Episode Book

EPISODE #6504

Why Voiceover Studios Obsess Over Repeatability

PUBLISHED	8 November 2026 at 15:56 UTC
DURATION	24:52
LISTEN	myweirdprompts.com/episode/voiceover-studio-repeatability-setup-06405/
HOSTS	Corn, Herman, Hilbert
SCRIPT MODEL	z-ai/glm-5.3
VOICES	chatterbox-regular (chatterbox-local)
PIPELINE	V5.26
SOURCES	25 cited by the research (listed at the end)
BOOK GENERATED	9 November 2026 at 08:09 UTC

SYNOPSIS

Professional voiceover setups are engineered around physical consistency, ensuring recordings made months apart match seamlessly. This deep dive covers microphone positioning geometry, off-axis plosive management, eye-level script placement to eliminate head drift, and why tablet reflections matter. It also details the acoustics of sitting versus standing, hands-free punch-and-roll workflows in Reaper, and why USB foot pedals share the same internal hardware architecture as common handheld gamepads.

TRANSCRIPT



Corn

The Attenborough voice. Everyone agrees it's the gold standard, and the usual conclusion is that the man is a national treasure and that's that. The boring truth is more interesting. That voice is a job. Somebody solved the plumbing around it decades ago, and the solutions are unglamorous and repeatable.



Herman

And Daniel's been listening to him, apparently, and it sent him down exactly this hole. He wants to know how professional narrators and voiceover people physically arrange themselves and their gear, standing versus sitting, mic position, where the script goes so you don't swivel off-axis, how pages get turned without hands, what to do about computer fans, and how a human being survives a six-hour session. And then he names his actual kit, because of course he does. A one and a half metre stand, a reflection filter that wraps around the mic, an AT2020, a Samson Q2U, and a tiny gamepad mapped to record and next. He's recording hundreds of short sentences for a voice clone.



Corn

Which is the detail that quietly rewrites every answer. He asked how the pros lay it out. The pros lay it out so nothing changes between takes. That's the whole thing.



Herman

That's the thesis I'd hang the episode on, honestly. A professional station isn't optimised for beautiful sound in the abstract. It's optimised for repeatability. Same distance, same room, same energy, on hour one and hour forty.



Corn

Which is a sloth-friendly doctrine. My entire physiology is a commitment to one posture.



Herman

I was going to say, you're the most professionally seated creature I know.



Corn

Hours of practice. Let's do the body and the mic first.



Herman

The standing question first, because it's the one everybody has an opinion on. The old rule is you always stand. Always. And it turns out that rule is a fossil. It comes from the old-school announcers, the big booming chest-voice radio men, where standing was supposed to let the celebrated diaphragm do its thing. Modern voiceover is conversational. Nobody's doing the bombastic read anymore, so the reason for the rule evaporated, but the rule survived.



Corn

Like the appendix of vocal technique.



Herman

It is. And the research is a mess in the best way. A Northwestern study from two thousand six found standing gave better lung capacity and expiratory flow than sitting, especially slumped sitting. Then Emory in twenty seventeen concluded head position, forward or back, mattered more than stance entirely. And a Michigan State pilot from twenty twenty-five found no aerodynamic or muscle differences across postures at all, in experienced singers. Three studies, three different answers.



Corn

So the science is telling us it doesn't matter, which means the advice should be pick whatever and stop thinking about it.



Herman

A working studio I read put it perfectly. You can sit or stand, but commit, because different body positions alter your sound. That's the version I'd give Daniel. It's not a posture question, it's a consistency question. Pick one, and never change it mid-corpus.



Corn

It's the same word you'd use about a marriage.



Herman

And there's a real cost to the standing side nobody mentions until you've paid it. Rosie Akerman, a UK voiceover artist, her first booth was a cupboard. Physically no room for a chair. So she stood, and her own words were, not so great if you're recording an audiobook for hours and days at a time. She now has a booth that fits a chair comfortably, and I suspect the chair was the upgrade she was proudest of.



Corn

The glamorous life. Standing in a cupboard for days.



Herman

Days. The stand follows the posture, by the way. Martin Allanson, another UK narrator, uses a swivel-mounted boom arm specifically so he can switch between standing and sitting with no screwing or messing about, and he said something I love. He's changed his mic stand and its positioning more than he's changed his mic.



Corn

That's a quietly radical sentence. The stand is the instrument.



Herman

And the cheap ones betray you. Extendable boom arms sag under mic weight. Jim, a studio guy who writes about gear, calls them dreadfully undersprung in many cases. You set your geometry in the morning and by lunch the mic has quietly sunk two centimetres and you haven't noticed.



Corn

Which for a normal person is an annoyance and for a clone corpus is contamination.



Herman

Right, because the model hears the drift. A TTS model treats everything in the recording as the voice. If your distance creeps, or your room changes between sessions, or your energy is different on a Tuesday, the model averages the variants or oscillates between them. The reference corpus everybody built on, the Edinburgh one, VCTK, had over a hundred speakers each reading about four hundred sentences with an identical setup, same mics, same hemi-anechoic chamber. The setup was the experiment.



Corn

The chamber was the point. Nobody's dream studio is a laboratory, and that's exactly what it was.



Herman

Mic distance, then. For a large-diaphragm condenser, the starting point is fifteen to twenty centimetres, six to eight inches. There's a lovely trick from a voice trainer named Trevor O'Hare. Make a fist, extend your thumb and pinky, the hang-loose sign, put your thumb on the mic, and your pinky lands roughly where your mouth should be. Instant calibrator, no ruler.



Corn

I have the opposite problem. My arms would need a second calibration.



Herman

Your reach is a fixed constant, at least. The dynamic mic is different. Something like an SM7B wants you two to four inches away, practically eating it. So Daniel's two mics want different geometries, and that's a real fork in the road. The AT2020 at a hand's span, the Q2U much closer.



Corn

He should pick one mic for the whole corpus and never touch the other.



Herman

Not just one mic. One posture, one distance, one room, one everything. Then the angle. You don't point the capsule straight at your mouth. You put it just above nose height, aimed down at the lips, or offset an inch or two to the side so you're speaking past it rather than into it. The physics is simple. Plosive air travels straight forward. If the capsule sits outside that airstream, the pop never lands, and your tone barely changes.



Corn

So the pop filter is a seatbelt and off-axis is the driving skill.



Herman

And then the pro move that nobody does at home. You mark your position physically. Gaffer tape on the floor where your feet go. Note exactly where the copy stand sits relative to the mic. One trainer's line was that auditions recorded in March should match pickups recorded in June.



Corn

That's a sentence with a whole career inside it.



Herman

There's a practical calibration too. Record the same sentence at roughly fifteen, twenty and thirty centimetres and listen back for clarity, low-end mud, breath thumps, room sound. Because proximity effect adds bass as you get closer, and more bass is not more clarity, especially across consecutive sentences. A sentence that sounds rich and intimate alone can sound muffled in a batch.



Corn

Which is exactly the batch Daniel's making. He's not recording one beautiful sentence, he's recording four hundred of them in a row.



Herman

And that's where the most common failure shows up. Drift within a take. The first sentence sounds full and present, then the actor turns to follow the script on a screen, or leans back on an exhale, and suddenly the voice is thinner and farther away. Every coaching source names this as the number one amateur tell.



Corn

And notice what that actually is. It's not a mic technique problem. It's a script placement problem. The head moved because the text was somewhere it shouldn't be.



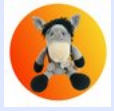
Herman

The script goes at mic height, directly behind or beside the mic, so reading never pulls you off-axis. If it's a tablet or a monitor, raise it to eye level. The failure is looking down at paper and back up at a screen. That changes your distance and your direction without you ever noticing, which is the worst kind of change.



Corn

The studio standard, I'm guessing, is not a stack of paper.



Herman

It's a music stand with an iPad. Printed pages are rare in narration now because they slow you down and they make noise. And the noise was a whole era. C.C. Hogan, a narrator who came up through radio drama, describes his constant battle with page turns. His solution in a directed session was to thump the talk-back button and mumble "paper." The engineer always knew what he meant, and would just rerecord the line with the rustle.



Corn

A one-word protocol. Efficient.



Herman

He also describes radio-drama actors who would lift the page they were reading delicately away in a big silent gesture, then bend down slowly like a ballerina and deposit it on the floor. A whole choreography existed just to move paper quietly.



Corn

Decades of professional training, spent defeating a tree.



Herman

The tablet kills all of it. One detail, though. You angle the tablet slightly so it doesn't reflect your voice back at you. Send the reflection off toward a wall, not toward the capsule. I read about one engineer who lines the capsule up with the top edge of the copy stand so reflections from the stand are geometrically impossible. Sound bounces like light off a mirror, and if the flat surface can't see the mic, it can't hit it.



Corn

There's markup too, isn't there. The script itself gets annotated.



Herman

Hogan's system is beautiful. Single slash for a breath point, double slash for a substantial pause, underline for emphasis, double underline for really important, plus pronunciation notes for the awkward names. And his point is that marking up is rehearsal. A good voice speaks the script out loud while marking it up. You're not preparing to read, you're already performing.



Corn

For Daniel's corpus that markup is basically free quality. Four hundred sentences, all pre-breathed in the same places.



Herman

Now the hands. The dominant pro technique for long-form is punch-and-roll. You make a mistake, you stop, spool back a few seconds, hear yourself, and start recording again at the last clean spot. You're editing while you speak. It's everywhere in the English-language audiobook world, and Reaper is one of the most popular tools for it.



Corn

And you can actually give the concrete setup.



Herman

You install the free SWS extension, set the nudge from its default to five seconds, and build two custom actions. One key stops, moves the cursor back five seconds, and plays, so you hear yourself. The other drops a marker and records. The marker means you can see exactly where you punched in afterwards. Ten minutes to configure, total.



Corn

Ten minutes. Cheaper than the microphone by a factor of everything.



Herman

And because keystrokes can land in the recording, pros put those two actions on a USB foot pedal. The de facto standard is the Infinity three-pedal transcription footswitch, about seventy dollars. Fun fact about that pedal, it's functionally a game controller. It registers as a HID joystick. Which means Daniel's tiny gamepad is not a hack, it's the same class of device the industry already standardized on. A gamepad mapped with a macro tool gives you fifteen, twenty distinct inputs, way more than the three pedals.



Corn

So the professionals spent seventy dollars to invent the controller he already owns.



Herman

He got there from the other direction. There is one caveat worth flagging. Gamepad-to-keystroke mapping through AutoHotkey can require an active AHK window with modern controllers. Test it before a long session, not during one.



Corn

Now the cost of punch-and-roll. Because it can't be free.



Herman

It isn't. The seams. If you re-set a spot two, three, four times, you end up with little snippets in a row because you never punch in at exactly the same place. Clipped breaths, because the take comes a tick too late. A sentence ending lost. A key press or a rustle right at the join. Pauses that don't match. And the deeper cost, part of your head is on the technique instead of the character.



Corn

Here's my heresy, then. Daniel shouldn't punch-and-roll at all.



Herman

I think you're right, and it's a unusual position for this show to land on. Punch-and-roll exists to deliver a clean continuous performance in one pass. Daniel is recording hundreds of discrete short sentences that will be sliced at sentence boundaries anyway. The seams nobody wants in an audiobook are irrelevant in his pipeline. His gamepad's record-and-next mapping, make a file, advance, make a file, advance, is arguably the better design for his job. He'll cut afterwards.



Corn

We did an episode once about whether an ageing voice breaks a fine-tune, and the same logic held there. It's not the voice changing that kills you, it's the voice being inconsistently captured. The corpus is the product.



Herman

There's a frame under all of this that I think explains the whole gear question. The pro studio's biggest luxury is a second person. The engineer drives the machine. If you trip, you pause, the engineer rolls back a few seconds, you pick up at the start of the sentence and match your tone. Every hands-free technique we've discussed, foot pedals, remotes, punch-and-roll macros, is one person buying back that second pair of hands.



Corn

The gamepad is Daniel's engineer.



Herman

A seventy-dollar-an-hour engineer with two buttons. Now the room and the machine, because none of the geometry matters if the fan is in the take.



Corn

Get the computer out. That has to be step one.



Herman

Get it out. One narrator's studio write-up says it flatly, I put my recording computer outside the booth so I don't pick up fan noise. June Yoon, a voice actor, published a whole budget recipe for it. USB keyboard and mouse, a USB hub, a ten-foot USB extension, a ten-foot HDMI, and a small monitor, all for around a hundred and sixty dollars or less. And her rule of thumb, always buy cables with extra length, because the cheap failure is the cable that's exactly long enough.



Corn

The exactly-long-enough cable is the one that gets tugged, and the tug arrives in the waveform.



Herman

And there are two noise sources, not one. Fan noise, which is the low hum, and hard-disc noise, which is the rattle. If you can't move the machine, there are halfway remedies. Tower under the desk, clean the dusty fans, acoustic foam baffles around it during the session, laptop sitting on a mouse pad to damp vibration. And then the test that costs nothing. Record twenty seconds of silence and listen to it.



Corn

Twenty seconds of silence as a diagnostic instrument. Very zen. I might adopt it for naps.



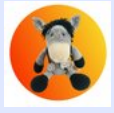
Herman

There's also the cheapest noise control of all, which is geometry. A cardioid mic's rear is its null, its least sensitive side. So you point the back of the mic at the noise source. The mic becomes its own noise gate.



Corn

Which leads us to Daniel's reflection filter, because I have doubts.



Herman

You should, and the research backs the doubt. The filter wraps around the back of the mic. It blocks reflections arriving from behind. But a cardioid's rear is already its null. The reflections that actually reach the capsule come from the front, off the walls and ceiling in front of the performer. The filter is defending a border that was already secure.



Corn

It's a moat on the side of the castle with no road.



Herman

And it carries costs. It adds mass to a one-and-a-half-metre stand, turning all of it into a lever arm that transmits floor vibration. It can introduce tonal coloration of its own in some setups. And the one that matters practically, people who build these stations warn that filters and stands can block the talent's view of the full page, so you crane off-mic to read around them. Daniel's script placement and his filter are fighting each other for the same physical space.



Corn

So treat behind the performer first, and let the polar pattern do the job it was designed for.



Herman

That's the standard advice, yes. Room before gear, always. Most rejected self-recorded audiobooks fail on the room, not the mic. Quietest room, soft furnishings, rug on hard floors, fans and AC and the fridge off. Though I'll flag a genuine disagreement, because it's rare and it's fun. The common advice is that a walk-in closet is a free vocal booth, all that hanging fabric. Home Brew Audio says no, quite firmly. Yes, it blocks outside noise, but the room sound will be boomy and boxy. Small parallel surfaces, all that energy bouncing with nowhere to go.



Corn

A closet is a booth the way a shoebox is a concert hall.



Herman

Which leaves the human being in the chair, or standing, for hour six. And this is where the most counterintuitive fact in the whole episode lives.



Corn

Hydration.



Herman

Hydration, but not the way anyone thinks. The glass of water on the mic stand is basically decorative, physiologically. Water you swallow doesn't touch the vocal folds on the way down. If it did, you'd be choking. It hydrates them from the inside, through the bloodstream, which takes hours. So the water you're sipping at three in the afternoon helps your dry mouth, and that's all. The water you drank all morning is what's keeping the folds supple at three in the afternoon.



Corn

So hydration is a logistics problem. You hydrate yesterday.



Herman

You hydrate all day before, basically. Eight to ten glasses, go easy on caffeine and alcohol, and when you do sip during a session, room temperature, not iced. Cold water shocks the cords. And not chugged, that just runs through you.



Corn

The rest of it. Breaks, warm-ups, the clock.



Herman

Breaks roughly hourly, a stretch every sixty minutes, with some sources saying forty to fifty minute blocks with ten minutes of genuine rest, no talking, and no whispering either, whispering is work. Warm up five to ten minutes before, gentle breath work, lip trills, a paragraph out loud. The Lion King cast does sirens and lip trills before every show, and they're not even recording. And the mouth noise thing, green apple. A green apple before recording helps cut down sticky mouth noise for many actors, which is why you'll find them in studios everywhere.



Corn

The fruit bowl in a recording studio is equipment.



Herman

It's the least mysterious mysterious object in any studio. Posture cues are the usual, long spine, relaxed shoulders, think conversation, with quick resets, breathing four counts in, two hold, six out, jaw release, shoulder drop. And one piece of advice I'd underline for a corpus, don't suppress breaths. Tense suppressed breathing damages your performance more than an audible breath damages your audio. You can always deal with a breath in post.



Corn

And the red flag, because you're the retired doctor and you'll want it on record.



Herman

I will. Any negative voice change lasting longer than two weeks, loss of quality, range, stamina, persistent hoarseness, gets a referral to a specialist voice clinic. That's the British Voice Association's line, and it's the right one. Voices are load-bearing. And the throughput reality check, for expectations. A new narrator typically spends about three hours of studio time per finished hour of audio, in sessions usually six hours long.



Corn

Three to one. So Daniel's few hundred sentences are a genuine workday, not an afternoon.



Herman

A real one, with a lunch break. And one through-line to all of it. The reference corpus people actually built on was recorded with identical setup for every speaker. The consistency isn't the garnish on the corpus, it's the corpus.



Hilbert

That part's right.



Hilbert

Though the pros I dealt with went further than that. You could rebuild their whole station in a hotel room in twenty minutes, from memory. That was the real test. I spent three weeks once getting one sound out of one room. Acoustic panels, a different chair, a second mic I didn't like. In the end I moved the lamp table eighteen inches to the left and that was it. Clean.



Corn

Eighteen. Did you measure that or did your eye do it?



Hilbert

I measured it. It's eighteen. The lamp table was walnut, it belonged to the studio, and eighteen inches was what the room wanted. Herman will tell you a room doesn't want anything, but I heard the difference and I don't need the panel to agree with me. Corn would have gotten it in one, the way he listens. He's slower, but he hears it.



Herman

The engineer in me wants to argue and the years behind this desk say don't.



Hilbert

I kept a photograph of that room. It's taped inside the flight case, the big one. People always assume it's a photo of the mic setup. It's the room. Empty. I've never explained it and I'm not starting for a podcast. That table did its job and the room deserved the record.



Corn

The photograph is of the empty room.



Hilbert

With the lamp table. Eighteen inches left of where the designer put it.



Corn

So the final word on Daniel's kit, as a working pro would give it. The AT2020 is a perfectly fine condenser for this. Twenty dB self-noise, cardioid, needs phantom power, all respectable numbers. If the room turns out to be the weak point, the Q2U is actually the safer choice, it's a dynamic and it just hears less of the room.



Herman

The stand with the filter clamped on is the thing I'd change. It's a lever arm waiting to pick up floor vibration, and the filter is solving a problem the cardioid pattern already solved while sitting exactly where the script wants to be. He can test it in twenty seconds of silence. And the gamepad he should absolutely keep, once he's confirmed the mapping holds over a long session.



Corn

None of that is the real advice. The real advice is pick one posture, one distance, one room, and never change it mid-corpus. It's the cheapest quality lever he has, and it costs nothing but discipline.



Herman

The question Hilbert's photograph leaves me with is the one I'd send back to Daniel. If the corpus is the product, is the recording station a studio or a laboratory? A studio is built for a performance. A laboratory is built so the experiment can be repeated. The pros straddle both, and he gets to choose which one he's running.



Corn

Produce every episode of this show, and our thanks as always to Hilbert Flumingtop at the desk. If you want more of this, try episode fifty-seven, From Lawyers in Limousines to Developers in Their PJs; and episode eight, Building Your Own Whisper. This has been My Weird Prompts. If you enjoyed this one, the review button takes four seconds and tells the algorithm we're a going concern. We'll be back soon. Until then, mind your eighteen inches.

SOURCES

What the research for this episode read before the script was written. Primary sources first.

1. **Lucent Audio, What to Expect in an Audiobook Studio, retrieved 2026-11-08.**
   https://www.lucentaudio.com/blog/what-to-expect-in-the-studio

2. **VO Trainer (Trevor O'Hare), 2026-07-06, retrieved 2026-11-08.**
   https://www.votrainer.com/blog/microphone-technique-tips-every-voice-actor-needs

3. **VocalCopyCat, Microphone Placement for Clear, Consistent Voice Over, retrieved 2026-11-08.**
   https://www.vocalcopycat.com/blog/microphone-placement-for-voice-over

4. **vodit pro, Punch and Roll in Reaper, updated 2026-10-23, retrieved 2026-11-08.**
   https://www.vodit-pro.com/ratgeber/punch-and-roll-reaper

5. **C.C. Hogan, Preparing an audiobook script for recording, retrieved 2026-11-08.**
   https://cchogan.com/audiobook-tips-a-short-guide/preparing-a-script-for-recording/

6. **C.C. Hogan, Recording my Audio Books, retrieved 2026-11-08.**
   https://www.cchogan.com/recording-my-audio-books/

7. **Jim, Voiceover Recording Studio Equipment and Gear, retrieved 2026-11-08.**
   https://justaskjimvo.studio/gear/

8. **Martin Allanson, Voiceover Recording Equipment | How to Setup a Studio, retrieved 2026-11-08.**
   https://martinallanson.com/how-to-setup-a-voiceover-recording-studio/

9. **Rosie Akerman, 2024-06-28, retrieved 2026-11-08.**
   https://www.rosieakermanvoiceover.com/2024/06/28/how-to-set-up-a-voiceover-home-studio/

10. **The Voice Realm (Sarah Robbins), 2026-10-24, retrieved 2026-11-08.**
    https://www.thevoicerealm.com/blog/sit-or-stand-for-voice-over/

11. **The Voice Realm, Vocal Health for Voice Actors, retrieved 2026-11-08.**
    https://www.thevoicerealm.com/blog/vocal-health-for-voice-actors/

12. **The Voice Realm, Positioning Guide For Recording Voice Overs, retrieved 2026-11-08.**
    https://www.thevoicerealm.com/blog/positioning-guide-for-recording-voice-overs/

13. **Gary Terzza, 2014-06, retrieved 2026-11-08.**
    https://vomasterclass.blogspot.com/2014/06/standing-up-or-sitting-down-whats-best.html

14. **Voiceover Gurus, Voice Over Myths, retrieved 2026-11-08.**
https://voiceovergurus.com/guru_blog/voice-over-myths-and-why-they-even-started/

15. **Dani Voiceovers, Vocal Health Tips for Voiceover Actors, retrieved 2026-11-08.**
https://danivoiceovers.com/en/vocal-health-tips-voiceover-actors/

16. **Topher Keene, Vocal Health for Voice Actors, retrieved 2026-11-08.**
https://www.topherkeene.com/articles/vocal-health-for-voice-actors

17. **ACTRA Toronto, Vocal Health for Extreme Voice Performance, retrieved 2026-11-08.**
https://actratoronto.com/vocal-health/

18. **British Voice Association, Lift your voice without hurting it (PDF), retrieved 2026-11-08.**
https://britishvoiceassociation.org.uk/wp-content/uploads/free-voice-care-literature/Lift%20your%20voice%20without%20hurting%20it.pdf

19. **Voiceovers & Soundbites, retrieved 2026-11-08.**
https://voiceoversandsoundbites.com/maintaining-vocal-health-as-a-voice-actor-the-complete-guide/

20. **June Yoon, How To Get Your Laptop (and the fan noise it makes) Out Of The Booth, retrieved 2026-11-08.**
https://www.linkedin.com/pulse/how-get-your-laptop-fan-noise-makes-out-booth-june-yoon

21. **LA Vocal Booths, Computer Fan Noise: Plan Your Booth and Cable Setup, retrieved 2026-11-08.**
https://www.lavocalbooths.com/blogs/booth-basics/computer-fan-noise-booth-cable-plan

22. **Voice Over Xtra (William Williams), Phantom Noise part 2, retrieved 2026-11-08.**
http://www.voiceoverxtra.com/article.htm?id=a326x8rf

23. **no diva. no drama., 2021-01-25, retrieved 2026-11-08.**
https://www.femalevoicenarrator.com/voiceover-blog/2021/1/25/electrical-low-voltage-wiring-for-audio-recording-studio

24. **Home Brew Audio, How to Build a Home Recording Studio: Part 2, retrieved 2026-11-08.**
https://www.homebrewaudio.com/10725/how-to-build-a-home-recording-studio-part-2-4-tips-for-preventing-noise/

25. **Audiobook Editing, Audiobook Home Studio Setup, retrieved 2026-11-08.**
https://audiobookediting.net/audiobook-home-studio-setup/