Episode Book

EPISODE #5847

Why One EQ Curve Can't Fix Two Voices

PUBLISHED	10 October 2026 at 19:56 UTC
DURATION	23:42
LISTEN	myweirdprompts.com/episode/multi-voice-eq-conflict-fix-87052/
HOSTS	Corn, Herman, Hilbert
SCRIPT MODEL	deepseek-flash
VOICES	chatterbox-regular (chatterbox-local)
PIPELINE	V5.3
SOURCES	14 cited by the research (listed at the end)
BOOK GENERATED	19 October 2026 at 17:40 UTC

SYNOPSIS

You've got a podcast where one host sounds warm and the other sounds harsh, and no single EQ preset can fix both. This episode unpacks the "proximity tax" that hits over 80% of interview recordings — why a cut that softens one voice thins out the other — and walks through what actually works: dynamic EQ, multiband compression, de-essing, and shared filtering for panel recordings. Plus the listener-side reality check: Poweramp EQ on Android only reliably processes Spotify, Apple Music, and YouTube Music, so your carefully dialed-in curve may be doing nothing at all. We cover AutoEQ speaker correction, Auphonic's Multitrack Adaptive Leveler, and the two genuinely different paths for fixing a mix you didn't create.

TRANSCRIPT



Corn

Okay, so I'm staring at a number that shouldn't be possible. Two hundred and forty hertz, a Q of one point two, minus three decibels — and the same cut gets applied to over eighty percent of interview episodes. Eight out of ten. That's not a mixing decision at that point. That's a tax.



Herman

It's a proximity tax. Every guest leans into the mic, every guest is sitting in a spare bedroom with a mattress against one wall, and the result is the same boom every time.



Corn

Which is exactly what Daniel's been fighting from the other end of the chain. He's listening on a cheap Anker Bluetooth speaker — one of the ones with no companion app, no EQ of its own — so he's running Poweramp EQ on Android, trying to dial the podcast in by hand, and he can't land it. Corn sounds pleasant. Herman has a slight nasally edge. And the fix for Herman's harshness introduces vocal mod across the whole mix. He's asking whether there's a way out of that, whether the same problem applies to real human recordings or just synthetic voices, and what a producer or a listener can actually do when you can't EQ speakers individually.



Herman

He's got the diagnosis right. That's the thing that gets me. He's describing a structural problem accurately, from the cheap seats, without the vocabulary.



Corn

He calls it the Goldilocks scenario. The parametric adjustment that makes one speaker sound better has the inverse effect on the second.



Herman

And it does. Every time. Because a static curve is applied to the whole file forever, regardless of who's talking.



Corn

So let's start with why this is so hard — and it's not because Daniel hasn't found the right preset.



Herman

The first thing to understand is that every voice has different problem frequencies, and those problem zones overlap. The map is roughly this: mud and proximity sit around a hundred and fifty to three hundred hertz. Boxiness, three hundred to eight hundred. Nasal honk, one to two kilohertz. Presence, two to five. Sibilance, five to eight. Different voices have different problems, and one speaker's mud frequency won't match another's.



Corn

Which is great when you've got one voice on one track. It becomes a negotiation the moment there are two.



Herman

A cut that fixes one host's boom will thin out the second host whose body lives in that same band. That's not a mistake. That's arithmetic.



Corn

So walk me through Daniel's specific case. Corn pleasant and relaxing, Herman nasally and harsh. What's actually happening in the frequency domain?



Herman

Nasality typically lives in the one to two kilohertz honk range — that's the band that gives a voice its pinched, reedy quality. Harshness, the thing you'd want to take down first, sits higher, two to four kilohertz. That's the presence region for a lot of speakers, which is why it's so tempting to reach for.



Corn

And Corn's pleasantness is partly a function of presence in that same band.



Herman

You take down two to four to soften my edge, and you're also pulling back the thing that makes Corn sound warm and forward. The mix goes dull. Daniel called it vocal mod — that's the correct description of what happens. The whole bed of the mix gets a hazy, distant quality that wasn't there before.



Corn

I don't love that you just described me as furniture.



Herman

Room tone. I meant room tone.



Corn

You said bed.



Herman

Bed is furniture.



Corn

Bed is where I make my living.



Herman

Alright, we'll come back to that. The point is that a static parametric EQ applies the same cut or boost forever, no matter who's speaking. When two voices need opposite treatment at the same frequency, no single static curve can satisfy both. That's not a skill gap. That's the tool.



Corn

So the answer isn't a better preset.



Herman

There aren't better presets. The preset is the problem. And it's not a marketing thing — the manufacturers of these tools say it themselves. Static EQs apply the same cut or boost forever. Dynamic EQ reacts only when it needs to.



Corn

There's the whole rest of the episode in two sentences.



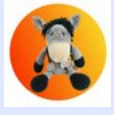
Herman

The producer's real answer isn't EQ. It's frequency-selective, level-triggered processing. Dynamic EQ, multiband compression. Waves' F6 is the clean example — you set a threshold at a specific frequency, and the cut only engages when the energy at that frequency crosses the line. Between the peaks, the band is untouched.



Corn

So for Herman's harsh two to four, you'd set a dynamic band that only moves when he's actually pushing.



Herman

And releases when he isn't. That's the whole trick. The tool is doing the thing Daniel's trying to do by hand — it's just waiting for the trigger instead of cutting blindly.



Corn

De-essing is the same idea applied to a specific problem band. Five to eight kilohertz.



Herman

And it's specifically *not* a static EQ job. Trying to fix sibilance with a static cut will make the voice dull between the harsh consonants. You hear it immediately — the esses get tamed, and everything else goes lispy and soft. A de-esser applies dynamic compression only in the band when energy peaks. So the esses get caught and the vowels get left alone.



Corn

The static EQ doesn't know when the S is coming. The dynamic one does.



Herman

The static EQ doesn't know anything. That's the point.



Corn

Okay, so we've established that dynamic processing is the right tool for the can't-separate case. But most podcast mixes aren't mastered with dynamic EQ per voice. What's the actual first move for a producer?



Herman

Separate first, then treat. That's the workflow. The recommendation across the board is: separate speakers, balance levels, de-noise, then EQ, then compress, then set loudness. Treat the two voices as two separate repair jobs before you treat them as a mix.



Corn

If you do it out of order?



Herman

You get a pumping, over-squashed mess. The compressor sees both voices as one signal, gets confused by level differences, and the whole thing breathes. It sounds like the mix is gasping.



Corn

That's a good description of about half of the interview podcasts on the internet.



Herman

And the other half sound like the two hosts are in different rooms.



Corn

Which is the panel case. Three or more speakers. What do you do when you've got a roundtable and every voice is different?



Herman

The recommendation is: apply the same high-pass and low-pass to every track first — eighty hertz on the bottom, fifteen kilohertz on the top — then treat each voice individually after that. The shared filtering is what makes the panel sound like it was recorded in the same room, even if it wasn't. The priority is consistency across tracks, not perfection on any individual track.



Corn

Right. Because if you chase perfect tone on voice three, the other five sound like they're phoning it in from a different building.



Herman

And listeners forgive flat tone. They don't forgive inconsistency. If one voice sounds closer and one sounds distant, all of it reads as amateur regardless of how well it's been EQed.



Corn

So what about the mastering-stage tools? The ones that work across all tracks.



Herman

Auphonic's Multitrack Adaptive Leveler is the one I'd point at. It uses signals from all tracks to correct level differences between speakers. It's not an EQ — it's a leveling problem — but it's aimed at the same underlying issue, which is that voices don't sit at the same level when they're coming from different people in the same room.



Corn

So it's a different lever pulling on the same multi-voice problem. You can't fix frequency per voice, so you fix the dynamics across voices instead.



Herman

Right. And it's one of the few tools where you actually get *more* signal from having all the voices available, not less. Most of the time the second voice is a complication. There it's the input.



Corn

That's interesting. Okay, so we've established the producer's side. Separate first, then treat. Dynamic EQ where you can. Shared filtering when you can't.



Herman

Yeah.



Corn

So what if you're not the producer? What if you're sitting on a sofa with a thirty-dollar speaker and a phone?



Herman

Then you've got Poweramp EQ, and you've got a much tighter box.



Corn

Let's talk about the box.



Herman

Poweramp Equalizer is actually a serious piece of software for something that costs five dollars. The premium tier gives you parametric mode with user-added bands — you pick the type, low pass, high pass, low shelf, high shelf, band pass, peaking, and you set channel, gain, frequency, Q. You can save per-device presets. It imports AutoEQ text files in both graphic and parametric form, and when you import a graphic one, it adds six decibels to the gains to compensate for the way AutoEQ exports.



Corn

The parametric mode was added a while back. Build eight ninety-nine through nine oh eight.



Herman

Yeah. The free tier gives you nineteen built-in presets, thousands of graphic AutoEQ presets, and device-specific EQs. Premium gets you twenty-five additional graphic and parametric presets, more AutoEQ presets, device-specific parametric EQs, plus a bass and treble dial and a compressor.



Corn

That's a lot of tool for a podcast listener.



Herman

It is. And here's the catch Daniel may not know about.



Corn

Of course there's a catch.



Herman

Poweramp EQ only reliably works with Spotify, Apple Music and YouTube Music. Tidal, Deezer, Qobuz, Netflix, YouTube, and Chrome are not recognized without experimental DUMP or ADB permissions.



Corn

So if Daniel plays the podcast through a browser, or through an app that isn't one of those three, the EQ may be doing nothing.



Herman

Nothing at all. The audio goes around it entirely.



Corn

"I'll just tune it." It's like adjusting the colour settings on a television that's off. You walk away convinced you've improved the picture.



Herman

That's a real failure mode for the whole project. You spend an evening dialing in a curve, and you've been listening to unprocessed audio the entire time.



Corn

How would you even know?



Herman

You'd toggle the EQ on and off while audio is playing. If nothing changes, the app isn't in the signal path.



Corn

Which is a five-second test, and most people never run it.



Herman

Most people assume the setting is applying. Which is a reasonable assumption. It just isn't always correct.



Corn

So let's say it *is* applying. Daniel's still stuck. What's the actual recommendation?



Herman

Two paths, and they solve different problems. The first is source-side EQ — get the producer to fix it before it ships. That's the cleanest fix, but it requires separate voice tracks, and it requires the skill to treat each voice individually. None of which the listener controls.



Corn

And it doesn't help any of the audio that's already been mixed and published. Which is most of it.



Herman

Most of it. The second path is speaker-specific correction — the AutoEQ approach.



Corn

Which is the thing Daniel asks about directly. "Speaker-specific presets."



Herman

AutoEQ is the canonical project here. It's a massive measurement database — thousands of headphones, IEMs, and speakers, including the Anker Soundcore line. Liberty 4 Pro, Liberty 5, Liberty 4 NC, Sleep A20, Space Q45. All measured. Each one has a correction curve that pushes it toward a target — Harman, IEF Neutral, diffuse-field, whichever you pick.



Corn

So the listener imports the correction for their exact speaker, and the speaker stops colouring the audio.



Herman

It stops colouring the audio. What it doesn't do is fix the multi-voice conflict inside the content.



Corn

Because if two voices need opposite treatment at two to four kilohertz, a flat speaker can't solve that. It just delivers both voices uncoloured, and they still disagree with each other.



Herman

Correct. AutoEQ corrects the *hardware* toward a neutral target. It helps every piece of content equally. It does not solve a mix problem.



Corn

So the two paths are: fix the content before it ships, or fix the speaker so it stops lying about what's actually there.



Herman

And neither one gives the consumer a way to EQ individual voices in a podcast they didn't produce.



Corn

Which is the actual shape of the problem. Daniel's asking whether there's a preset he's missing. The answer is no. There's no preset that does per-voice correction on a two-voice mix, because per-voice correction requires per-voice isolation, which is exactly what the podcast doesn't hand him.



Herman

And there's a negative finding worth saying plainly. There's no standalone product shipping "speaker-specific podcast EQ presets" — per-podcast, per-speaker. AutoEQ covers hardware, not content. The tool Daniel is imagining doesn't exist yet.



Corn

That's worth knowing. That's a real answer. "You're not missing a preset, the preset doesn't exist."



Herman

Yeah.



Corn

Okay, second half of his question. TTS versus humans. Does the same mastering problem apply, and does the answer change?



Herman

The mastering question is identical. The root causes differ.



Corn

Unpack that.



Herman

The research literature treats TTS quality through MOS — mean opinion score — naturalness testing. That's how TTS systems get evaluated. The SOMOS dataset is the biggest example — twenty thousand synthetic utterances of the LJ Speech voice, generated by two hundred different TTS systems. That's a naturalness benchmark. It's not about EQ. Nobody's publishing "how to EQ a TTS voice."



Corn

So the literature doesn't answer Daniel's question directly.



Herman

It doesn't. But it points at the underlying mechanics. TTS artifacts are model and vocoder-dependent. They come from training data limits or information loss during distillation. Not from a microphone. Not from a room.



Corn

Which means they're baked in.



Herman

They're baked in at synthesis time. The vocoder is going to do what the vocoder is going to do, and the listener can't reach back into the model to fix it.



Corn

Compare that to humans. Real human recorded in a bad room — the boom at two hundred and forty hertz is a proximity effect plus a room mode. You could, in principle, go back and re-record. Put the guest further from the mic. Hang a blanket.



Herman

Or you treat it dynamically.



Corn

Either way, the root cause exists in the physical world and can be addressed there. The TTS artifact is not in the world. It's in the model.



Herman

So in some ways TTS voices are *easier* to EQ. No room modes. No proximity effect. Consistent every episode. The same voice, at the same level, at the same tone, on Tuesday as it was on Monday.



Corn

Which should be great for a static preset.



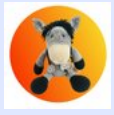
Herman

It should. In theory, once you've got the curve right for a TTS voice, it stays right forever. Nothing drifts.



Corn

Meanwhile, in reality.



Herman

In reality, the artifacts that are baked in can't be fixed at the source by anyone downstream. And different TTS systems bake in different artifacts. So the preset that works for one model's output won't work for another's.



Corn

Which is the exact opposite of what a preset is supposed to be. A preset is a saved curve that keeps working. If every model needs a different one, you're not saving settings. You're just starting over every time.



Herman

Yeah. And that's the interesting asymmetry. Humans are harder to EQ because everything varies. TTS is easier to EQ because nothing varies — but only within a single model, and only after someone has figured out what that model's specific artifacts are.



Corn

And that someone isn't the listener.



Herman

No. The listener gets whatever the API spits out.



Corn

So if you're in Daniel's position — cheap speaker, Poweramp, both hosts' voices fighting each other inside a single mix — what are you actually supposed to do?



Herman

The honest answer is that you've got two moves available and neither is a complete fix. First, make sure the EQ is actually applying at all. If you're listening through a browser, you might be doing nothing.



Corn

Step one. Turn the dial and check the lights are on.



Herman

Step two, run the speaker correction so the cheap hardware stops adding its own colour on top of the problem. That won't fix the mix, but it removes a second layer of distortion. Then, from there, you're doing damage control — a gentle dynamic cut in the harsh band if Poweramp's compressor can help, and after that you're out of moves.



Corn

Whereas the producer side has a much bigger toolbox, and it still can't make one curve work for two voices.



Herman

It can't. The tools are better. The structural constraint is the same.



Corn

That's the part I keep coming back to. Both ends of the chain are being asked to solve a multi-variable problem with a single knob.



Herman

Which is why the answer to Daniel's question is "sort of, but not really." There is a solution for the producer. There is a partial mitigation for the consumer. There is no solution for the consumer that reaches backward into the mix and cleans up the voice that's bothering them.



Corn

Before we wrap up. Hilbert has something he wants to say about this — and I'm not sure it's going to help.



Corn

It's a hundred and twelve.



Hilbert

The high-pass. You said a hundred. It's a hundred and twelve. I've had that written on a card in a drawer since I started doing this. Twelve is where the plosives stop. You go to eighty, you start losing body on the lower voices. You go to a hundred and forty, you start sounding thin. Twelve. Every time.



Corn

Hilbert, that's — okay. That's specific.



Hilbert

It's the kind of thing you only learn by doing it wrong nine times. Anyway. You're both wrong about the EQ.



Herman

In what sense?



Hilbert

You've been treating this like the knobs are the problem. I once spent a whole weekend trying to make a two-host thing sound right in a car. It was a ninety-seven Escort. It sat in a service bay and I sat in it because the waiting room had terrible lighting and I had a two-hour repair on the books.



Corn

Two-hour repair.



Hilbert

It was a two-hour repair once I finished what I did under the hood. Let's say the estimate was ambitious. The stereo in that thing had three knobs. Bass, mid, treble. That was the entire EQ. Three knobs on a plastic face, and one of them had a scratchy shaft.



Herman

And you got the podcast sounding right?



Hilbert

I got it sounding *better*. Here's what you do. Midrange all the way down. Treble all the way up. Then adjust the bass until the two voices stop fighting. That's it. You don't need parametric anything.



Corn

That's a smiley-face curve. That's literally the loudness button.



Hilbert

That's not the loudness button. The loudness button is a marketing feature. This is a technique.



Herman

The two voices were fighting, you said?



Hilbert

Trying to occupy the same parking space. That's what it is. Both voices want to park in the same spot. The midrange is the spot. You take the midrange out of the equation, they've got no spot to fight over, and you're left with the words and the bass. Words and bass. That's all a podcast is.



Corn

What was the podcast?



Hilbert

It was the hold music for a chain of car dealerships. Two guys. One did the local news bits, one did the financing offers. But if you EQ'd it right, they sounded like they were in the same truck.



Herman

The hold music was the podcast you were trying to EQ?



Hilbert

It was already playing on the dealership's PA system. Which is how I ended up testing it. I got the service manager to let me plug the CD player into the PA. That took some doing.



Corn

How much doing?



Hilbert

About forty-five minutes and a lunch I wasn't going to eat anyway. The CD player had a skip-protection buffer that had failed. Every time someone walked past the doorway of the service bay, it would stutter. So I had to sit perfectly still, which is why I brought the wrench down with me — I was using it to brace the tray.



Herman

Wait, so the CD player was sitting on a wrench?



Hilbert

On top of a socket rail. Same idea. The point is that I got the EQ right. I sat there with the two knobs — mid all down, treble all up, bass adjusted — and it worked. Both voices locked in.



Corn

And the service manager?



Hilbert

He was my cousin. He got fired about a month later. Unauthorized audio testing. That was the actual phrase on the write-up, in the actual box on the actual form. Unauthorized audio testing.



Herman

Hilbert.



Hilbert

It's a real phrase. I saw the form.



Corn

And the hold music?



Hilbert

It got a lot better. For about three weeks.



Herman

And then?



Hilbert

And then the dealership switched to a new system. No physical knobs. All touchscreen. You couldn't adjust anything without going through four menus, and by the time you found the bass slider they'd already put you on hold for a different reason. Modern cars. That's the problem.



Corn

Okay. So where does that leave us?



Herman

It leaves us where Daniel's already standing, honestly. Two static ends of the chain, one curve that has to serve both voices, and neither side can fully solve it with the tools that either side can reach.



Corn

The producer's answer is per-voice processing — dynamic EQ, multiband compression, the level-triggered stuff that only moves when it needs to. The consumer's answer is a speaker correction and a fight with whichever app is allowed to touch the audio stream. Neither one of those things talks to the other.



Herman

And the gap widens the more synthetic voices ship. TTS is consistent within a single model, which means a preset that works today may work forever. But the different models all bake in different artifacts, so the preset that fits one doesn't fit another, and the consumer has no way to know which is which without listening.



Corn

The question worth leaving people with is this: is the Goldilocks scenario actually solvable for a listener, or is the real answer always to go upstream and get the source audio right — either at the microphone or at the model?



Herman

Does the rise of TTS make that better or worse? More consistency, but artifacts you can't get behind.



Corn

I'll take either answer from a listener. Thanks as always to Hilbert Flumingtop for producing — even the parts we didn't ask for.



Herman

Especially those.



Corn

For more along these lines, there's episode fifteen, AI Gets Personal; episode nine, Benchmarking Custom ASR Tools - Beyond The WER; and episode four, If Your Voice Ages, Does Your Fine-Tune Become Useless. This has been My Weird Prompts, the human-AI collaboration podcast. We're on the website at my weird prompts dot com — the feed's there, the archive's there, all of it. If you've been wrestling with your own Goldilocks EQ problem, send us your own prompt on Telegram at t dot me slash MWP listener bot.



Herman

We'll be back soon.

SOURCES

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

1. **Poweramp forum, 2021-06-29 [primary]**
https://forum.powerampapp.com/topic/21506-poweramp-equalizer-builds-899-908/

2. **AutoEQ repo [primary]**
https://github.com/jaakkopasanen/AutoEq

3. **Auphonic Multitrack docs [primary]**
https://us.auphonic.com/help/algorithms/multitrack.html

4. **Podigy, Podcast EQ Settings: The Complete Guide, updated 2026-03-17**
https://podigy.co/blog/podcasters-eq

5. **Android Authority, 2023-07-02**
https://www.androidauthority.com/third-party-vs-companion-eq-apps-3330999/

6. **GitHub, created 2026-03-22**
https://github.com/DAPAAADF/AutoEQ-Parametric-64band-for-Poweramp

7. **Waves F6**
https://www.waves.com/plugins/f6-floating-band-dynamic-eq

8. **Initial Audio Dynamic EQ 2**
https://initialaudio.com/dynamic-eq-2/

9. **Sonnox Oxford Dynamic EQ**
https://www.uaudio.com/products/sonnox-oxford-dynamic-eq

10. **SplitBySpeakers**
https://splitbyspeakers.com/playbooks/clean-up-podcast-audio

11. **AudioForgePro**
https://audioforgepro.com/blog/how-to-make-two-speakers-sound-the-same-volume-podcast-interview

12. **SOMOS dataset, INTERSPEECH 2022**
https://arxiv.org/abs/2204.03040

13. **MOS-prediction TTS training, IEEE Access 2022**
https://arxiv.org/abs/2011.01174

14. **MixChops frequency map**
https://mixchops.com/learn/mix-problems