Episode Book

EPISODE #6423

How Spoken Safety Rules Became Machine Training Data

PUBLISHED	3 November 2026 at 15:43 UTC
DURATION	33:15
LISTEN	myweirdprompts.com/episode/voice-recognition-air-traffic-control-98784/
HOSTS	Corn, Herman, Larry, 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	4 November 2026 at 09:31 UTC

SYNOPSIS

High-stakes professions relied on spoken discipline long before artificial intelligence arrived. Examining decades of development across air traffic control reveals how rigid phraseology designed for human reliability ended up creating ideal training grammars for speech recognition engines. From 1989 MIT studies and Eurocontrol trials to recent automated transcription and European aerospace tests, the technical challenge was never just recognizing sound, but translating highly structured, mission-critical speech into trustworthy operational data.

TRANSCRIPT



Corn

Eighty-one point two percent. That's the number that got my attention this morning, and I hadn't even had my first nap.



Herman

Context for that, because it deserves it: 1994, Eurocontrol runs a simulator study, eight controllers, forty-four sessions, about twenty-six hundred voice inputs into a speech recognition system, and the machine got it right first time for eighty-one percent of them.



Corn

Which in 1994 was a triumph, and today would get a product recalled.



Herman

And that's exactly the tension Daniel sent us. His framing is that voice tech is going mainstream now, but there are four workforces that have lived by the spoken word for decades — air traffic control, military command and control, emergency dispatch, medical dictation. The original voice-first professions, as he puts it. He wants to know for each one: when did speech recognition actually show up, what did the work run on before it, and how has the job changed since. And the gear question alongside — who supplies each world, and how much of the kit is bespoke purpose-built stuff versus ordinary commercial technology.



Corn

So the question isn't really when voice tech arrived. It's what these professions built to make voice reliable enough to bet lives on before the machines could help — and what happened when the machines finally caught up.



Herman

The order matters here. Air traffic control first, because the stakes are highest and the discipline is most rigid. Then military, then dispatch, then medical dictation, which is the field where this whole story already ended.



Corn

Already ended?



Herman

Ended, twenty-some years ago. Medical dictation is the one where speech recognition arrived and won. The other three are still mid-fight.



Corn

Fine. But before we get there — the thing I keep noticing in all of this is that none of these professions waited around for recognition to get good. They engineered the speech itself. Air traffic control has ICAO phraseology and the phonetic alphabet. The military built voice command around the idea that the operator's eyes are busy. Dispatch had radio logs and card systems. Medicine had the transcription pool — a room full of typists turning a doctor's rambling into a chart.



Herman

Right, and that's the thread that pays off later, so let me say it once properly: all of that — the phraseology, the phonetic alphabet, the tape, the typists — that was reliability engineering for human ears. And here's the twist. That same rigid discipline is why the machines can do it now. The phraseology wasn't just the interface for humans. It's the training data. The safety script became the grammar.



Corn

A grammar that limits what a controller can say to a few thousand sanctioned constructions.



Herman

A few thousand constructions with hard slots — call sign, altitude, heading, clearance type. When you shrink the space of legal sentences, you shrink the space the recognizer has to search. Low perplexity, in the jargon.



Corn

So the very thing that makes controllers sound like robots is what makes them legible to robots.



Herman

Couldn't have put it better as a mechanism. Let's do air traffic control properly then. The pre-recognition workflow, the one that ran for half a century — voice radio, and on top of it, the standard. In the US it's codified in FAA Order JO 7110.65, which is the controller's bible. Current version is 7110.65BB, issued February of last year, and they republish it roughly every six months because the phraseology keeps being tuned.



Corn

A rulebook for talking. Republished twice a year. I want you to sit with that, listeners, the way I had to.



Herman

It specifies — down to the word — how you say an altitude, how you hand off an aircraft, what "cleared" means versus "climb and maintain." Every ambiguity that ever killed somebody, somebody wrote a rule about. The phonetic alphabet — Alfa, Bravo, Charlie — exists because "B" and "D" don't survive a bad radio, but "Bravo" and "Delta" do.



Corn

So recognition research reached ATC early. How early?



Herman

Nineteen eighty-nine. MIT's Flight Transportation Laboratory, a paper by Joakim Karlsson, making the case that speech recognition could improve safety and efficiency in the tower — and then the honest sentence: the available devices, he wrote, were "not sufficiently capable for operational ATC use." MIT had actually evaluated two commercial recognizers, the Votan VPC 2000 and the Verbex Series 5000.



Corn

That's a very polite way of saying the machine couldn't tell "cleared to land" from static.



Herman

Then five years later, Eurocontrol runs the study behind your number. The Verbex system, the same family, in a TRACON simulator — TRACON being the terminal radar approach control environment, the busy bit around the airport. Eight controllers, forty-four sessions, twenty-six hundred sixty-two voice inputs. Eighty-one point two percent recognized first time.



Corn

And the other nineteen percent?



Herman

Repeated, corrected, or abandoned. Which sounds tolerable until you remember that in a real tower, the other nineteen percent is a pilot flying into something. That's why it stayed in the lab.



Corn

Because ATC has a parallel escape hatch that the other three professions don't have — the radio isn't the only channel. There's a whole datalink story running alongside this.



Herman

Yes, and it's the road not taken that people forget. ACARS, the aircraft communications addressing and reporting system, goes operational in 1978 — text messages between aircraft and ground. FANS-1/A in 1995 brings controller-pilot datalink communications over oceanic routes. VDL Mode 2, the VHF digital link, goes operational in 2000. ARINC provides the ATN over VDL service for the FAA's first CPDLC build at Miami Center in 2002.



Corn

So for forty years the industry's answer to unreliable voice was: use less voice.



Herman

And NASA tried to bridge the two, which I find beautiful. The Controller Interface work in the late nineties — a controller speaks a clearance out loud, the recognition system parses it, converts it into a DO-219 datalink message, and it goes up over Mode S data. Speech in, text out, radio never touched.



Corn

A machine whose whole job is eavesdropping on the controller and doing the typing.



Herman

That is a surprisingly accurate summary of where the field ended up twenty years later. Modern FAA and MITRE work: remote audio access to a hundred thirty FAA facilities, more than two million hours of audio, two billion transmissions transcribed per year. That was the 2021 figure. The point of that corpus isn't transcription — it's training and safety analysis.



Corn

Two billion transmissions a year. Someone is indexing every "roger" in American airspace.



Herman

And the applications got serious. MITRE's Closed Runway Operation Prevention Device — built to catch the exact scenario where a controller clears an aircraft onto a runway that's closed. Field-tested at Dulles in the summer of 2014, ninety-two point seven percent true intent accuracy on clearance detection.



Corn

Ninety-two point seven. Nineteen ninety-four's machine got eighty-one. Thirty years for eleven points.



Herman

It's worse than that, actually — it's better, but it's worse. The Verbex number was a speaker-dependent system: each controller trained it on their own voice, in a simulator, in a quiet room. The Dulles number is real controllers, real frequencies, real accents. That's the jump that matters.



Corn

The European side has the numbers I actually believe, because they're cruel.



Herman

The DLR work, the German aerospace center, under SESAR. The PROSA validation ran twelve Vienna approach controllers through twenty-eight hours of simulation — about a hundred twenty thousand spoken words, nine thousand transmissions, eighteen thousand commands. Word error rate around three percent. Command recognition ninety-two percent. Call sign recognition ninety-seven.



Corn

The cruel one.



Herman

The cruel one is the comparison the researchers ran alongside. Without the recognition support, eleven percent of the controllers' logged commands were missing or wrong. With the recognition system backing them up, four percent. The machine reading the human beat the human reading the human.



Corn

Hold on. That's not automation taking a task. That's automation catching the controller's own mistakes.



Herman

On that narrow, well-structured task — logging what was said — yes. And the efficiency numbers were absurd. The mouse-clicking time, the manual entry into the electronic flight strips, fell from twelve thousand eight hundred seconds across the simulation to four hundred.



Corn

Twelve thousand eight hundred to four hundred. Someone divided by thirty-two and framed that.



Herman

The earlier AcListant-Strips work put it in fuel: seventy-seven seconds of flight time and sixty litres of fuel saved per aircraft, because the clearance gets logged and acted on faster. Multiply across a busy approach sector and it stops being a rounding error.



Corn

Before we leave ATC — the FAA's own prototype, the one from 2024.



Herman

CASPER. October 2024. A computer-simulated pilot that has to reconstruct the controller's command. Eighty-six percent of the time it produced the correct pilot command overall — and ninety-six percent when the controller used standard phraseology.



Corn

There's the whole episode in one figure. Deviate from the script and you lose ten points.



Herman

Which is either an argument for better models or an argument for controllers to stop improvising, and honestly it's both.



Corn

Here's where I have to say the thing about the gear, because we've been talking as if the tower runs on magic. The console a controller sits at, the radio, the voice switch — that's purpose-built, certified kit governed by that 7110.65 world. And the recognition running on top at the research end is increasingly ordinary commercial machinery — Whisper, fine-tunes of Whisper, the whole open ecosystem. Which raises the question of whether the moat is the hardware or the data.



Herman

The data, and increasingly only the data — but hold that, because it deserves its own stretch.



Corn

Agreed. And speaking of gear that hears what you meant rather than what you said —

L

Larry

I want to tell you about a customer of mine, a satisfied customer, name of Gary, no wait, it was Gerald, Gerald from Bakersfield, and Gerald was a controller, an air traffic controller, and he told me the old phraseology was wearing him down, all that roger and wilco and say again, so he switched to WhisperWing Pro, the headset that hears what you meant and not what you said.

L

Larry

Gerald says WhisperWing Pro turned his whole tower voice-first, it replaces the phonetic alphabet with a friendly hum, and now every clearance he mumbles goes out clean and recognized first time, eighty one point two percent of the time every time, and he's never once had to key a microphone again.



Corn

Eighty-one point two percent of the time, every time.



Herman

I have never once contacted anyone about running an ad on this show. Not Larry, not anyone named Larry, I don't know a Larry, I've never owned a headset with a hum, and I would want to see the certification paperwork before any tower went voice-first on a product whose selling point is skipping the microphone.



Corn

The denial had four clauses, Herman.



Herman

It needed four.



Corn

Military next. And I'll confess this is the one I knew least about going in.



Herman

The military story starts earlier than the civil one and delivers later, which is a very military pattern. The framing you find in the literature by 1990 — there's an ACL paper on military systems — is battle management command centers needing rapid access to large, rapidly changing databases in what they call an "eyes-busy environment." The commander's eyes are on displays. Hands are busy. Voice is the free channel.



Corn

The free channel. That's the entire pitch, and it's a good one.



Herman

And they had a limited-vocabulary recognizer working at a weapons-control workstation, and the paper's assessment of the users was that they were — the phrase is "optimistic despite limited capability."



Corn

Which should be engraved above every defense procurement office.



Herman

There's also a Naval Postgraduate School thesis from March 1985 that I like because it's so stubbornly practical. It compared continuous speech, discrete speech, and keyboard input for an interactive warfare simulation, across different light and noise conditions. Continuous speech and keyboard beat discrete speech in all conditions.



Corn

Discrete speech being the pause. Between. Every. Word.



Herman

The technology of the era, because recognizing a pause was easier than recognizing a sentence. The finding was that if you could do continuous at all, it was better even in noise. And speed versus error rate was a real tradeoff against manual input — faster, but wrong more often. That tradeoff is still the whole argument today.



Corn

How long did the research program run?



Herman

CERDEC, the Army's communications-electronics research center at Fort Monmouth, had been at tactical speech recognition for over twenty years as of a 2004 paper — so from the early eighties. Including whispered-speech recognition for covert operations, which I find remarkable — recognizing a whisper is brutally hard and they were funding it in that era.



Corn

DARPA's end of it.



Herman

The 1988 spoken language systems program set a target of real-time systems handling up to ten thousand words at better than ninety-five percent word accuracy. Which, for context, is roughly what consumer dictation hit about a decade later — on a machine with a million times the memory, indoors, with one cooperative user.



Corn

And the thing that actually shipped.



Herman

TRANSTAC, from 2004 — spoken language communication and translation for tactical use. The Army's machine foreign language translation suite descended from it, and it runs on Android smartphones. Soldier holds up a phone, speaks English, a local speaks Arabic or Dari or whatever the language set is, and it translates.



Corn

A program that began as command-and-control research ends as a soldier holding a phone to a stranger in a village. That's a bigger pivot than most companies manage.



Herman

The procurement history around it is its own saga — a statement of need from Army Pacific in 2001, the Army designated lead service for language technology that spring, joint requirements by 2003, a language transformation roadmap by 2005. Twenty years of paperwork to get a phone in a soldier's hand.



Corn

Dispatch, then. The newest convert.



Herman

And the fastest mover, which surprised me. The pre-recognition workflow is the one everyone of a certain age remembers from police shows: radio logs, handwritten cards on a spinning rack, then computer-aided dispatch — CAD — which became the backbone of every 911 center in the country.



Corn

Who makes CAD?



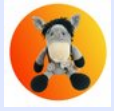
Herman

Four names dominate the American market: Motorola Solutions with PremierOne, Tyler Technologies with New World, Hexagon with I/CAD, and Central Square. That's the proprietary, purpose-built end of this whole episode, in one list.



Corn

Motorola making both the radio and the software that logs the radio. Comfortable arrangement.



Herman

And sticky. A March 2023 NHTSA report looked at how well emergency communications centers share CAD data with each other, and the answer was: barely. Limited interoperability between centers that routinely hand off the same incident.



Corn

Two neighboring counties, same car chase, systems that can't talk. The voice-first profession, defeated by databases.



Herman

Which is partly why AI arrived so fast in dispatch — the pain is enormous and the staffing is desperate. CISA, in March of last year, was already describing AI in emergency communications centers as a decision-support tool — transcription, keyword detection, background noise reduction, translation. Not as a replacement. As triage.



Corn

What's it actually doing?



Herman

The NTIA put out a white paper last month, based on site visits and interviews, and the numbers are startling for a field this conservative. Jeffcom 911 in Colorado — AI cut their non-emergency call volume by up to forty percent, stabilizing around a sustained third. Monterey County in California, roughly thirty-six percent on average, peaking near thirty-nine. Monterey now resolves over a hundred non-emergency calls a day without a human touching them.



Corn

A hundred calls a day resolved by software. What are they, people reporting a parking complaint at two in the morning?



Herman

Mostly exactly that — the non-urgent tier, noise complaints, abandoned vehicles, that layer of demand that eats dispatcher hours without needing a dispatcher. And the Orleans Parish district in New Orleans — cloud-native systems, two and a half hours of downtime since mid-2022, AI triage averaging forty-plus events a day, and other callers getting answered up to forty seconds faster.



Corn

The retention number was the one that got me.



Herman

Monterey's academy retention — dispatch is a brutal field for attrition, you spend months training someone and they quit. Their average across the six academies before the AI tools was forty-three percent. The three most recent academies, with AI fully in the workflow: sixty-seven percent.



Corn

Because the job stops being an unbroken stream of misery with a headset on.



Herman

That's the humane reading and I'll take it. The vendors are piling in — Intrado launched AI voice-to-text translation for emergency centers in forty-plus languages last summer, Prizym claims a hundred eighty-five, and Wahltek's ReVI is purpose-built for public-safety answering points, tuned for overlapping speakers and background noise.



Corn

Tuned for the reality of a 911 call, which is three people shouting over a passing ambulance. That's the one acoustic environment no consumer product is tested against.



Herman

And that gets us to medical dictation, and I get to be the retired doctor here, so let me have this.



Corn

By all means. You lived it.



Herman

I dictated for years before any machine listened. The workflow was: you finish a consult, you dictate into a recorder, the tape goes — later, a digital file — to the transcription pool, and a human being, usually a woman, usually paid per line, turns your mumbling into a formal note by the next morning. The transcriptionist was the reliability layer. The whole system ran on human ears that had heard every doctor in the department and every drug name in the formulary.



Corn

And the machine that ended that.



Herman

Dragon NaturallySpeaking, 1997. The first continuous speech dictation system — no pausing between words — and Dragon Medical followed as the first continuous dictation product built for healthcare specifically. By January 2008, fifty to sixty thousand physicians were using Dragon Medical, including seven thousand in the Veterans Administration. By August 2008 it crossed seventy thousand licenses — more than a tenth of all US physicians.



Corn

Doctors, historically famous for embracing new software, adopting at that speed.



Herman

The push was the 2009 HITECH Act — electronic records became effectively mandatory, and every note now had to be typed into a system. The choice was type it yourself, dictate for a transcriptionist, or dictate to Dragon. Nuance was doing around four hundred fifty million dollars a year in Dragon medical sales by 2010.



Corn

And the transcriptionist?



Herman

The Bureau of Labor Statistics counts about forty-two thousand medical transcriptionists today, median pay around forty thousand dollars a year, and projects the field to shrink about four percent over the next decade. And the BLS says why, in plain language: speech recognition and natural language processing let physicians document in real time.



Corn

The one profession on Daniel's list where the machine didn't just assist. It took the job.



Herman

Took the job and then handed half of it back. There was a 2025 study — not one from my world, worth saying — finding AI draft notes contained errors in seventy percent of cases. The doctor still has to read the note before signing it. The work didn't vanish; it moved from transcription to verification.



Corn

Same doctor, same note, except now the error is hidden in fluent prose instead of a missing tape. I'd rather catch a blank page than a confidently wrong one.



Herman

And the modern continuation is ambient — Nuance's DAX Copilot listens to the whole consultation and drafts the note afterward. Built, as they say themselves, on the Dragon Medical franchise, with the ambient work starting in 2017.



Corn

So the arc across all four: ATC spent thirty years getting from eighty-one percent to ninety-two and still isn't operationally deployed for controllers. The military spent two decades in the lab and shipped a phone. Dispatch went from no AI to a third of its call volume in about two years. Medicine had its revolution in 1997 and is now on the second one.



Herman

Pull the lens back, then, to Daniel's gear question, because the pattern in those four arcs isn't random. ATC and dispatch run on purpose-built proprietary kit. Medicine and the research frontier of ATC run on commercial foundation models — Whisper, distil-Whisper, the fine-tuning stack.



Corn

Why the split?



Herman

Safety criticality, mostly. A CAD system and a controller console are certified, procured on multi-year cycles, and locked into standards. A note that a doctor signs is a documentation workflow — a wrong word is caught by the human who reads it before it matters. The loop closes with a person, so the loop can absorb an eighty-five percent-reliable component. ATC voice is the loop.



Corn

The NTIA named the barriers, didn't they.



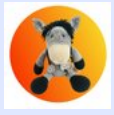
Herman

Verbatim: "rigid procurement processes, a lack of national standards, and vendor lock-in." That's about emergency dispatch, but it describes ATC equally well.



Corn

Which produces the strange situation we're in — the field with the best data is the slowest to use it.



Herman

The moat question, then. If the gear is certified consoles and Motorola CAD, the vendor's advantage is the certification and the install base. If the value is a fine-tuned model, the advantage is the corpus — the hours of domain speech nobody else has. That's where it's



Herman (continued)

all moving.



Corn

Daniel would say the mainstream is catching up with the niche, and he's been watching it from the other end — he's building his own fine-tune and buying dictation microphones, and his observation is that for good gear, the Philips medical dictation mics, you still go hunting outside consumer channels. The mainstream catches up and the good stuff stays niche.



Herman

Because the good stuff was never mainstream gear. It was built for a professional with a workflow, not a consumer with a podcast.



Corn

His Bluetooth experiences being hit and miss, which is the polite version.



Herman

Bluetooth is a consumer convenience protocol doing a professional job. Latency and dropouts matter less for a phone call than for dictation. That's the whole gear split in miniature.



Corn

The accent problem, though — that's the downstream effect I think this whole episode turns on.



Herman

The Singapore work from this year. ASTRA is an air traffic controller training simulator with autonomous simulated pilots, and they measured off-the-shelf recognition systems on Singaporean-accented aviation speech: word error rates up to a hundred seven point eight percent.



Corn

A rate above a hundred. The machine scores worse than a coin-flip of simply guessing nothing.



Herman

Worse than useless — it confidently emits wrong text. The same project's fine-tuned pipeline reaches twenty-three point four five. And a paper from early 2025 got under ten percent on Southeast Asian accented ATC speech with fine-tuning.



Corn

So an off-the-shelf model trained on Western voices is actively dangerous in a Singapore tower, and a fine-tune of the same underlying model is fine. The moat isn't the model. It's the accented hours.



Herman

The ATCO2 project — five thousand hours of transcribed ATC speech from over ten airports, built precisely because that data didn't exist publicly. Seventeen point nine percent word error rate on public ATC datasets as of the 2023 write-up, improving since.



Corn

And the newest answer to scarcity is manufacturing it.



Herman

This year's synthetic audio work — text-to-speech, voice conversion, and the clever bit, accent conversion, in both directions — generating accented ATC speech to train on, because you can synthesize a thousand Singaporean-accented clearances for a fraction of what collecting them costs.



Corn

Training the machine on speech no human ever spoke, so it can understand the speech humans actually speak.



Herman

If you'd proposed that at the 1994 Eurocontrol study they'd have quietly ended the session. Which, I should note, we have — we've covered the general ASR capability jump before, back when we did the episode on how dictation went from frustration to magic, and I said then that the transformer did in five years what the previous thirty couldn't. The part I underweighted: the domains that had been failing since 1989 were positioned to benefit most, because they'd already done the hard part themselves.



Corn

They'd written the grammar.



Herman

The grammar, the accents, the data — the fields did the prep work without knowing it. There's one more gap I want on record, because it's the thing I couldn't nail down and I don't like not nailing things down. Nobody has a public accounting of how much 911 speech recognition is proprietary versus commercial foundation models under the hood. The vendors all say "purpose-built tuning." None of them say what they tuned.



Corn

They say purpose-built the way a restaurant says homemade.



Herman

And for a public-safety system, that opacity is a finding in itself. The procurement story is years ahead of the transparency story. When Monterey's triage bot makes a wrong call, who knows what model made it?



Hilbert

They're not wrong about that.



Hilbert

Herman, you said the phonetic alphabet is for clarity on a bad radio. It's a compression scheme. Alfa, Bravo, Charlie takes longer than A, B, C, but it carries more redundancy, and redundancy is what survives noise. The pre-machine world was designed for the worst channel, not the best one, and you both undersold that.



Corn

Undersold how?



Hilbert

Everything ran on the assumption the channel would fail. I have a customer, Dennis, comes into the shop — I still do two mornings a week at the radio place on Agrippas — and he wanted a dictation program trained to take police-style traffic over the counter, phonetic alphabet and all, for his towing business. Weeks I spent on it. Weeks. The software kept turning Echo into "eggo" and X-ray into "excretory." I have the list still, every misrecognition, handwritten, eleven pages. Better reading than most poetry published since.



Herman

Eleven pages of the machine mishearing NATO.



Hilbert

The list is the only part of the job I kept. The software, the microphone, the whole setup went back. Dennis ended up paying a woman in Ramat Gan to type his calls, which she did faster and without the attitude. The turtle here would say that's the whole industry in one anecdote, but I wasn't going to be the one to say it.



Corn

I was going to say it differently.



Hilbert

You were going to take four minutes. The other thing — I read the 7110.65 aloud to myself for about a month, evenings, because the phraseology settles me down. This was before your show existed and I'm not apologizing for it. I can still recite the clearance delivery script. A headset sits wrong on a snout anyway, so it was read from paper, as intended.



Corn

Eleven pages and a memorized FAA order, and the microphone went back.



Hilbert

The microphone worked fine. The list was the point.



Corn

Of course it was.



Herman

The thing his list actually demonstrates — and I say this as the man who spent the segment defending the grammar — is that the grammar only pays off when the machine has been taught it. Echo to "eggo" is what happens when you feed a consumer model a professional dialect.



Corn

Which is the fine-tuning argument wearing a towing-company uniform.



Herman

Where that leaves us, then — the proprietary vendors' hardware advantage is eroding, because the models are commodities and the mics, as Daniel keeps proving, you can buy. What they actually hold is the certified install base and the domain hours. The certification regime may keep them in place a decade longer than the technology alone would justify.



Corn

And the open question is whether the accent gap gets closed by synthetic data or whether it takes procurement pressure — whether a Singapore regulator has to refuse a Western model before anyone fine-tunes for the towers that need it.



Herman

The transcriptionist-to-verifier shift is the other one I'd watch. Forty-two thousand people moving from typing the words to catching the machine's mistakes, and a seventy percent error rate in draft notes means the verification isn't ceremonial. Whether that's a stable resting place or a waypoint on the way to something with no human in it, nobody's data reaches.



Corn

Produce this one properly, Hilbert Flumingtop. For more along these lines, there's episode ten, How ASR Went From Frustration To … Whisper Magic. This has been My Weird Prompts. If the episode earned its runtime, leave us a review wherever you get your podcasts — it's how the archive finds new ears. We'll see you soon.

SOURCES

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

1. **FAA/MITRE, 2021-09-16 [primary]**
https://www.faa.gov/air_traffic/flight_info/aeronav/atiec/media/Presentations/2021/ATIEC_2021_ATC_Speech_Rec-MITRE-2021-09-16.pdf

2. **FAA CASPER, Oct 2024 [primary]**
https://hf.tc.faa.gov/publications/2024-computer-automated-simulation-pilot/casper_final_202410.pdf

3. **Karlsson, MIT, 1989 [primary]**
https://ntrs.nasa.gov/api/citations/19910009712/downloads/19910009712.pdf

4. **Eurocontrol TRACON/Pro study, 1994 [primary]**
https://www.eurocontrol.int/archive_download/all/node/10179

5. **SESAR JU, 2023-03-31 [primary]**
https://sesar.eu/news/reduced-flight-time-and-fuel-burn-cards-controller-voice-recognition-tool

6. **NASA CPDLC Controller Interface [primary]**
https://ntrs.nasa.gov/citations/19980210625

7. **CISA, March 2025 [primary]**
https://www.cisa.gov/sites/default/files/2025-03/25_0328_s-n_ai-implemen-ecc_infographic_508C.pdf

8. **NTIA, Sept 2025 [primary]**
https://www.ntia.gov/sites/default/files/2025-11/ntia-ai-in-9-1-1-white-paper.pdf

9. **Intrado, 2025-07-22 [primary]**
https://www.intrado.com/news-releases/intrados-ai-voice-translation-and-transcription-technology-speeds-9-1-1-response

10. **Nuance Dragon timeline [primary]**
https://www.content.shi.com/cms-content/accelerator/media/pdfs/nuance/nuance-051823-evolution-of-dragon-medical-one-infographic.pdf

11. **BLS Occupational Outlook Handbook, 2025 data [primary]**
https://www.bls.gov/ooh/healthcare/medical-transcriptionists.htm

12. **Wikipedia, last edited 2026-10-13**
https://en.wikipedia.org/wiki/Speech_recognition

13. **Wikipedia, last edited 2026-01-03**
https://en.wikipedia.org/wiki/Aeronautical_phraseology

14. **Wikipedia, last edited 2026-06-06 (7110.65BB, 2025-02-20)**
https://en.wikipedia.org/wiki/FAA_Order_7110.65

15. **Wikipedia, last edited 2026-10-28**
https://en.wikipedia.org/wiki/Medical_transcription

16. **ATCO2 lessons learned, Aerospace 2023**
https://doi.org/10.3390/aerospace10100898

17. **FedScoop**
https://fedscoop.com/mitre-air-traffic-conversation-ai-tool-faa-dot/

18. **Ohio University/NASA, Sept 2001**
https://ntrs.nasa.gov/api/citations/20010090826/downloads/20010090826.pdf

19. **ARINC CPDLC/ADS-C milestones**
https://www.icao.int/sites/default/files/sp-files/SAM/Documents/2010/SURAUTOSEM/17%20Arinc_CPDLC%20ADSC.pdf

20. **CERDEC tactical speech recognition, 2004-06-01**
http://oai.dtic.mil/oai/oai?identifier=ADA465047

21. **Sears, DARPA, JASA 1988**
https://pubs.aip.org/asa/jasa/article-pdf/84/S1/S188/11455906/s188_2_online.pdf

22. **Opportunities for Advanced Speech Processing in Military Systems, 1990**
https://aclanthology.org/H90-1103.pdf

23. **NPS thesis, 1985-03-01**
http://hdl.handle.net/10945/21199

24. **MFLTS on Android**
https://aclanthology.org/W18-1923.pdf

25. **NIST/TRANSTAC**
https://www.nist.gov/publications/evaluating-speech-translation-systems-applying-score-transtac-technologies