

Speech Quality Benchmarking for Communications Between NASA Advanced Reference Spacesuits (ICES-2026-574)

Steven Shawn Smith (GeoControl Systems/Amentum/NASA Audio Lab)

Additional NASA suit audio system info: ICES-2024-64

Trade names are used in this presentation for identification only. Their usage does not constitute an official endorsement, either expressed or implied, by the National Aeronautics and Space Administration, Amentum, or GeoControl Systems.

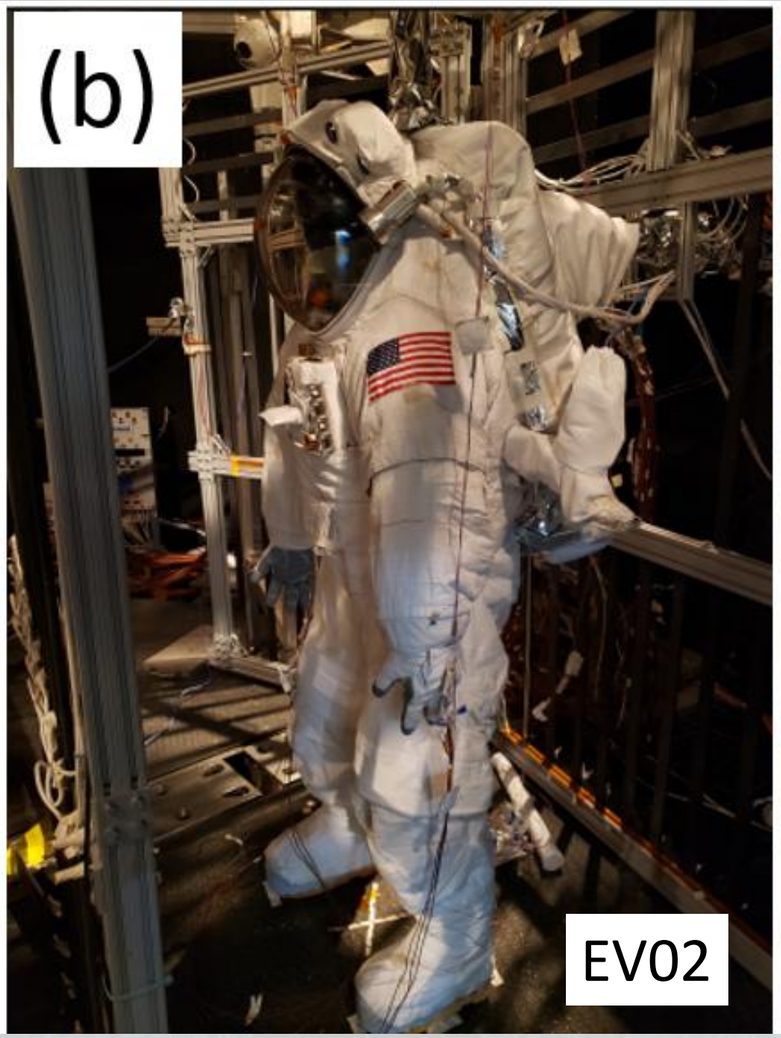
Executive Summary

- Compare audio in & between two NASA advanced suits
- Excessive vs intended background noise conditions
- End-to-end and segmented communications path (helmet, mics, electronics)
- Three benchmarks of
  - intelligibility vs signal quality
  - low vs high information density
- Demonstrate
  - transmitting suit's helmet/noise greatest issue
  - Electronics/microphone quality good – doesn't degrade audio

Equipment: SxEMU



“Short” xEMU with Human Metabolic Simulator & PLSS



Second suit w/nitrogen and extensive thermocouples



Audio test manikin with voice and measurement microphones

COMMCAP FREE

Key Audio/Acoustics Concepts

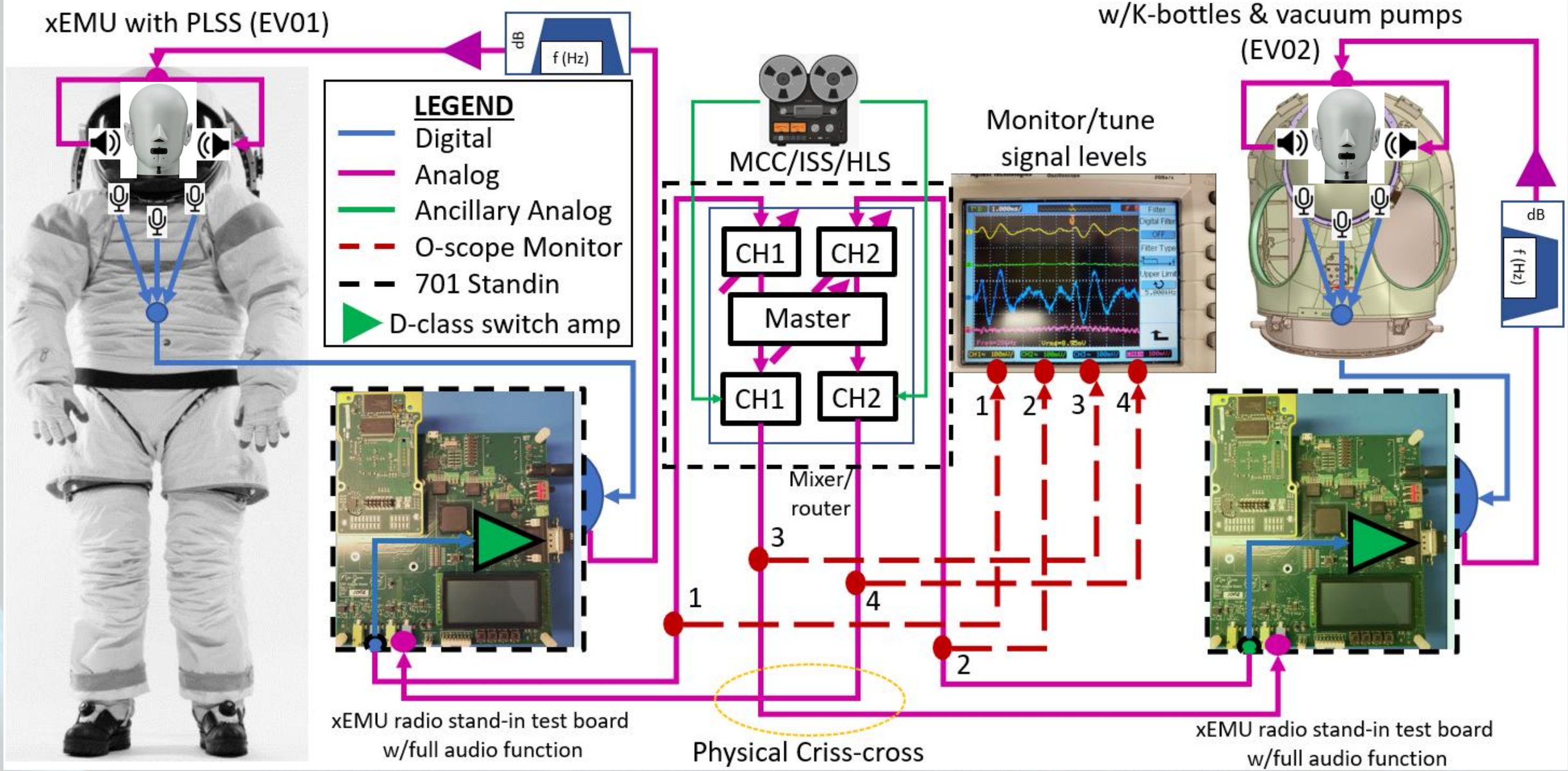
- Suit is a filter (speakers, suit volume, noise sources)

- Three interacting acoustic regions: HUT, Helmet, PLSS

- Acoustic Echo Canceller (AEC) – NO COMMCAP
  - multiband adaptive Digital Signal Processing (DSP)
  - models how suit changes inbound comms, tones
  - subtracts from outbound comms
  - prevents echo and feedback
  - passes crew speech (good) but also PLSS noise (bad)

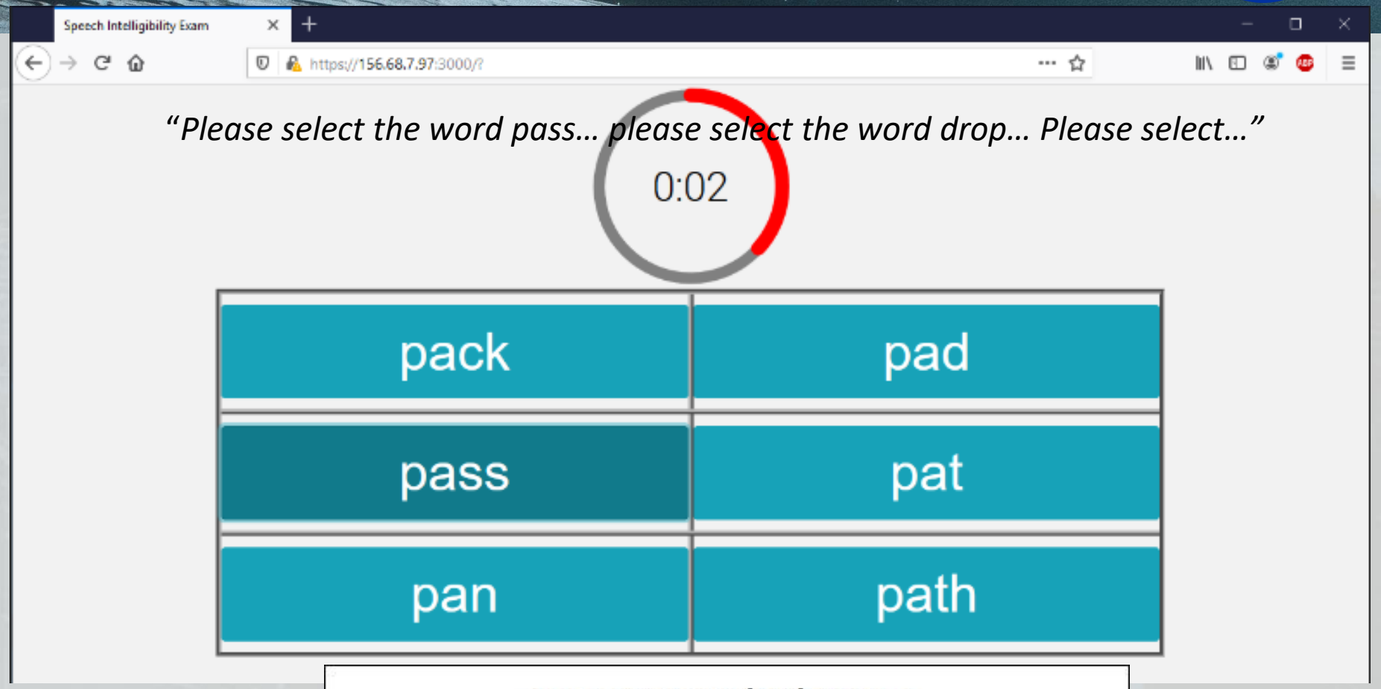
2-suit comms via simple, continuous loop

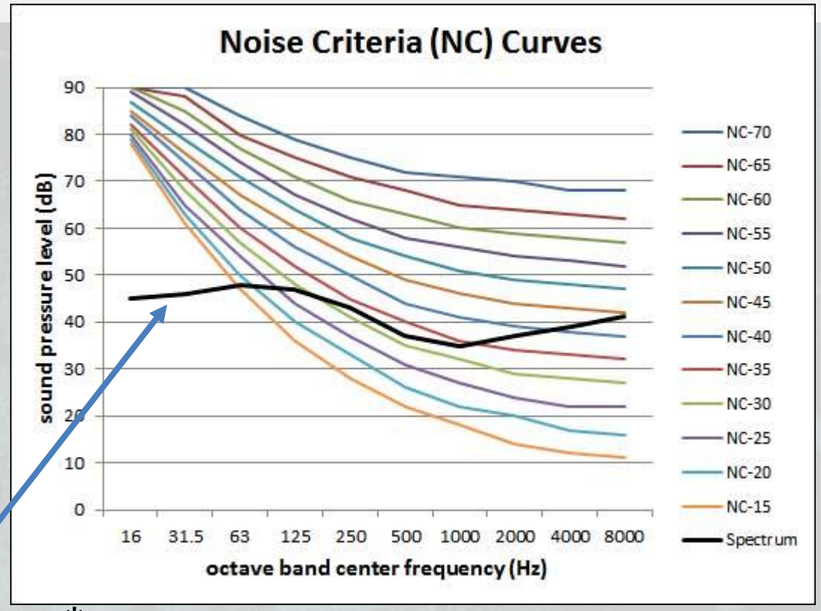
COMMCAP FREE



SPEECH/AUDIO BENCHMARK 1: MRT

- Modified Rhyme Test (MRT)
  - ANSI/ASA S3.2-2009 standard
  - subjective listening jury
  - scores intelligibility at single point
  - Single score: % of words recognized
  - LOW info density
- Suit scores (single numbers)
  - xEMU + PLSS: 63% worst case (85 dB)
  - xEMU + External Fan: 88% best case (60 dB roughly “NC-52” sound levels*)





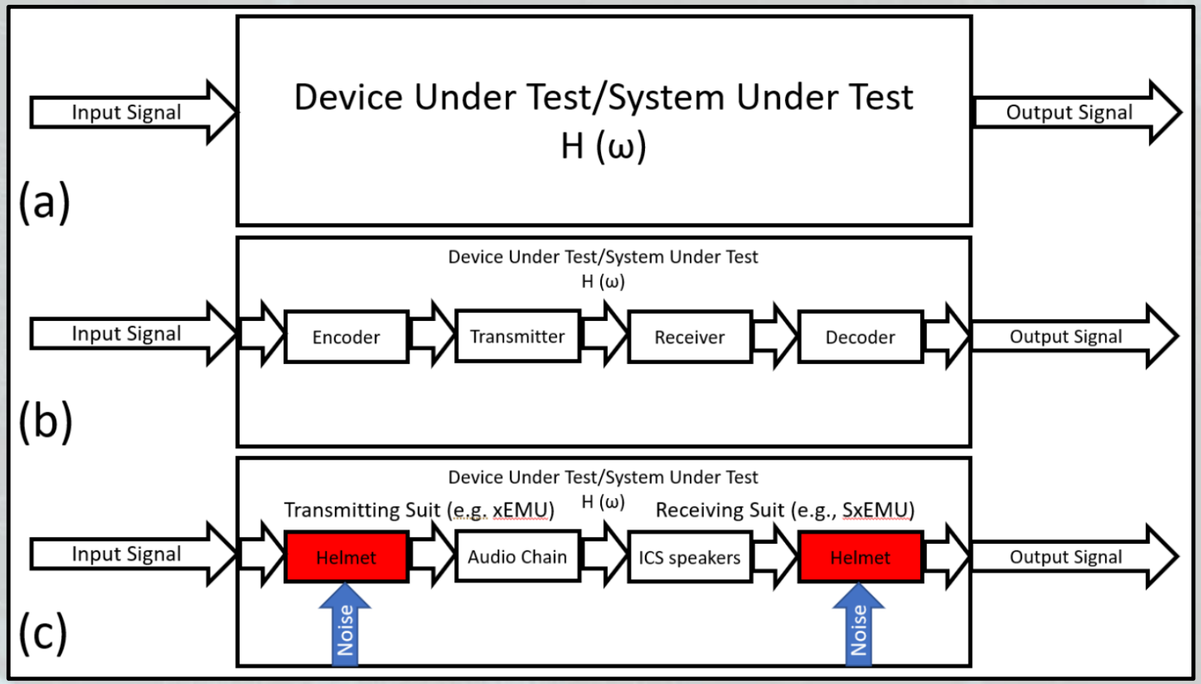
Graphic: Wikipedia

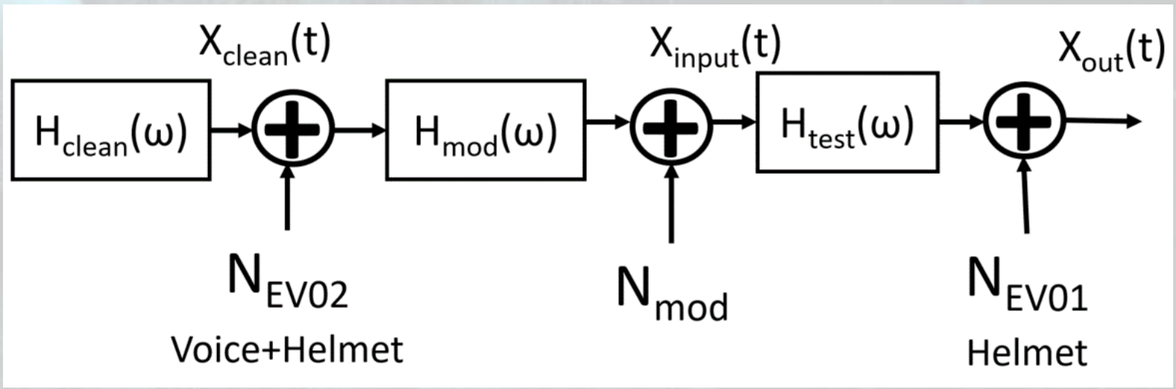
Black curve = example measurement
Highest value at 8000 Hz hits NC-45 curve
Thus, system is NC-52 rated

*NC-XX Curves are guideline dB vs frequency curves for the design of various acoustic spaces (classrooms, theaters, homes, etc.)

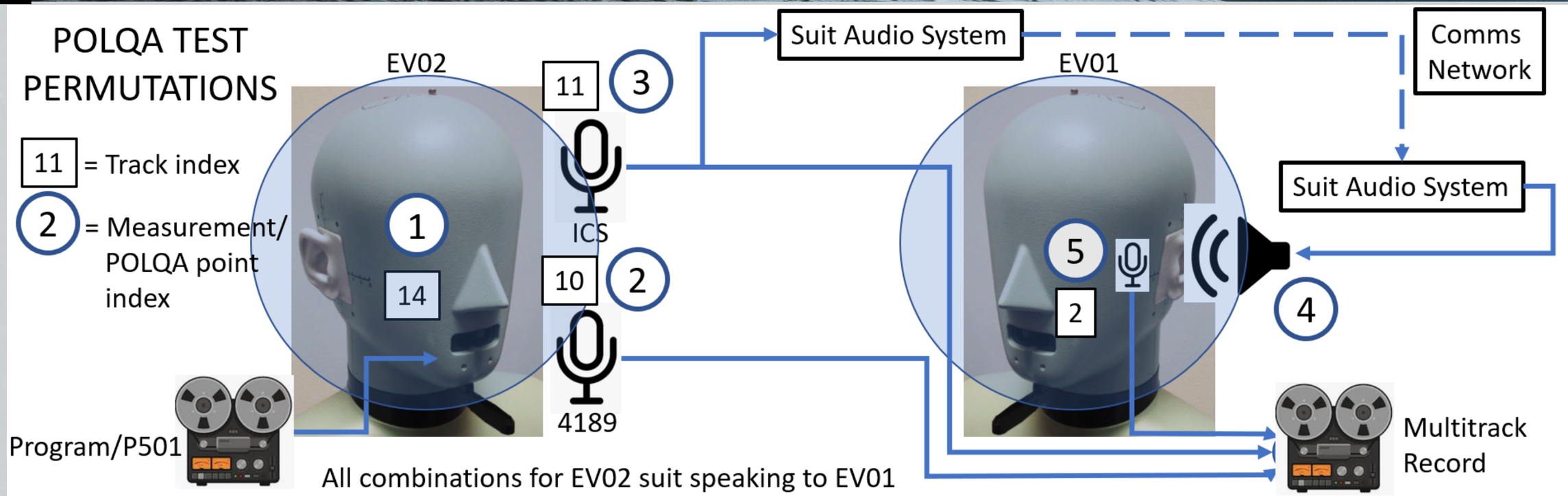
SPEECH/AUDIO BENCHMARK 2: POLQA

- Preliminary Objective Listening Quality Analysis (POLQA)
  - International Telecommunication Union
  - "Mean Opinion Score" (MOS): "pseudo-objective": numerical model/algorithm from extensive human testing (proprietary/licensed)
  - Single score: 1 to 4.5 (Fs=8 kHz)
  - Clean input vs time-aligned output
  - Closed component or chain FIGURE (a,b)
  - Applied here in segments – acoustic not suggested: FIGURE (c)
  - Segment input taken to be "clean" reference





POLQA: Applied across segments



All combinations for EV02 suit speaking to EV01

Index	Ref Point	Meas Point	Ref Track	Meas Track	MOS xPLSS	MOS "NC-52"	Result
1	EV02 Voice in (1)	EV02 B&K mic (2)	14	10	1.95 (56.7%)	2.83 (37.0%)	Manikin mouth + suit degradation of clear speech at B&K mic
2	EV02 Voice in (1)	EV01 L-EAR Out(5)	14	2	1.41 (68.7%)	2.17 (51.8%)	Manikin mouth + suits + audio system degradation
3	EV02 ICS mic (3)	EV01 L-EAR Out(5)	11	2	2.53 (43.8%)	2.54 (43.5%)	System + receiving suit/speaker degradation
4	EV02 B&K mic (2)	EV01 L-EAR Out(5)	10	2	2.41 (46.4%)	2.44 (45.8%)	ICS speaker degradation (no audio system) – this is a signal-to-signal comparison
5	EV02 Voice in (1)	EV02 ICS mic (3)	14	11	1.97 (56.2%)	2.64 (41.3%)	Manikin mouth + suit degradation of clear speech at summed ICS mics

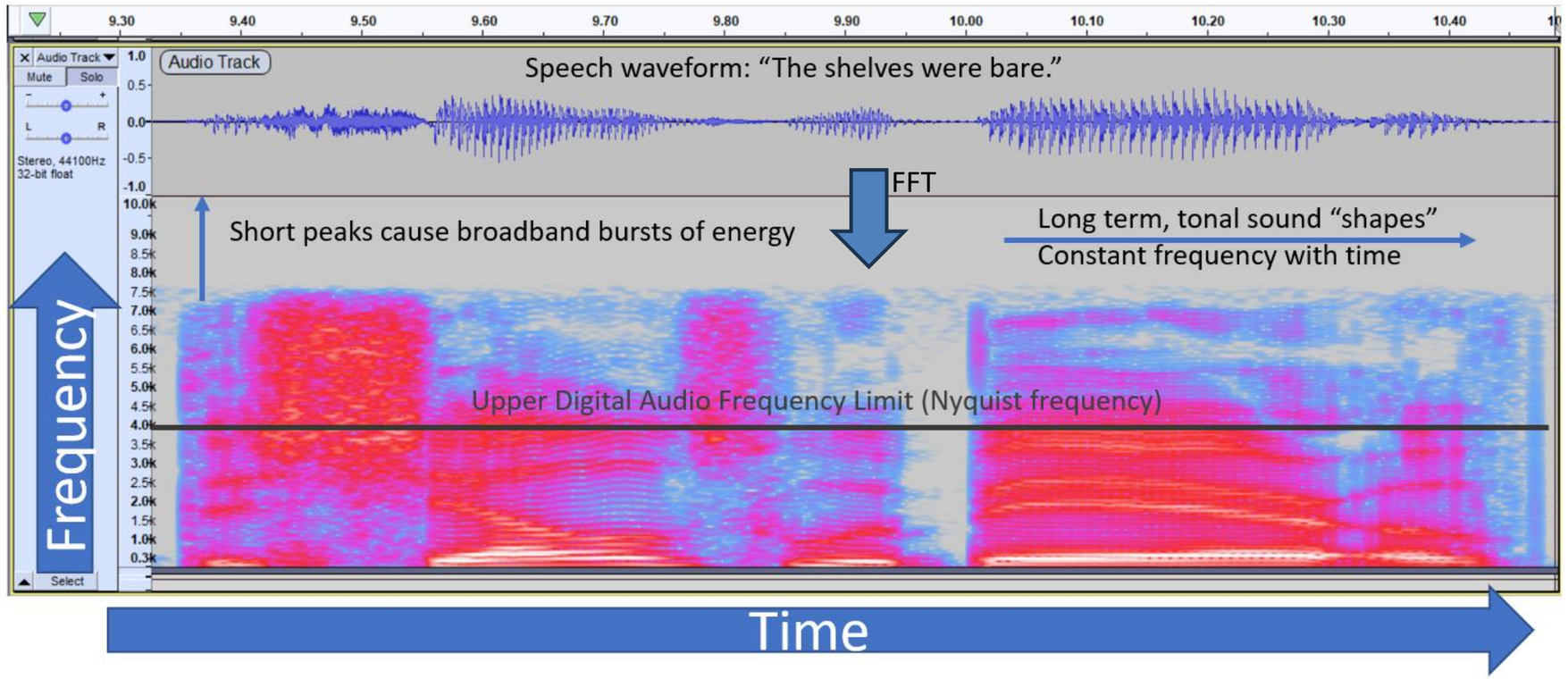
Worst Case | Best Case

POLQA “suit tour”: point-to-point degradation

- Goal: Isolate and Identify segments causing most degradation
- Helmet + PLSS give bulk of end-to-end degradation (Row 2)
- B&K meas. mics and suit mics comparable (Row 3)
- electronics degradation small (Rows 4&5)

Type	PLSS-case Degradation	NC52-case Degradation	Transmit Point	Receive Point	Comment
End-to-End	69%	52%	1, EV02 mouth	5, EV01 ear	Full path mouth to ear, max degradation
Mouth to Mics: Helmet & PLSS noise	56%	41%	1, EV02 Mouth	2,3 EV02 collar mics	Collar & Reference mics same 1 to 4 kHz, assume manikin mouth eq. flat. BULK OF TOTAL DEGREDATION
Mic compare: Meas. Mic vs Suit Mics	7%	7%	1, EV02 Mouth	Difference of 2,3 EV02 mics	Small difference for measurement vs suit mics small: suit mic measurement confidence good
Audio Electronics	(69-56)=13%	(52-41)=11%	3, EV02 ICS mics	5, EV01 ear	End-to-end minus manikin/PLSS → suit-mics-to-EV01 manikin ears. Audio degradation small.
Split electronics: one branch w/electronics, one w/o	2.6% between mics	2.3% between mics	3, EV02 meas. mic vs 2, EV02 suit mics	5, EV01 ear	ICS mic branch has audio electronics, B&K 4189 branch does not. SPECIAL CASE. Infer low electronics degradation.

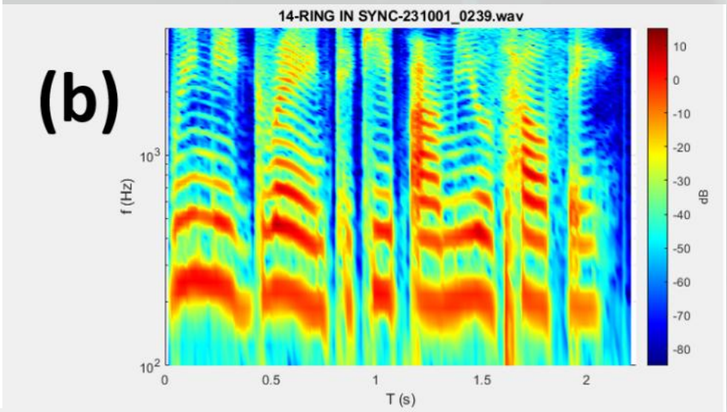
Spectrograms



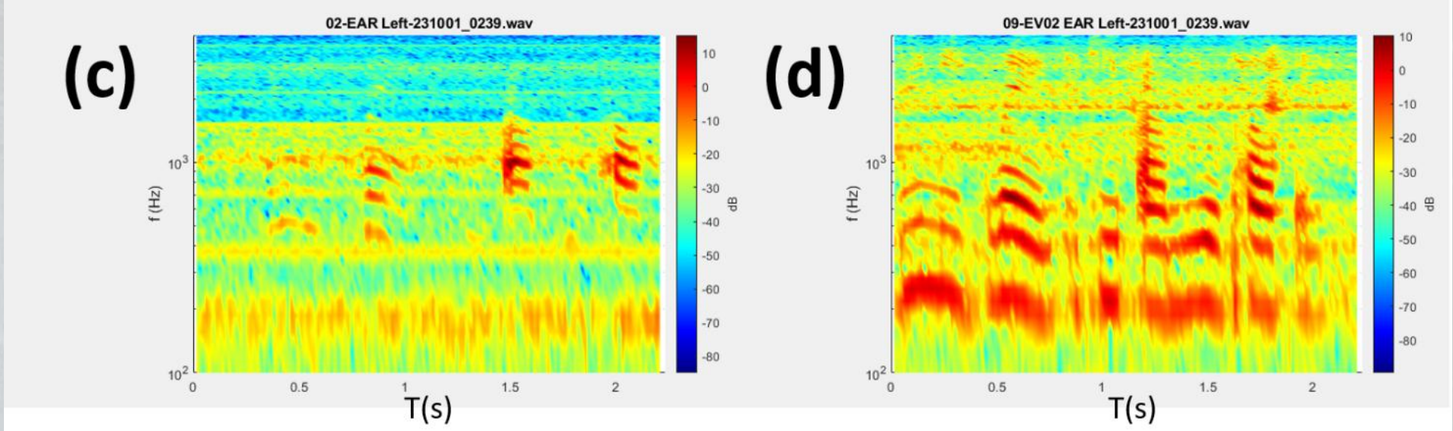
(b) = clean speech

(c) = received speech

(d) = clean speech+PLSS



- Amplitude vs frequency vs time
- Linear/sloped structures = voice
- Patchy higher “blobs” = unvoiced
- Vowels (AEIOU), fricatives (SH, F), Plosives (P, Bu)

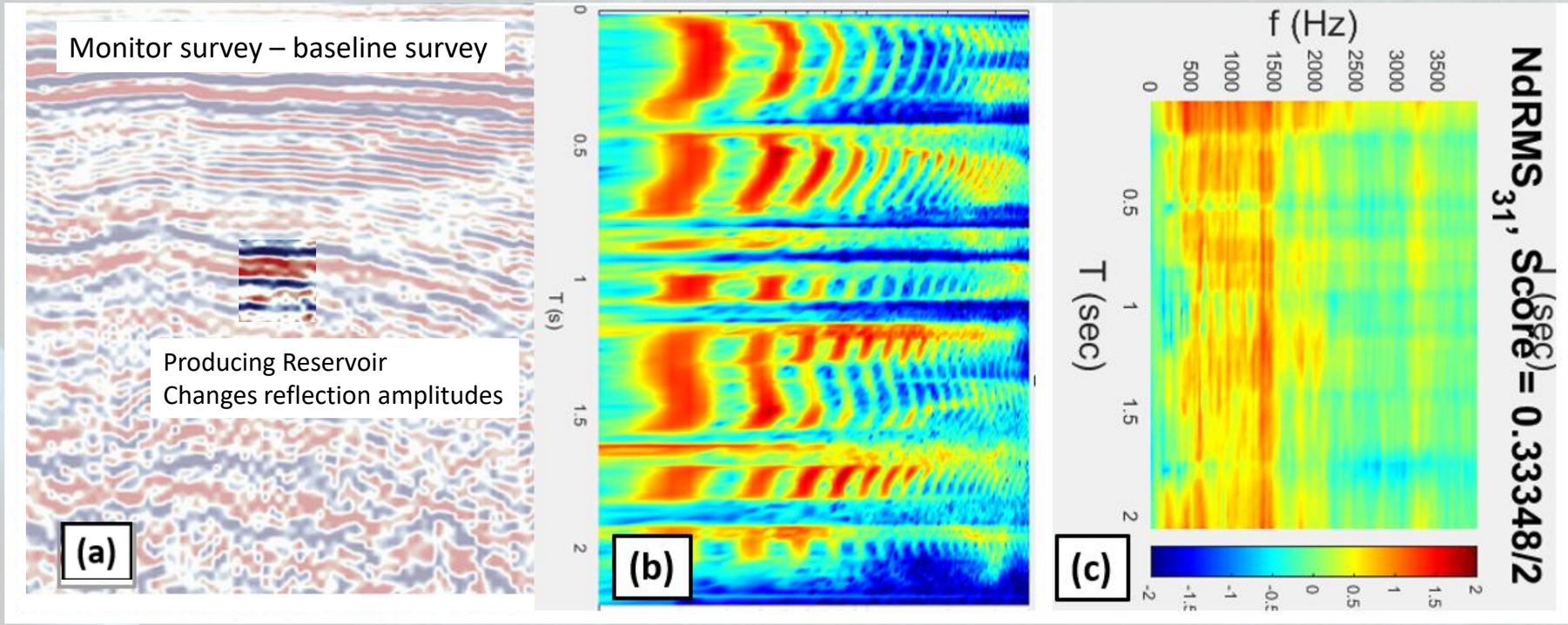


NdRMS: Normalized diff of Root Mean Squares

- 100% objective – compares signals directly
- Any two signals/segments -- unrestricted
- Uses spectrogram frequency traces
- Quality benchmark from -2 to 2
- High information density → advises development/corrections

$$NdRMS = \frac{RMS2 - RMS1}{0.5 * (RMS2 + RMS1)}$$

$$Score = \frac{\sum_{k=0}^{n} | NdRMS(f,t)_k |}{2 * n}$$

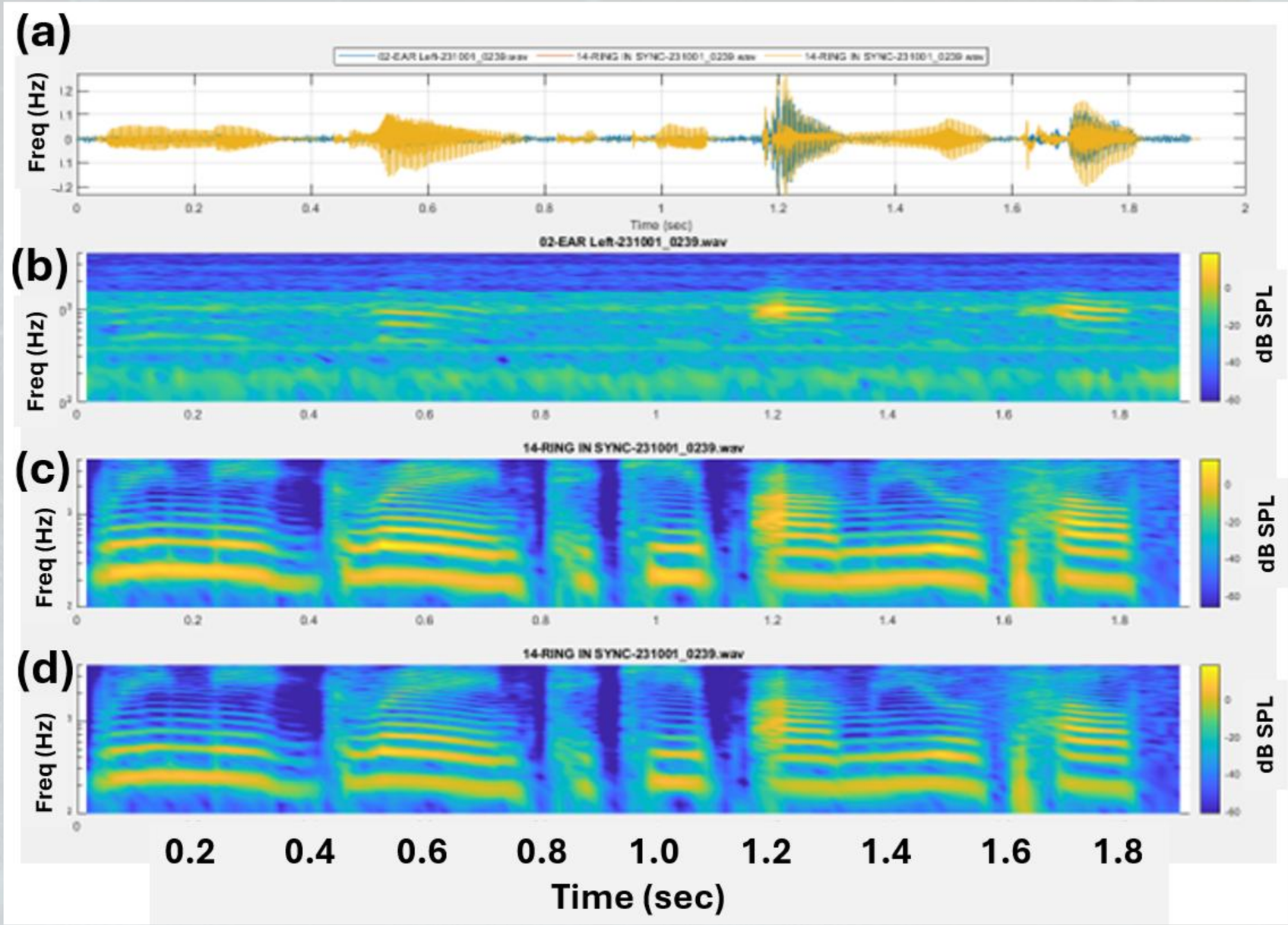


- (a) Compare seismic images: amplitude diff shows oil/water movement
- Applied to audio spectrograms (b) shows signal diff in freq and time (c)

Seismic: https://www.usgs.gov/media/images/seismic-reflection-profile-0

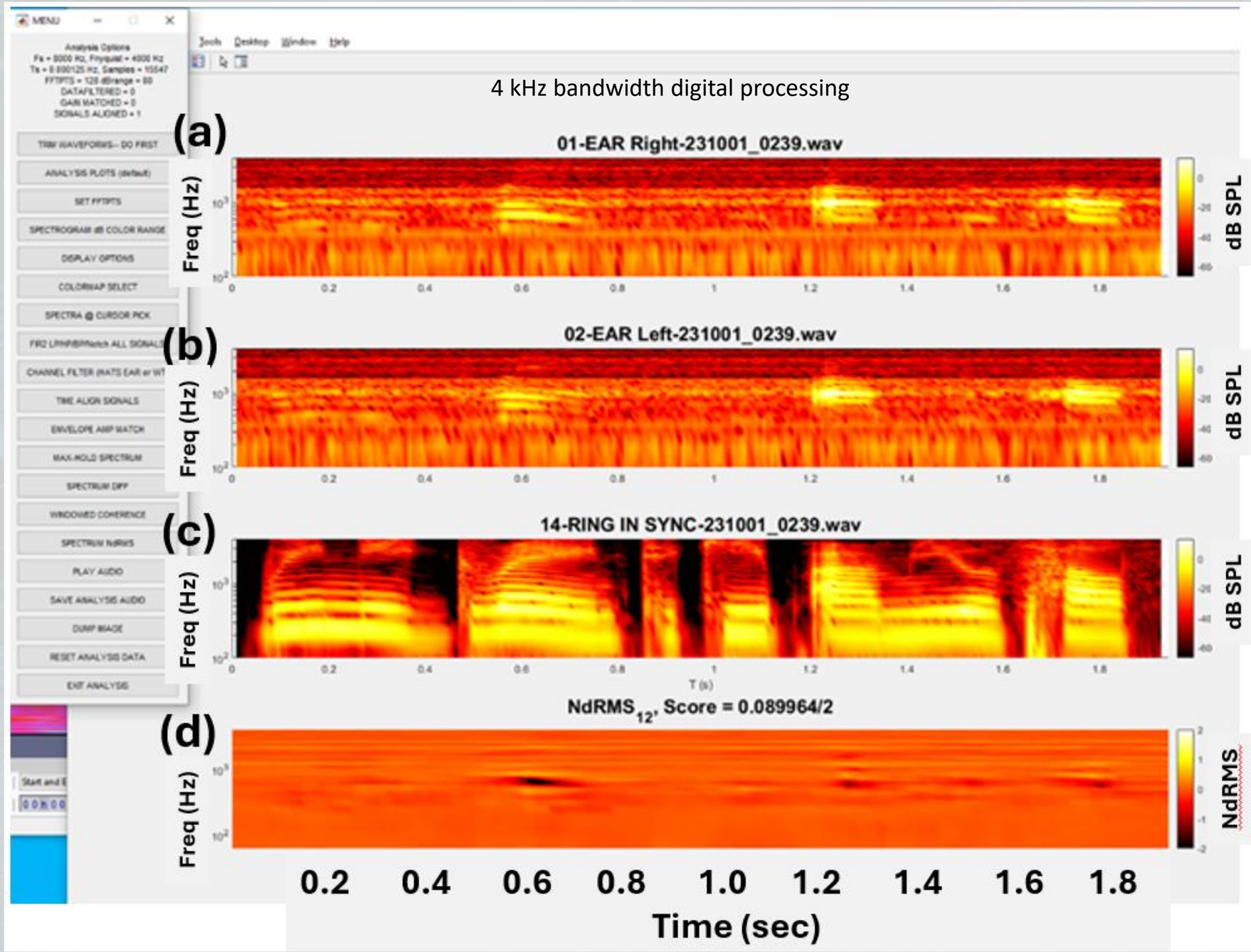
Time alignment & amplitude scaling

- (a) Time aligned, amplitude matched spoken and recorded speech waveforms
- (b) EV01 SxEMU received speech (b) (noisy) – uses same FFT points, window, overlap as input speech (c & d)



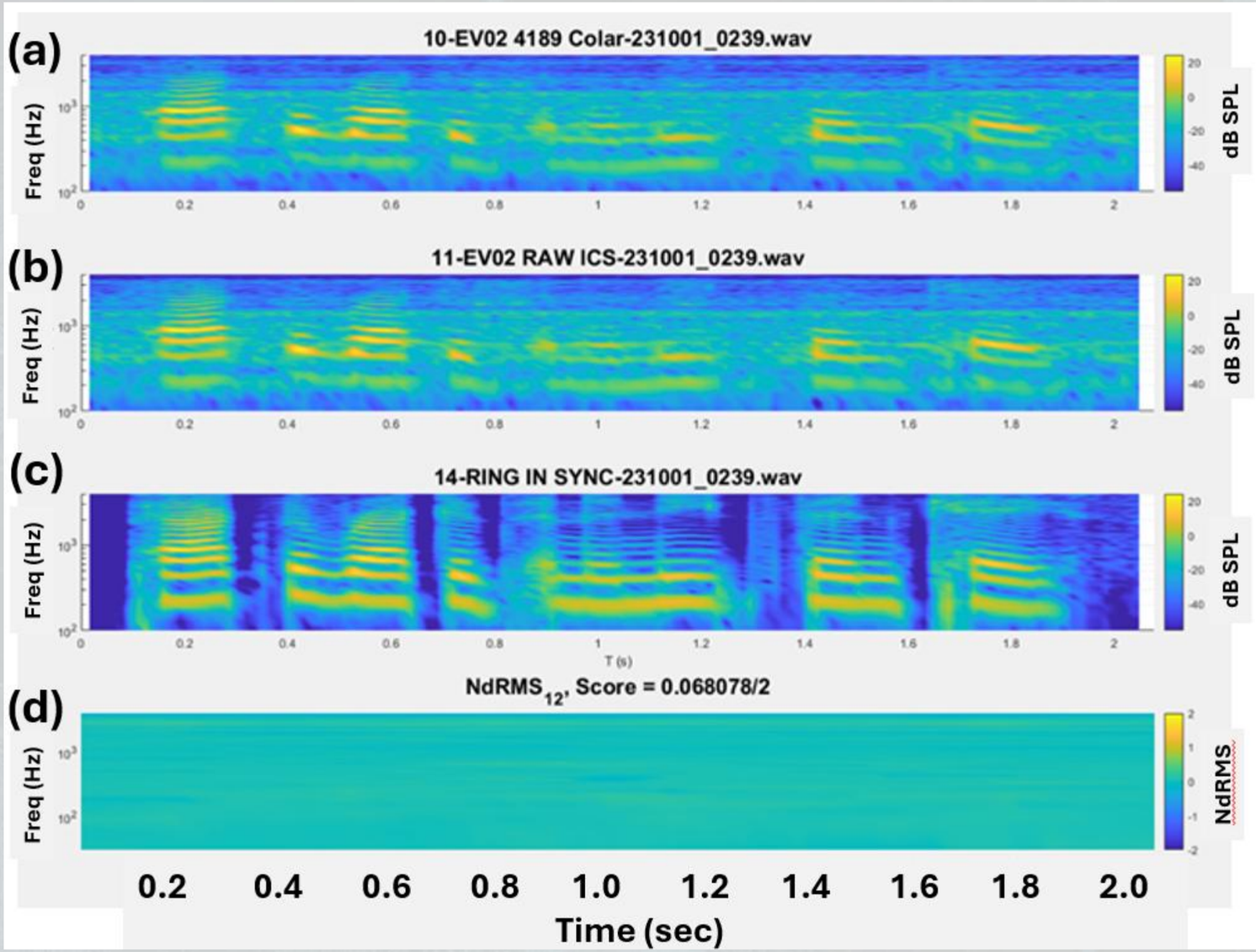
NdRMS “suit tour” – receiving manikin ears

- Manikin ear “tracking”
- (a) and (b) received speech at SxEMU L/R ears – note DSP bandlimited
- (c) EV02 manikin input speech – clean – full bandwidth
- (d) NdRMS=(-2) artifacts + hi-freq variations



Speaking manikin mic comparison (EV02)

- Compares suit mics to B&K measurement mic
- (a) B&K 4189 measurement mic – flat frequency response 10-10,000 Hz
- (b) MEMS I2S digital microphones (3) on suit collar
- (c) clean input speech to EV02 manikin
- (d) Low NdRMS – confirms POLQA 4189 & ICS mic response agreement

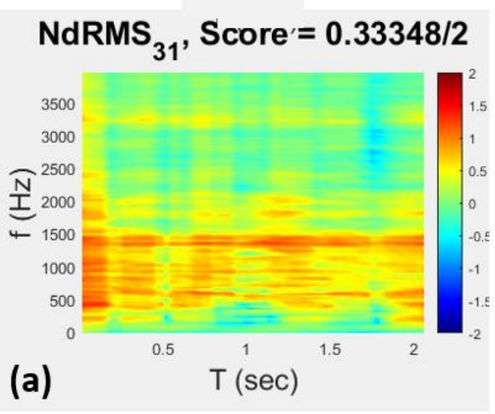


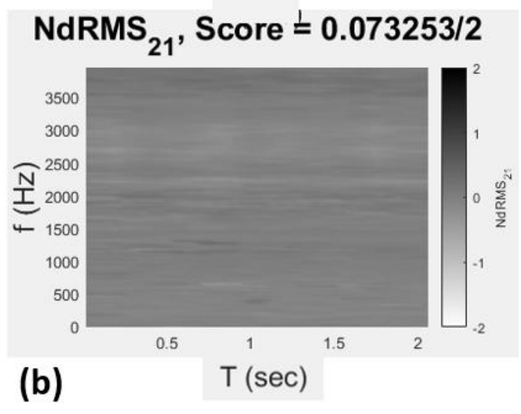
"Is there really no difference between the microphones?"

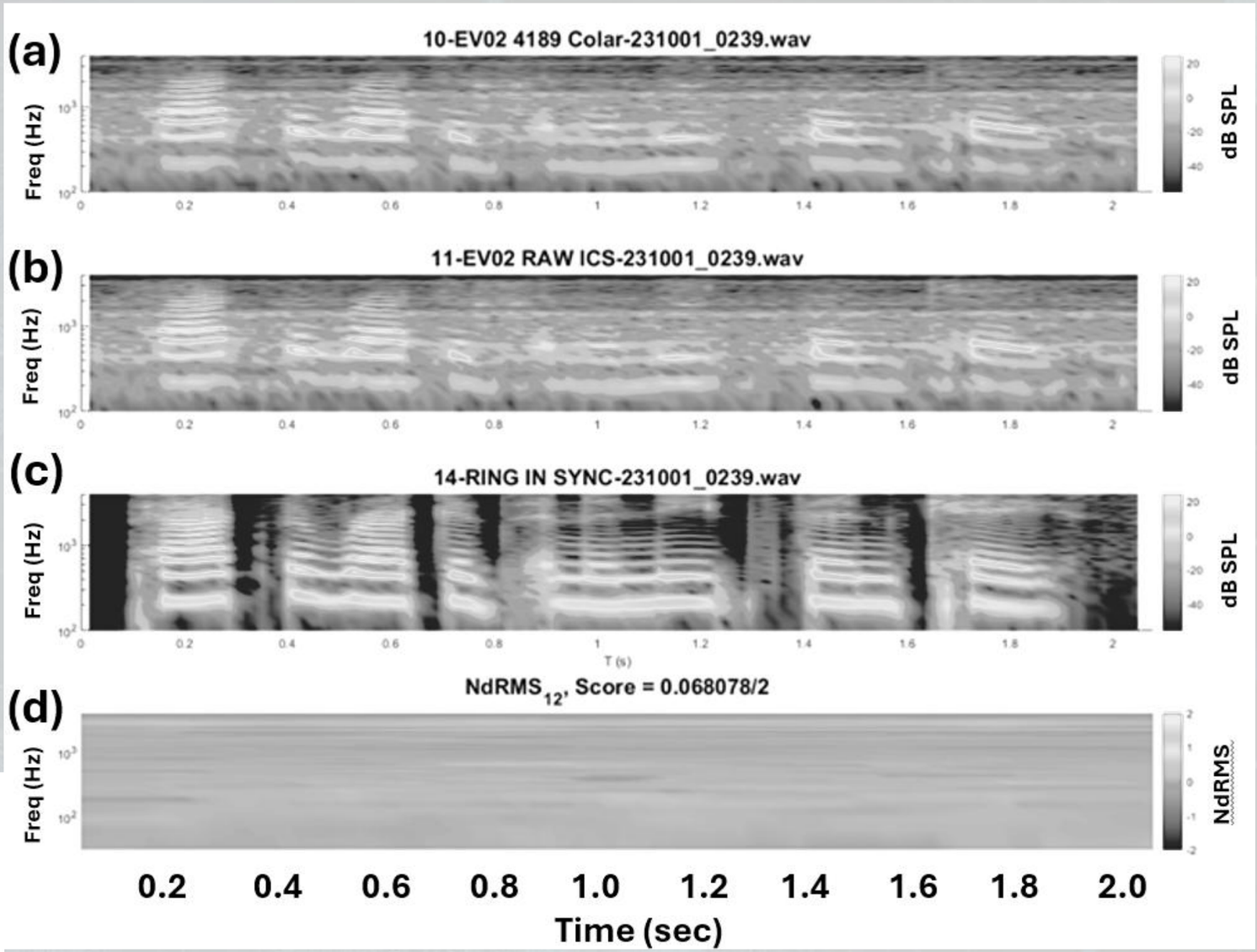
Greyscale reveals minor 4189/ICS differences

- Grayscale + saturation shift reveals some time dependent variation between mics
- ICS mics are distributed
- No direct 4189 vs suit one-for-one co-location

Use colormap changes to show small variations

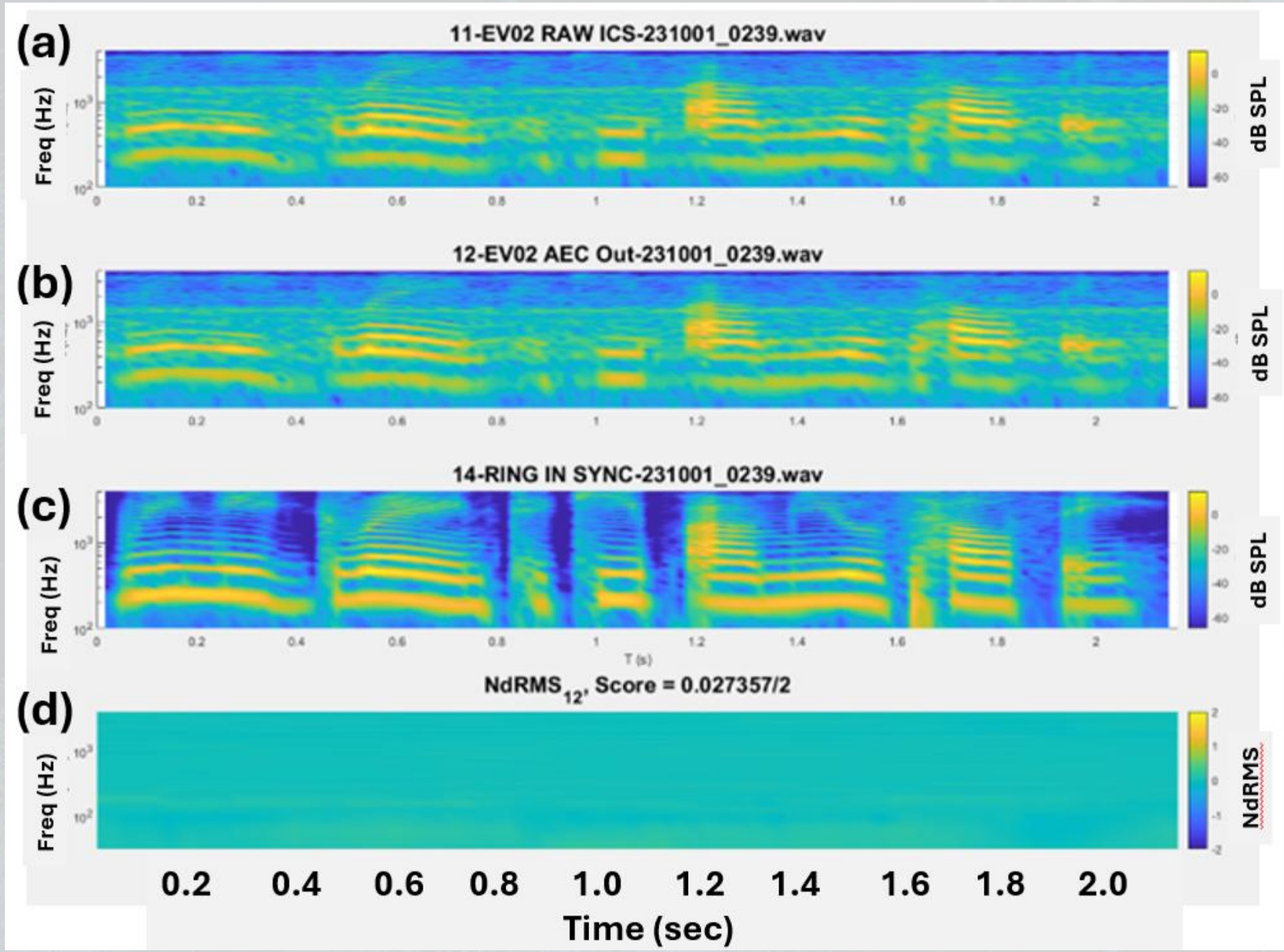






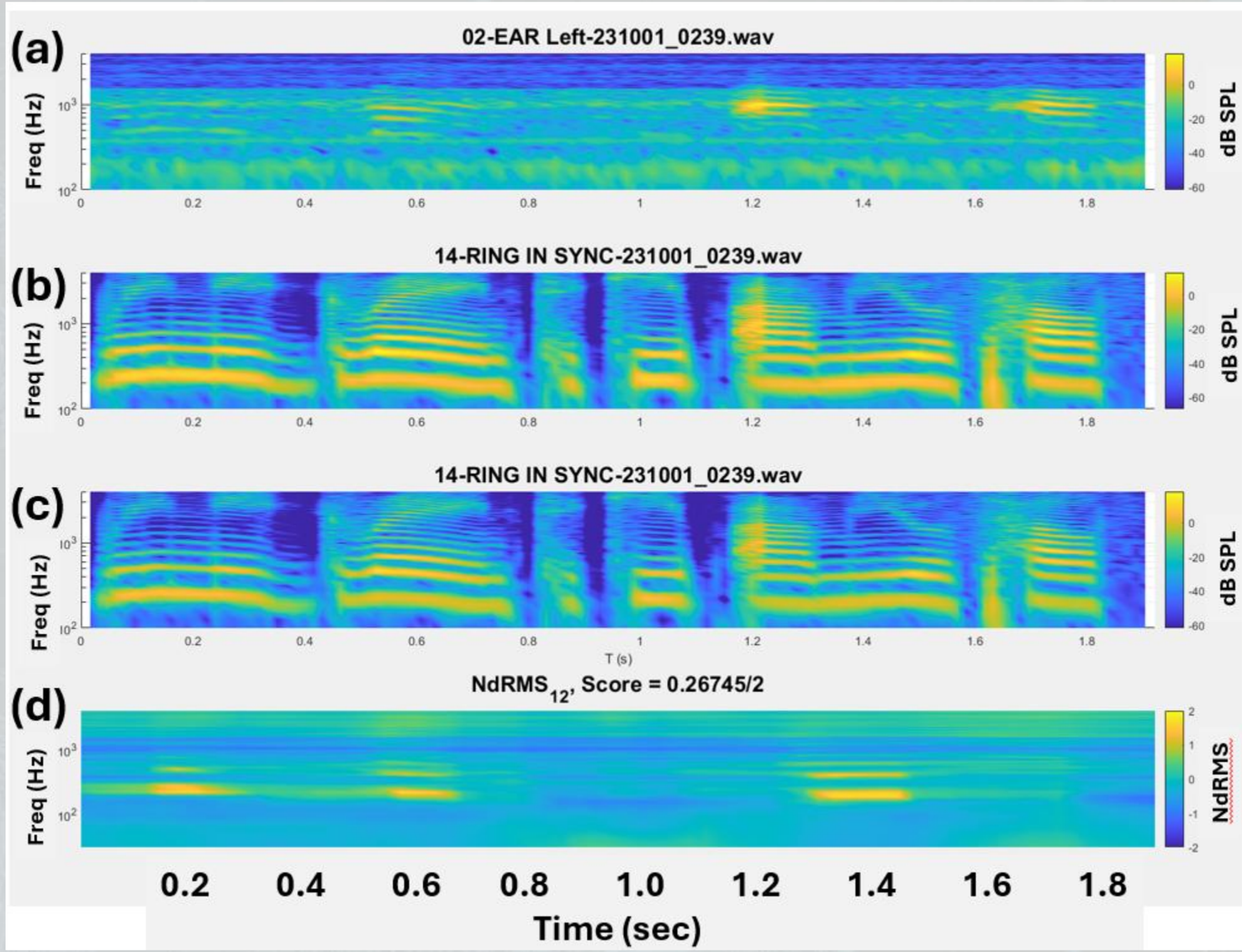
EV02 Electronics: Suit Mics to COMMS output

- (c) Clean input to EV02 transmitting manikin mouth
- (a) EV02 manikin/crew spectrogram recorded by microphones
- (b) Manikin/crew spectrogram out to comms
- (d) Effectively no distortion from audio electronics, NdRMS → 0



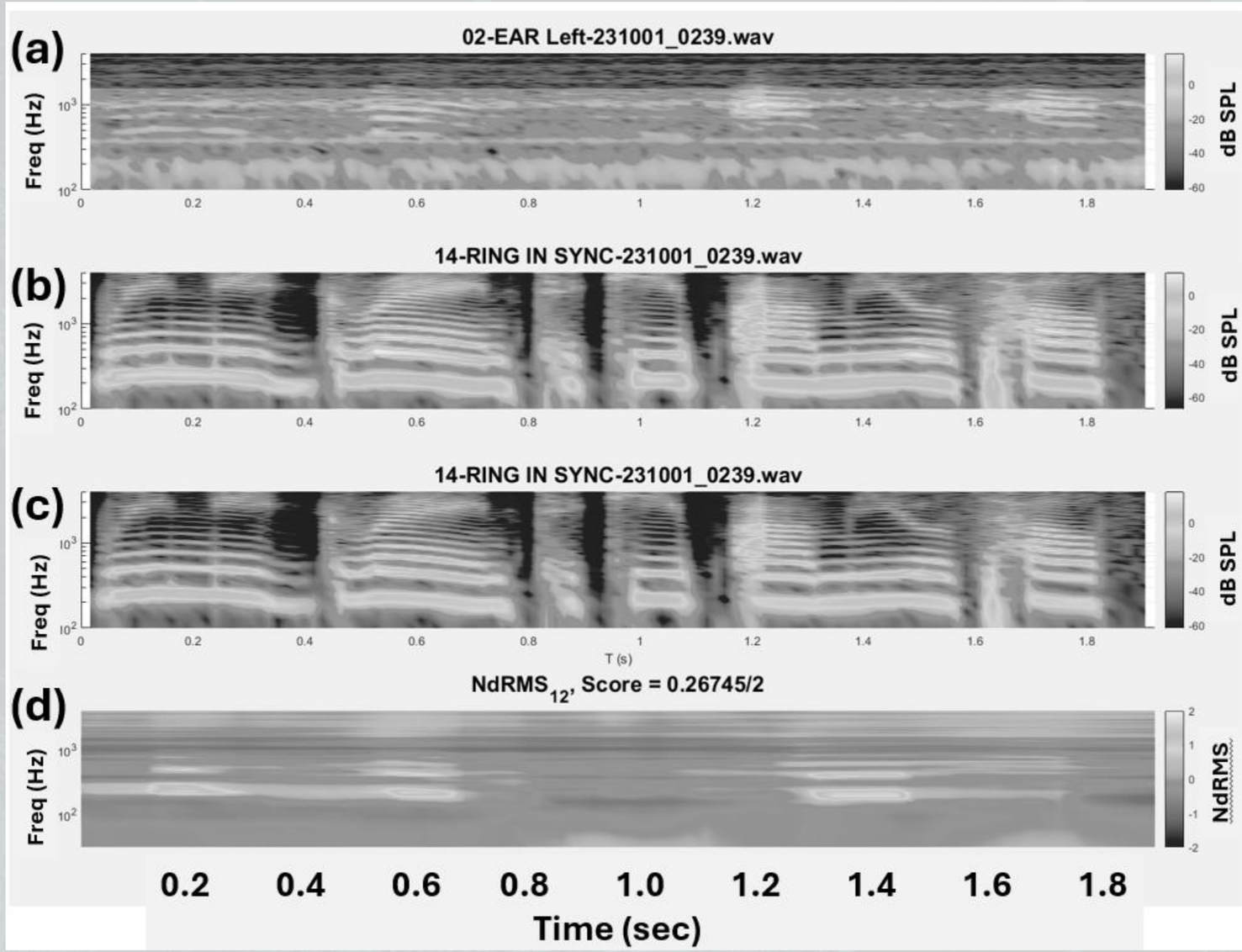
NdRMS: Manikin mouth to receiving suit ear

- End-to-End
- (a) received speech spectrogram at EV01 ear
- (b and c) SAME transmitting EV02 manikin clean speech
- (d) NdRMS of (a) vs (b)
- Observations
  - NdRMS (0.26/2.0) = 13%
  - Low level distortion of strong vocal speech (vowels) at 0.2, 0.6, 1.4 seconds, approx. 400 Hz
  - Continuous strong distortion at approx. 1 kHz (fan, water pump)



NdRMS: Manikin mouth to receiving suit ear

- End-to-End
- (a) received speech spectrogram at EV01 ear
- (b and c) SAME transmitting EV02 manikin clean speech
- (d) NdRMS of (a) vs (b)
- Observations
  - NdRMS (0.26/2.0) = 13%
  - Low level distortion of strong vocal speech (vowels) at 0.2, 0.6, 1.4 seconds, approx. 400 Hz
  - Continuous strong distortion at approx. 1 kHz (fan, water pump)



Same as previous slide, but with grayscale enhancement

Summary: Suit comms and analysis methods

- Two NASA suits in open comms loop with PLSS/air, speakers, mics, AEC – no commcaps
- Modified Rhyme Test (MRT): Subjective Intelligibility
  - Single 63% and 88% scores for high- and low-suit noise cases (90% target)
  - Low information density
- POLQA: Quality model from extensive listener studies (“pseudo-objective”)
  - Single “Mean Objective Score” (e.g. 4.11/5.0): Low information density
  - End-to-end and by-segment use to isolate comms issues
  - Significant degradation from transmitting suit helmet + PLSS noise
  - Suit and measurement microphone response agreement
  - Low or insignificant degradation from audio electronics
- NdRMS: Objective Quality
  - Normalized difference score from -2 to 2
  - Time/frequency representation → high information density guides engineering decisions
  - Unrestricted application for any two signals/spectrograms/comms segments
  - Agrees with POLQA – helmet noise, mic match, audio quality
- Future work: NdRMS to POLQA score mapping

Questions?





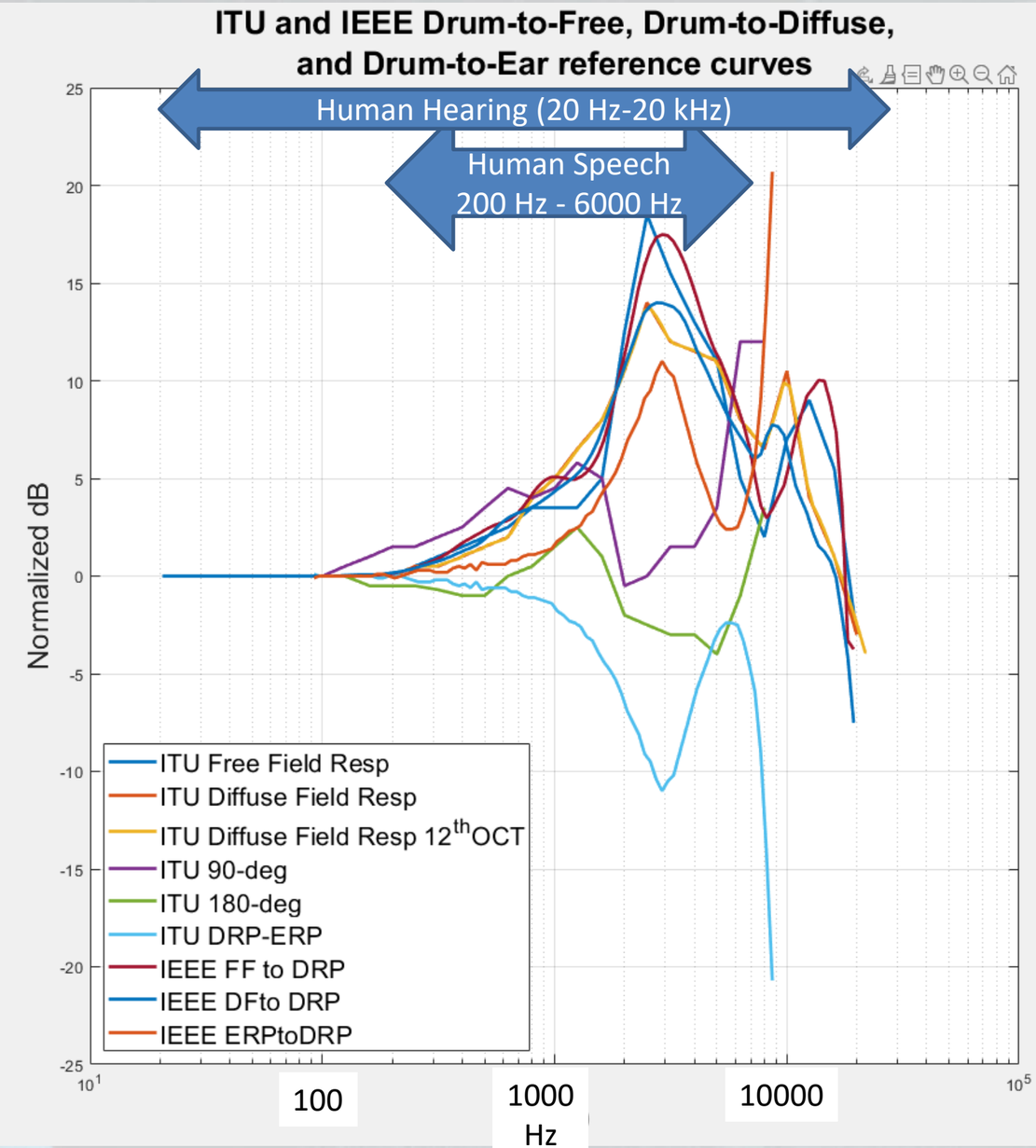
amentum

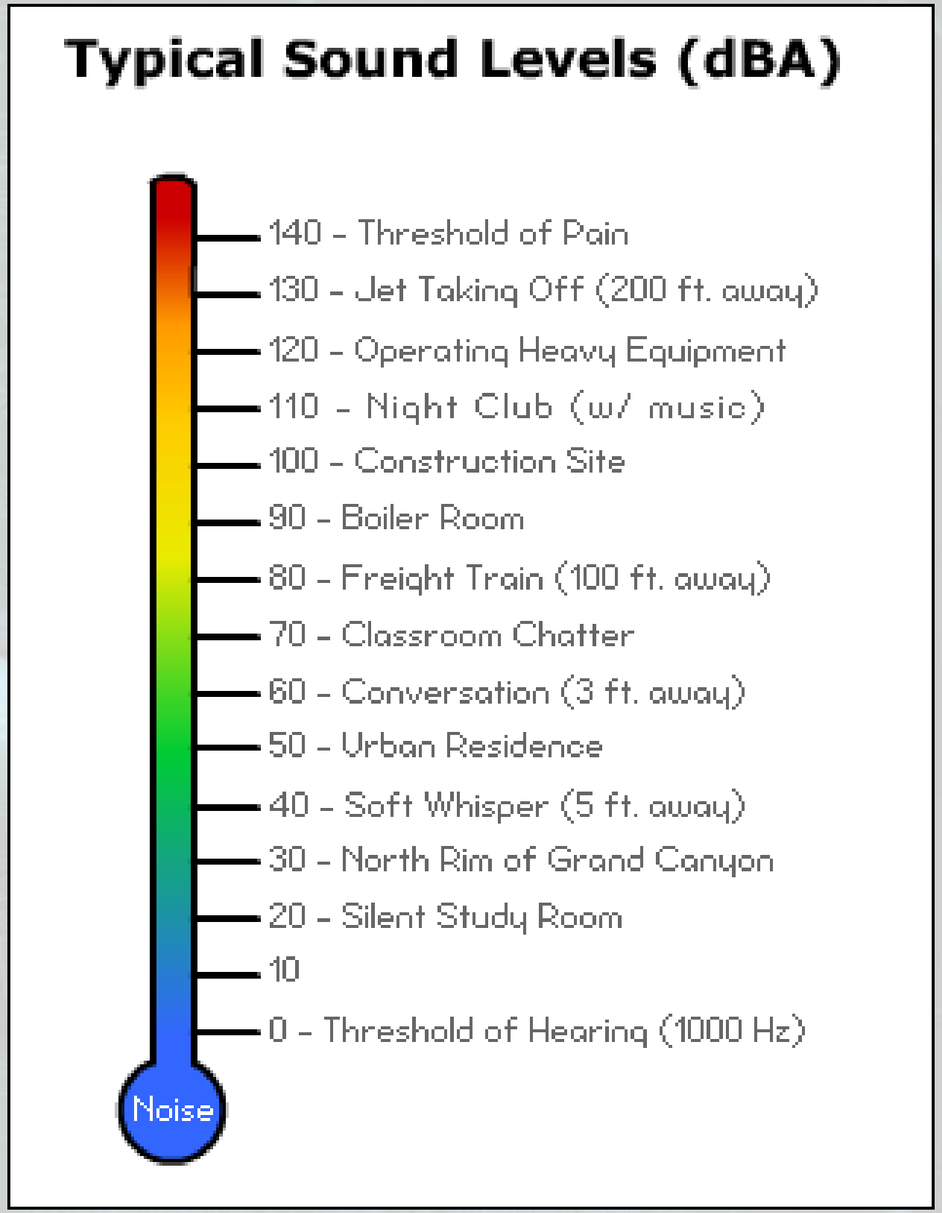


GEOCONTROL SYSTEMS, INC.

Contact Info:
ICES-2026-547, ICES-2024-64
Steven Smith, GeoControl Systems/Jacobs/NASA Audio Dev Lab
Suit audio/DSP/audio-acoustics
steven.s.smith@nasa.gov
“No commcaps were used in the making of this technology”

Sound Levels, corrections, bandwidth





Source: OSHA Technical Manual (OTM) - Section III: Chapter 5 | Occupational Safety and Health Administration