WebRTC Audio Troubleshooting - Field Runbook

Five ordered checks, the number to read at each, and the fix - on one page.

Step 0 - classify before you touch anything

One-way or two-way?	Isolates the failing direction - a send path or a receive path.
Silence / distortion / echo?	Silence -> Check 1. Distortion -> Check 2 or 4. Echo -> Check 3.

The five checks - run top to bottom, stop at the first that fires

1 Device & permissions
Read: sender audioLevel ~ 0; permission 'denied'
Fix: Unmute OS; re-grant; pick the right input device.

2 Network path
Read: packetsLost rate; jitter; concealedSamples rate
Fix: Enable FEC/RED; let the jitter buffer adapt.

3 Echo cancellation
Read: echoCancellation = false? on speakerphone?
Fix: Enable AEC; the reliable fix is headphones.

4 Sample-rate mismatch
Read: rate at capture vs codec (48 kHz) vs AudioContext
Fix: Insert a resampler at the boundary; never force a rate.

5 Packet capture
Read: webrtc-internals dump; Wireshark RTP stream
Fix: Find dropped/mismatched media on the wire (TURN, SDP).

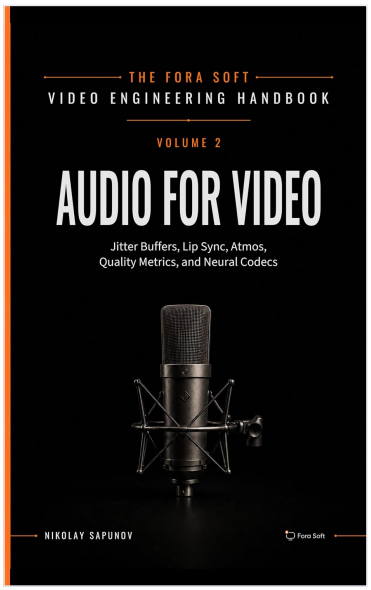
The dashboard: which getStats field is which

audioLevel	Sender media-source. 0 = silence, 1.0 = 0 dBov. Flat 0 while talking = dead mic.
packetsReceived	Inbound. Not climbing = no media arriving at all (firewall/NAT).
packetsLost / jitter	Inbound. Loss rate + arrival-timing spread (RFC 3550 A.8, seconds).
jitterBufferDelay	Inbound. / jitterBufferEmittedCount = avg time audio waited.
concealedSamples	Inbound. Synthesized audio covering gaps; climbing fast = distortion.

Remember

- [ ] Classify first: one-way vs two-way, and silence / distortion / echo.
- [ ] Run the five checks in order - ~7 in 10 tickets end at Check 1 or 2.
- [ ] Sample every getStats value TWICE a second apart; use the rate of change, not the total.
- [ ] Flat sender audioLevel while speaking = muted/wrong device - the network is irrelevant.
- [ ] Loss < 1% inaudible; 2-5% noticeable clicks; > 5% severe without redundancy.
- [ ] Echo with AEC on? Ask 'speakerphone or Bluetooth?' before touching code.
- [ ] Wrong pitch/speed with CLEAN network stats = sample-rate mismatch, not loss.
- [ ] Reach for Wireshark LAST, not first - it is the most expensive check.

Companion to: 'Diagnosing Audio Problems in Production: A Runbook'
www.forasoft.com/learn/audio-for-video/articles-audio/diagnosing-audio-problems-runbook



GO DEEPER · VOLUME 5 OF 7

Audio for Video: A Complete Guide to Sound in Video Products

From Jitter Buffers and Lip Sync to Atmos, Quality Metrics, and Neural Codecs

By Nikolay Sapunov, CEO at Fora Soft · 1,288 pages · Kindle & paperback on Amazon

Read it on Amazon → amazon.com/dp/B0HH6XPY46

This runbook classifies a broken call, runs five ordered checks and maps each getStats field to a fix. Volume 5 goes deeper into the causes: jitter buffers and NetEq, loss concealment, FEC and RED, timestamps and clock domains, and incident analysis by symptom.

WHAT YOU WILL FIND INSIDE

- Incident analysis by symptom
- Jitter buffers, NetEq and concealment
- FEC and RED against packet loss
- Timestamps, clocks and sample rates

THE FULL HANDBOOK · 7 VOLUMES · 9,000+ PAGES · ALL ON AMAZON

Vol. 1 · Video Encoding — pixels and color to H.264, HEVC, AV1, quality metrics and AI
Vol. 2 · Video Streaming I — network transport, contribution and ingest, HLS, MPEG-DASH, CMAF
Vol. 3 · Video Streaming II — ABR, CDNs, players, WebRTC, DRM and the economics of delivery
Vol. 4 · Audio for Video I — PCM, codecs, loudness, accessibility, real-time voice processing
Vol. 5 · Audio for Video II — jitter buffers, lip sync, Atmos, quality metrics, neural codecs
Vol. 6 · AI for Video I — models and inference, computer vision, speech, search, generative video
Vol. 7 · AI for Video II — real-time pipelines, agents, serving, cost, compliance, verticals

Users reporting silence, echo or robotic audio in your calls?

We will run the five checks on your production stats, find the failing stage and fix it in your client or media server. 250+ video products shipped since 2005.

Book a 30-min call → **WhatsApp →** **Email us →**