worksheet: Project Character Budget

Fill this in before you start any real project so you never stall mid-build with an empty balance. Use roughly 1,000 characters per minute of audio, 1 credit per character in Multilingual v2, and about half that in Flash or Turbo.

Project name and target length (e.g. 3-minute explainer voiceover)

Total script character count (paste into a character counter)

Expected regenerations of tricky sections (assume 2 to 4)

Draft model and cost per pass (Flash, approx 0.5 credit per character)

Final model and cost per pass (Multilingual v2, 1 credit per character)

Total estimated characters for the project

Hard character cap for this project (stop and review when reached)

Checklist: Account and Cost Readiness

- [] Created an ElevenLabs account and confirmed the current plan character allowance on the pricing page
- [] Located the model selector in the Text to Speech tool and identified Multilingual v2, Turbo, Flash, and v3
- [] Found the usage panel showing remaining monthly credits
- [] Completed one full text-to-speech generation end to end
- [] Saved both starter clips into a named project folder for later reuse

Voices and the Settings That Control Them

Cast the right voice from the Voice Library and master the four settings that shape every read.

Exercise: Audition a Voice Shortlist on Your Real Script

Open the Voice Library and filter for voices that fit a project you actually make. Add three or four to your VoiceLab, then generate the same real sentence from your script with each. Cast the winner the way you would cast a narrator.

- Which real sentence from your script did you audition (use your words, not hello)?

-
- List the three or four voices you tested and one note on each (accent, age, energy, fit)

-
- Which voice did you cast and why does it match the content and audience?
-

Exercise: Hear What Stability Does

Take your baseline starter clip and regenerate it at three Stability settings — low, medium, and high — keeping voice, Similarity, Style, and Speaker Boost fixed. Listen to the three back to back to train your ear on the slider.

- Describe how the low-Stability read differed from the high-Stability read

-
- At which Stability value did the read sound most reliable and professional?

-
- Did any setting introduce artifacts, wandering tone, or flatness — at which value?
-

Worksheet: Voice Settings Recipe Card

Lock a repeatable recipe for one cast voice. Record the exact values so you reproduce the identical read in any later session and any patched line matches.

Voice name (as saved in your VoiceLab) and source (default, Library, cloned)

Model (Multilingual v2, Turbo, Flash, or v3) — keep it fixed for the project

Stability value (e.g. 55%)

Similarity value (e.g. 75%)

Style Exaggeration value (keep low) and Speaker Boost (on/off)

Use case this recipe is cast for (e.g. Doc Narrator, Brand Warm)

Checklist: Consistency Quality Gate

- Cast one voice per speaker and committed to it for the whole project
- Recorded the exact Stability, Similarity, Style, and Speaker Boost values
- Kept the same model across every clip so reads match
- Generated in coherent chunks (paragraph or section) rather than line by line
- Regenerated any fix with identical voice and settings so seams disappear

Voice Cloning and Directing Delivery

Clone a voice the right way, stay inside the consent rules, and direct pace, emotion, and pronunciation from the script.

Exercise: Instant-Clone Your Own Voice

Record two to three minutes of clean, natural reading in a quiet, non-echoey room with a decent mic. Create an Instant Voice Clone of your own voice and generate the starter script with it. Hear text you never spoke in your own voice.

- How clean was your source recording (room noise, echo, clipping) and what would you fix next time?

-
- How faithful is the Instant clone to your real voice on a scale of 1 to 5?

-
- For your real projects, would Instant cloning be enough, or do you need Professional cloning — why?

Exercise: Direct One Paragraph With Punctuation and Tags

Take one paragraph of your script and make a direction-rich version: add deliberate commas and dashes for pacing, fix any tricky name or acronym with a phonetic respelling, and if you are in v3, add one or two emotion or audio tags. Generate plain and directed versions and compare.

- What punctuation changes did you make and how did pacing change?

-
- Which word did you respell phonetically (e.g. a brand or name) and did it fix the pronunciation?

-
- If you used audio tags, which ones, and did the directed read beat the plain read?

Worksheet: Voice Consent Record

Before cloning anyone but yourself, document consent. Keep one record per cloned voice so your use is defensible and inside ElevenLabs' rules and the law.

Voice owner full name (whose voice is being cloned)

Relationship and confirmation this is your own voice or you have permission

Specific permitted use (e.g. brand explainers, internal training, this campaign)

Cloning method (Instant or Professional) and verification captcha completed (yes/no)

Date consent given and where the written consent is stored

Disclosure plan (where you will tell listeners the voice is AI-generated)

Checklist: Ethics and Cloning Quality Gate

- Cloned only your own voice or a voice with explicit, documented permission
- Recorded clean, consistent source audio with no background noise or clipping
- Completed the voice-verification captcha honestly for any Professional clone
- Confirmed the use is not deceptive, fraudulent, or impersonating a real person
- Decided how and where to disclose AI-generated voice to the audience

Long-Form, Cleanup, and Exporting for Video and Podcast

Narrate long scripts in Studio and export clean, properly formatted audio ready for your editor.

Exercise: Run a Script Through Studio

Create a Studio project, import a short multi-paragraph script, lock your cast voice and settings, and generate the whole document. Then deliberately regenerate just one paragraph to feel how isolated fixes save time and characters.

- How did per-paragraph regeneration compare to re-rendering the whole script?
-
- Did the long-form read hold together as one performance across paragraphs?
-
- Roughly how many characters did the full script consume against your plan?
-

Exercise: Export Two Formats and Normalize

Export one finished clip twice — once as MP3 at 44.1 kHz and once as PCM WAV at 44.1 kHz. Drop both into your audio or video editor, run a loudness normalization, and compare them.

- Could you hear any difference between the MP3 and the PCM WAV master?
-
- What loudness target did you normalize to (e.g. minus 16 LUFS podcast, minus 14 LUFS video)?
-
- Which format will you use as your delivery file and which as your editing master?
-

Worksheet: Export and Delivery Spec

Lock the export and assembly settings for your project so every clip and episode comes out consistent and platform-ready.

Destination (YouTube voiceover, podcast episode, course video, ad)

Editing master format and sample rate (PCM WAV at 44.1 kHz)

Delivery format and bitrate (MP3 at 44.1 kHz, 192 kbps or higher)

Loudness target for the destination (e.g. minus 16 or minus 14 LUFS)

Music bed level relative to voice (e.g. 15 to 20 dB below) and ducking on/off

Editor used for assembly (CapCut, Premiere, Resolve, Descript)

Checklist: Finished Voiceover Quality Gate

- Produced long-form narration in Studio with one locked voice and settings
- Regenerated only the paragraphs that needed fixing, keeping settings identical
- Exported a clean PCM WAV master at 44.1 kHz for editing

- [] Placed a music bed 15 to 20 dB below the voice and ducked it under narration
- [] Ran a final loudness normalization on the full mix to the platform target
- [] Produced one complete voiceover end to end as a portfolio piece

Your Action Plan

1. Create your ElevenLabs account and confirm the current plan's character allowance before planning any project
2. Run the starter script twice in Multilingual v2 to internalize that identical text produces different reads
3. Fill in the Project Character Budget worksheet and set a hard character cap for your first real voiceover
4. Audition a shortlist of Voice Library voices on a real sentence and cast one per speaker
5. Tune Stability, Similarity, Style, and Speaker Boost one slider at a time and save a recipe card
6. Instant-clone your own voice from a clean two to three minute recording before cloning anyone else
7. Document consent and a disclosure plan before cloning any other person's voice
8. Direct delivery from the script with deliberate punctuation, phonetic respellings, and audio tags where supported
9. Narrate long scripts in Studio and regenerate only the paragraphs that need fixing
10. Export a clean PCM WAV master, assemble it with a ducked music bed, normalize loudness, and finish one piece

