PSCR Speech Intelligibility Work – Past, Present, Future
2014 Public Safety Broadband Stakeholder Meeting
June 5, 2014

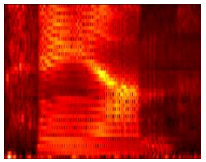
PSCR

Public Safety Communications Research

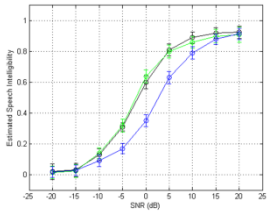
Stephen Voran and Andrew Catellier

Overview

- Speech Intelligibility testing by Modified Rhyme Test (MRT) 2008-2012
  - MRT example
- Automated MRT: ABC-MRT
  - How it works
  - ABC-MRT Demonstration

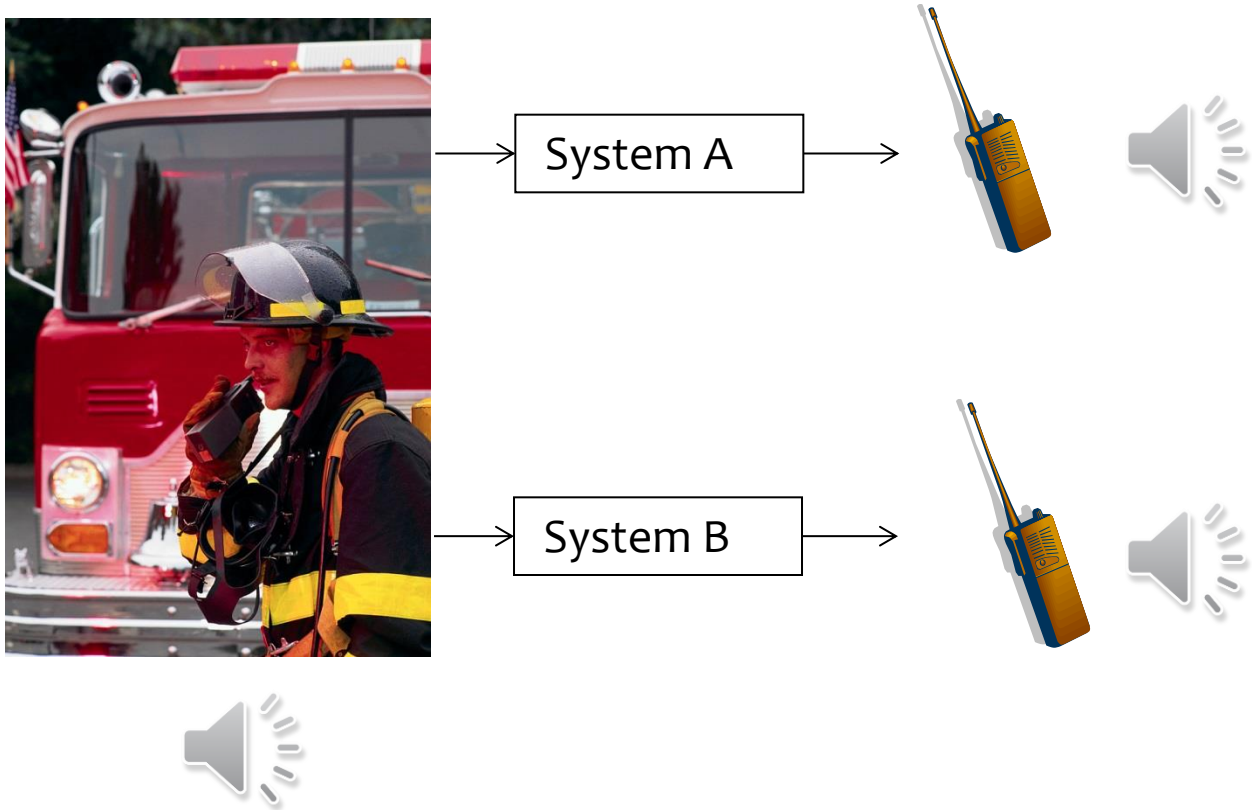


- Upcoming Intelligibility Tests for Mission Critical Voice applications over LTE



PSCR

Intelligibility is Critical



Ability to measure intelligibility is also critical.

PSCR

Assessing Intelligibility

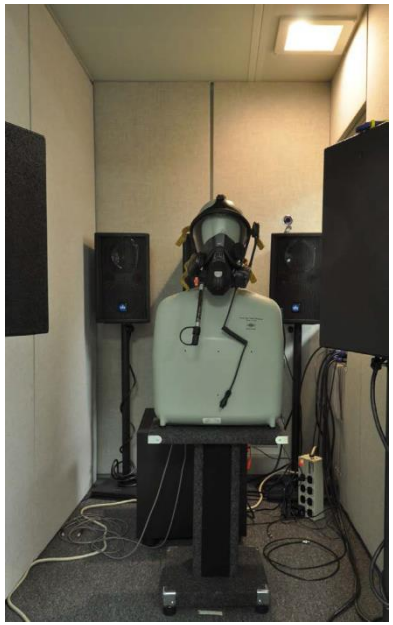
- Many ways to assess intelligibility
- US National Fire Protection Association specifies the ANSI S3.2 Modified Rhyme Test (MRT)
- Stakeholders endorsed this test
- PSCR has used this test since 2008

PSCR

Preparation of Recordings

Background Noise Recordings

Prepare Head and Torso Simulator (HATS)

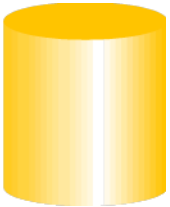


HATS in Sound-Isolated Recording Chamber

MRT Recordings (.wav files)

PSCR

Scoring of Recordings



MRT Recordings

Scoring by First Responders

In Sound-Isolated Playback Chamber

PSCR

MRT Basics

- MRT uses sets of words that sound similar
- Rate of correct word identification forms a measure of intelligibility

PSCR

MRT Trial 1

Please select the word _____.

- ○ Went
- ○ Sent
- ○ Bent
- ○ Dent
- ○ Tent
- ○ Rent

PSCR

MRT Trial 1

Please select the word _____.

- ○ Went
- ○ Sent
- ● Bent
- ○ Dent
- ○ Tent
- ○ Rent

PSCR

MRT Trial 2

Please select the word _____.

- ○ Hold
- ○ Cold
- ○ Told
- ○ Fold
- ○ Sold
- ○ Gold

PSCR

MRT Trial 2

Please select the word _____.

- ○ Hold
- ○ Cold
- ○ Told
- ○ Fold
- ○ Sold
- ● Gold

PSCR

MRT Trial 3

Please select the word _____.

- ○ Pat
- ○ Pad
- ○ Pan
- ○ Path
- ○ Pack
- ○ Pass

PSCR

MRT Trial 3

Please select the word _____.

- ○ Pat
- ○ Pad
- ● Pan
- ○ Path
- ○ Pack
- ○ Pass

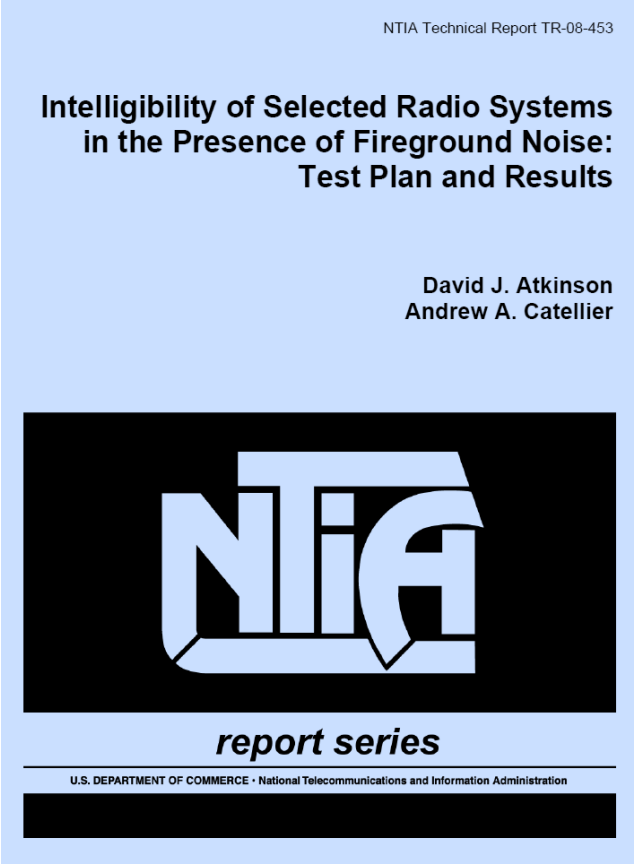
PSCR

PSCR 2008 MRT

30 Scenarios Tested (comm. system x background noise)

P25, Analog FM

http://go.usa.gov/ByTB


NTIA Technical Report TR-08-453

Intelligibility of Selected Radio Systems in the Presence of Fireground Noise: Test Plan and Results

David J. Atkinson
Andrew A. Catellier

report series

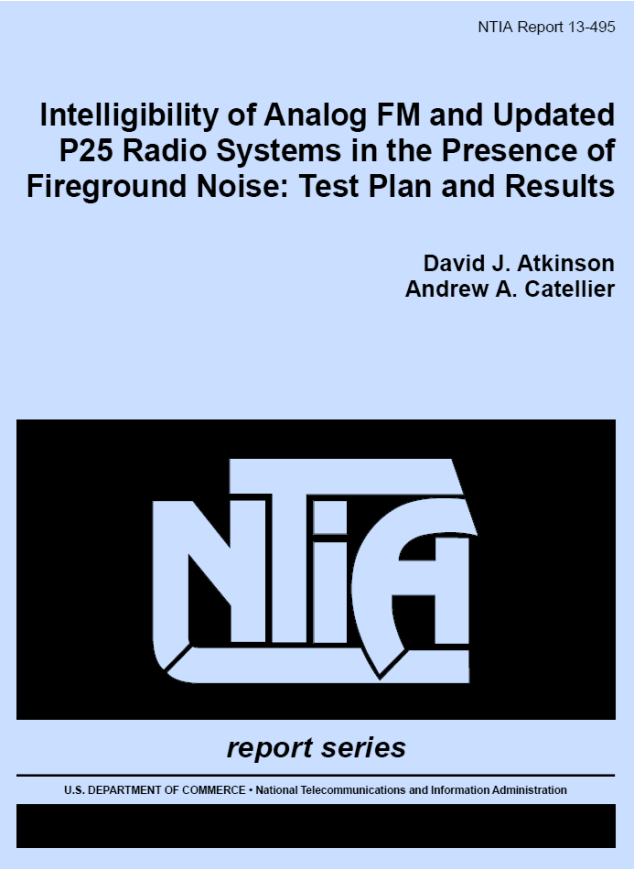
U.S. DEPARTMENT OF COMMERCE • National Telecommunications and Information Administration

MSCR MRT 2010

56 Scenarios Tested

P25, Analog FM

http://go.usa.gov/ByTQ


NTIA Report 13-495

Intelligibility of Analog FM and Updated P25 Radio Systems in the Presence of Fireground Noise: Test Plan and Results

David J. Atkinson
Andrew A. Catellier

report series

U.S. DEPARTMENT OF COMMERCE • National Telecommunications and Information Administration

PSCR

PSCR MRT 2012

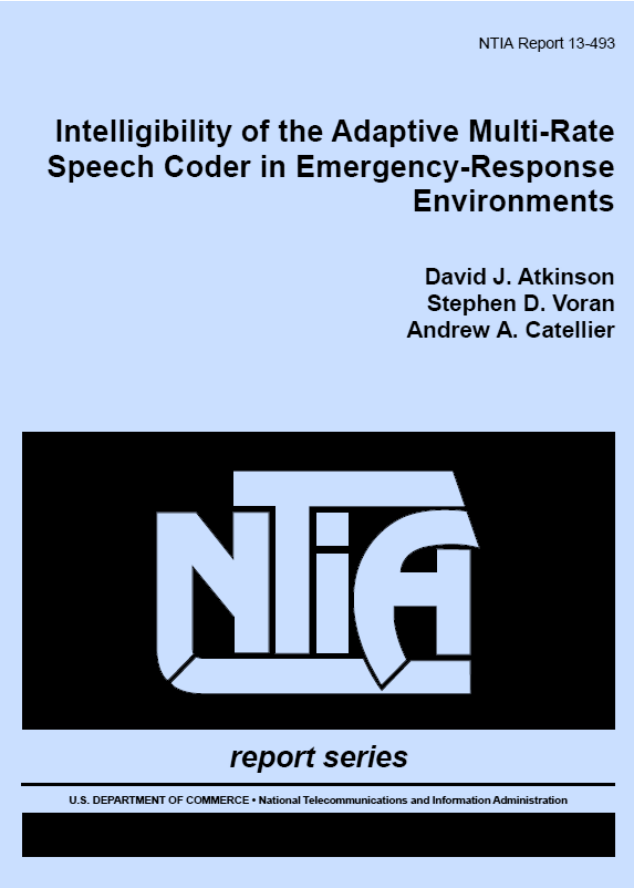
28 Scenarios Tested

AMR, P25, Analog FM

AMR 12.2 kbps intelligibility is same as or higher than P25 4.4 kbps (ver 1.6)

AMR 12.2 kbps intelligibility is higher than Analog FM in 3 noise conditions, but lower in 2 noise conditions

http://go.usa.gov/ByTw


NTIA Report 13-493

Intelligibility of the Adaptive Multi-Rate Speech Coder in Emergency-Response Environments

David J. Atkinson
Stephen D. Voran
Andrew A. Catellier

report series

U.S. DEPARTMENT OF COMMERCE • National Telecommunications and Information Administration

PSCR

MRT Reflections ↑

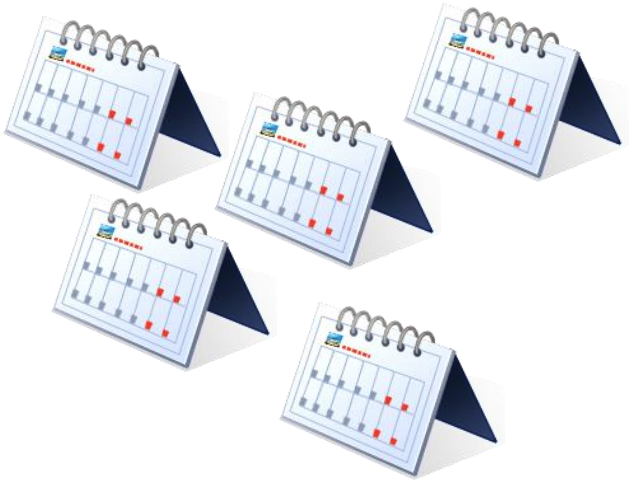
- PSCR MRT work has provided valuable information
- Key results for low rate speech coding in high noise environments
- 3 reports, many thousands of speech files and MRT scores
- All available at pscr.gov

PSCR

MRT Reflections ↓

- MRT requires intensive and extensive effort

- E.g. 2008 MRT with 30 scenarios
  used 54,000 Trials (1200 trials per subject x 45 subjects)



PSCR

Observation: MRT is Speech Recognition

Can we use automatic speech recognition (ASR) to automate the MRT?

PSCR

Yes, but we have to Start Over

ASR Goals

- Large vocabulary
- Speaker independence
- Maximally robust to impairments

- Mel Cepstrum, Δ's, …
- HMM, DNN

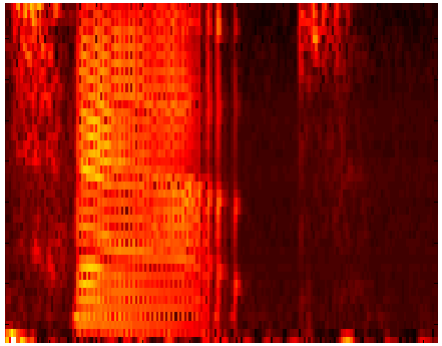
MRT Goals

- 6 words
- A few specific speakers
- Robustness similar to human robustness

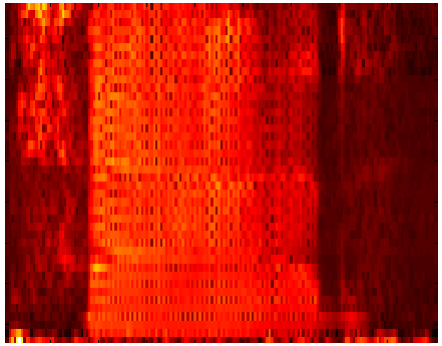
- Articulation Index (AI) Bands
- Correlate loudness patterns across time
- Articulation Index Band Correlation MRT (ABC-MRT)

PSCR

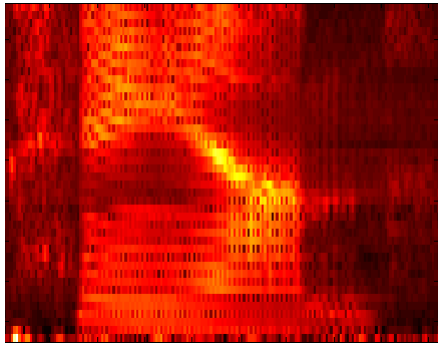
ABC-MRT T-F Patterns
6 Original Words



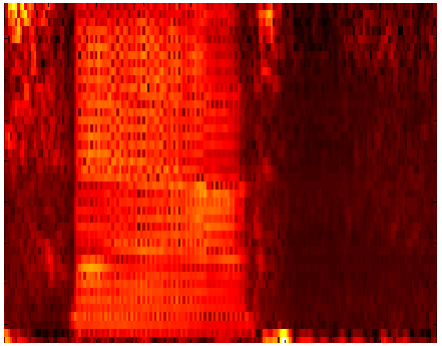
Pat



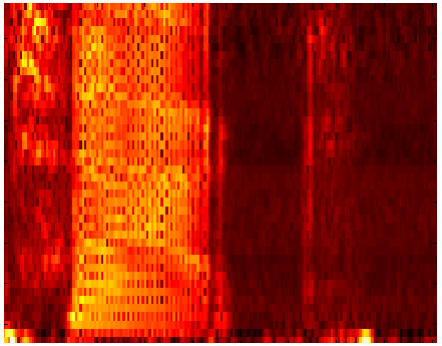
Pad



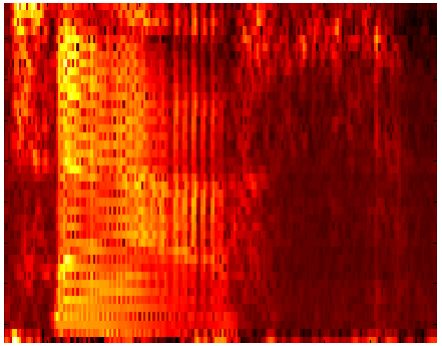
Pan



Path



Pack

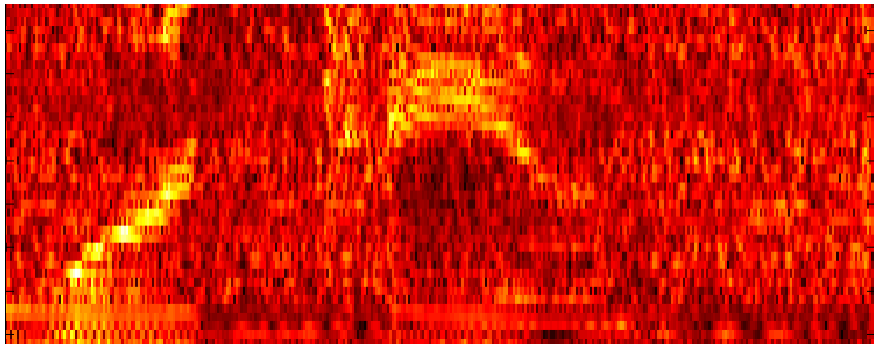


Pass

PSCR

Match SUT Signal to One of Six Originals

- Correct answer is known
- But what does the algorithm say?

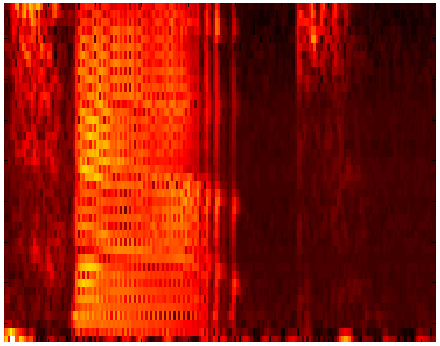


SUT Signal

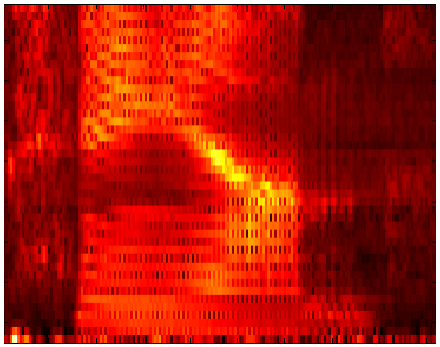
?

?

?

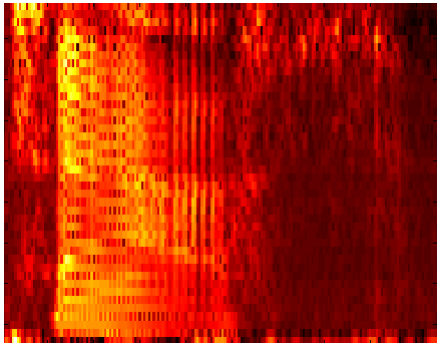


Pat



Pan

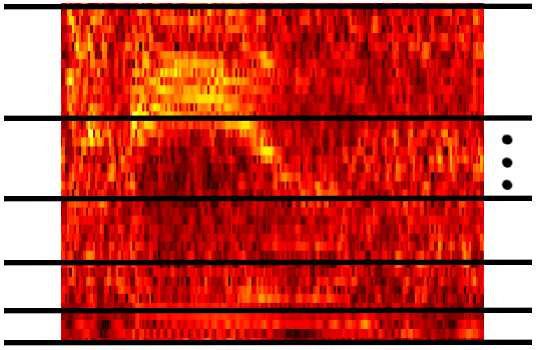
• • •



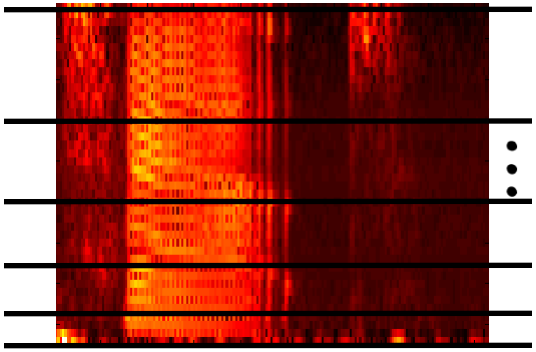
Pass

PSCR

AI Band Correlations across Time



SUT Signal

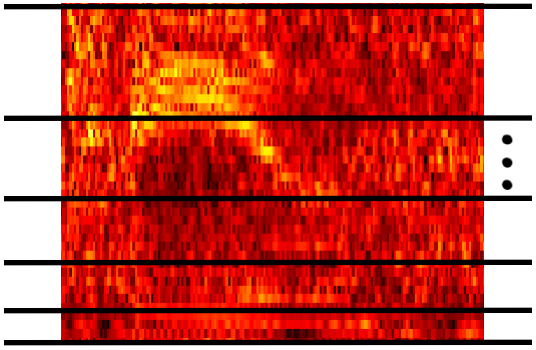


Pat

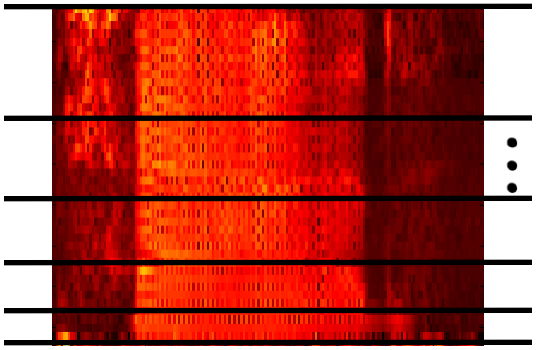
AI Band	Pat
1	.55
2	.53
3	.36
…	…
17	-.07

PSCR

AI Band Correlations



DUT Signal

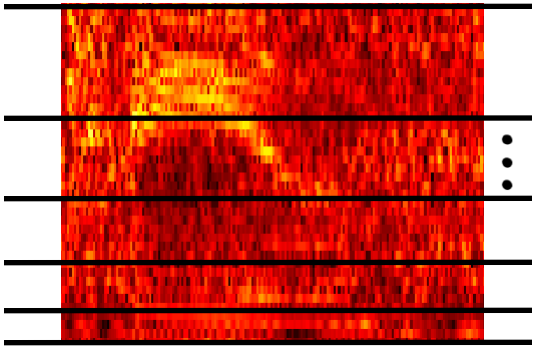


Pad

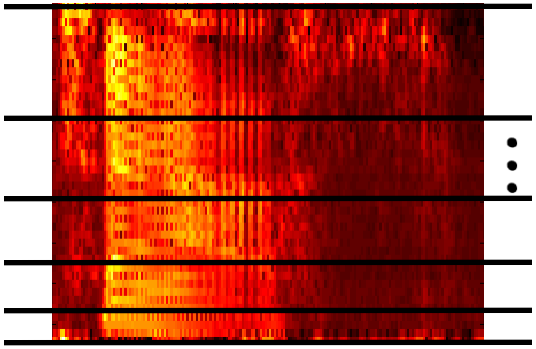
AI Band	Pat	Pad
1	.55	.37
2	.53	.55
3	.36	.41
…	…	…
17	-.07	.30

PSCR

AI Band Correlations



DUT Signal



Pass

AI Band	Pat	Pad	Pan	Path	Pack	Pass
1	.55	.37	.33	.62	.58	.54
2	.53	.55	.72	.43	.60	.59
3	.36	.41	.59	.54	.46	.47
…	…	…	…	…	…	…
17	-.07	.30	.26	.11	.03	-.17

PSCR

Select Highest Correlation in each AI Band

AI Band	Pat	Pad	Pan	Path	Pack	Pass	Winning Word
1	.55	.37	.33	.62	.58	.54	Path
2	.53	.55	.72	.43	.60	.59	Pan
3	.36	.41	.59	.54	.46	.47	Pan
…	…	…	…	…	…	…	…
17	-.07	.30	.26	.11	.03	-.17	Pad

PSCR

Correct Answer is "Pan"

AI Band	Pat	Pad	Pan	Path	Pack	Pass	Winning Word
1	.55	.37	.33	.62	.58	.54	Path
2	.53	.55	.72	.43	.60	.59	Pan
3	.36	.41	.59	.54	.46	.47	Pan
…	…	…	…	…	…	…	…
17	-.07	.30	.26	.11	.03	-.17	Pad

PSCR

Find Success Rate

AI Band	Pat	Pad	Pan	Path	Pack	Pass	Winning Word	Correct?
1	.55	.37	.33	.62	.58	.54	Path	N
2	.53	.55	.72	.43	.60	.59	Pan	Y
3	.36	.41	.59	.54	.46	.47	Pan	Y
…	…	…	…	…	…	…	…	…
17	-.07	.30	.26	.11	.03	-.17	Pad	N

$$c = \frac{6}{17} = 0.35$$

PSCR

Apply Correction for Guessing

AI Band	Pat	Pad	Pan	Path	Pack	Pass	Winning Word	Correct?
1	.55	.37	.33	.62	.58	.54	Path	N
2	.53	.55	.72	.43	.60	.59	Pan	Y
3	.36	.41	.59	.54	.46	.47	Pan	Y
…	…	…	…	…	…	…	…	…
17	-.07	.30	.26	.11	.03	-.17	Pad	N

$$c = \frac{6}{17} = 0.35$$

$$c' = \frac{6}{5}(c - \frac{1}{6}) = 0.22$$

Estimated Intel. (scale runs 0 to 1)

PSCR

ABC-MRT Process Summary

- Play MRT recording into System under Test
- Record output of SUT
- Apply the ABC-MRT algorithm
- Corrected success rate is the estimated intelligibility
- Repeat with different talkers and word sets
- Calculate average and uncertainty for scenario

PSCR

Portable Real-Time ABC-MRT Demo

System under Test is

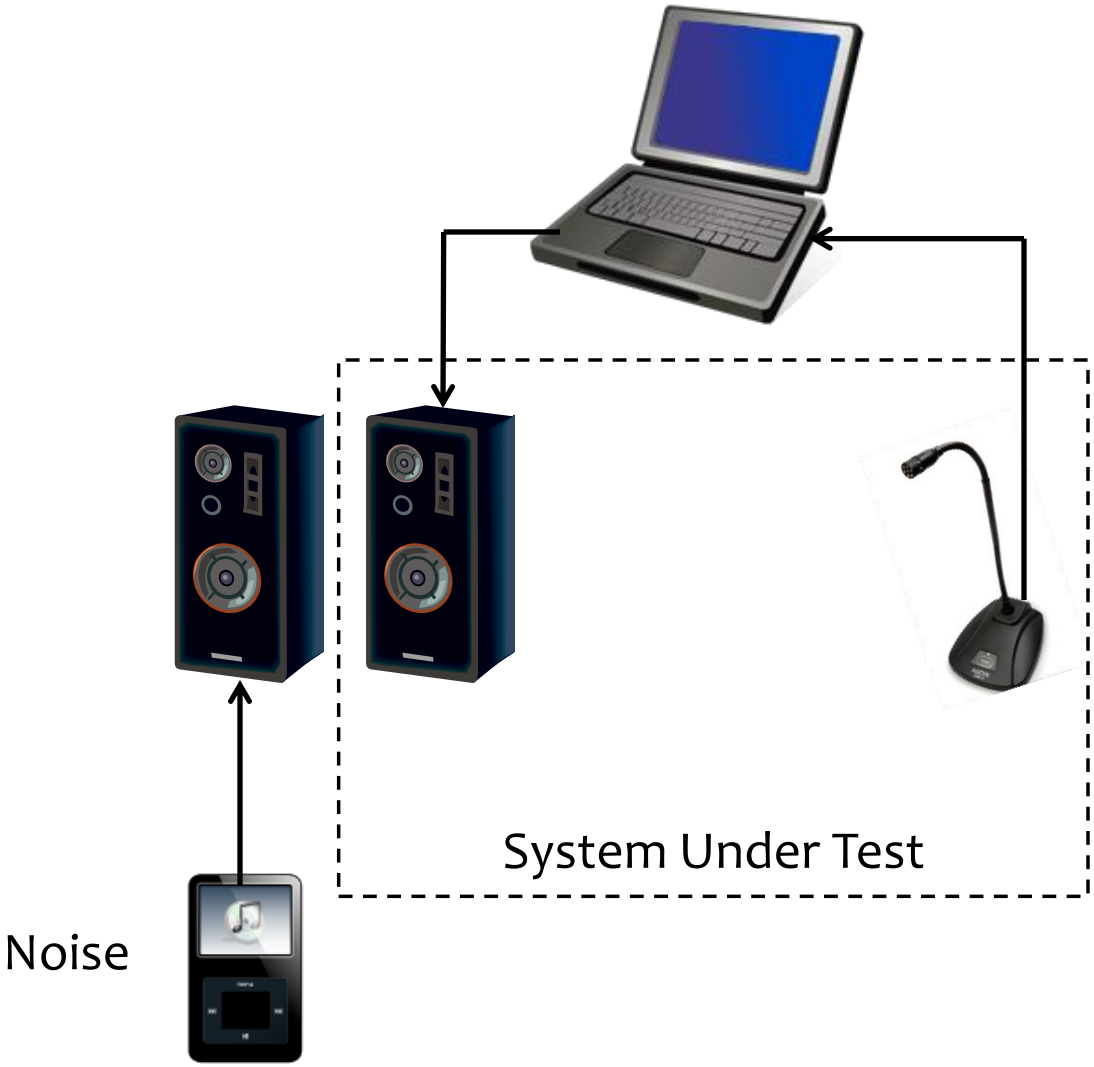
- Speaker
- Acoustic Path
- Microphone

PSCR

Portable Real-Time ABC-MRT Demo

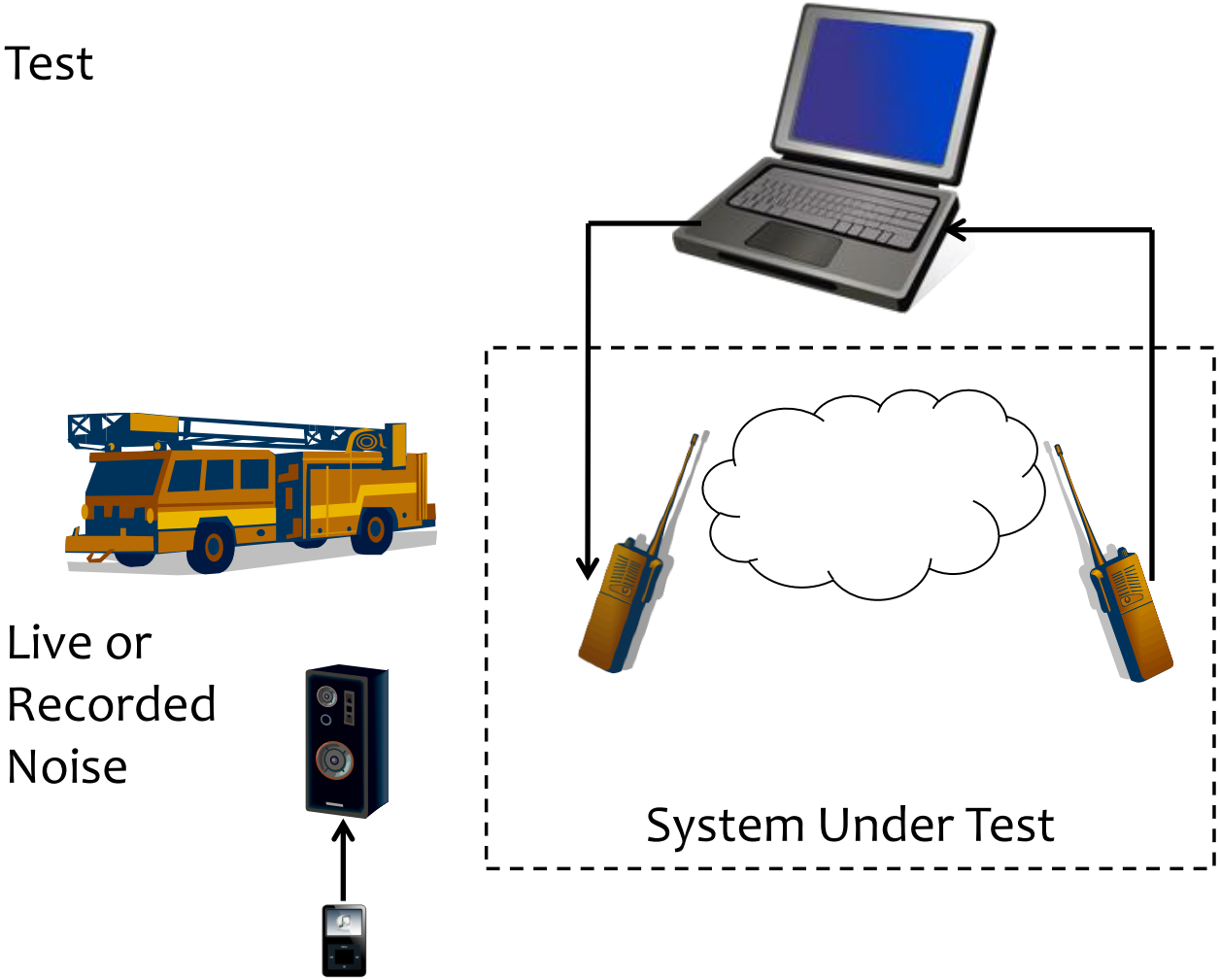
System under Test is

- Speaker
- Acoustic Path
- Microphone
- Noise



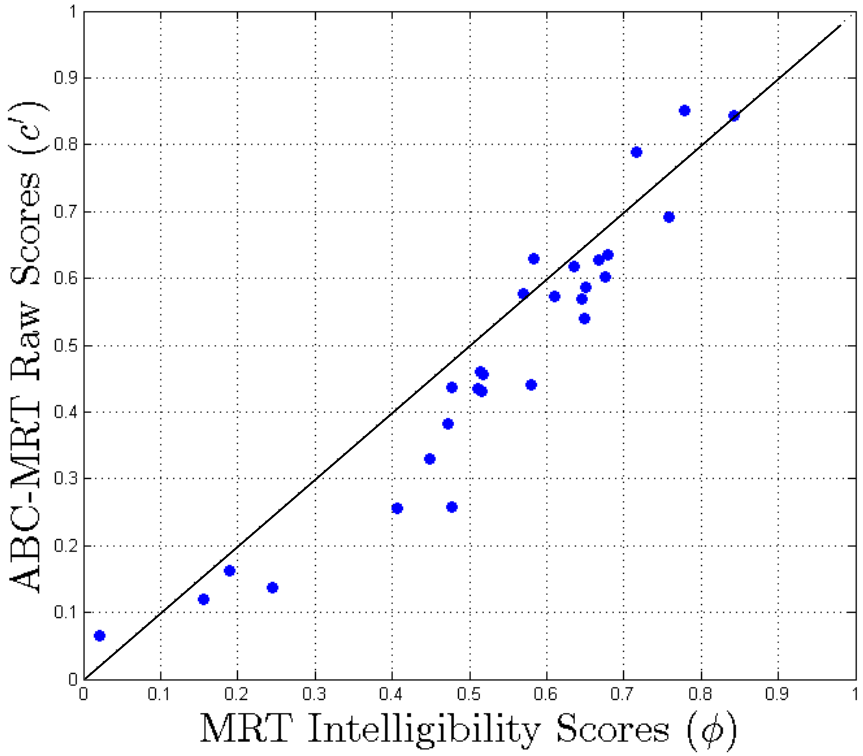
Example ABC-MRT Application

Real System under Test



How good are the Estimates?

- Development used PSCR 2012 Test
- 28 scenarios, .02 ≤ MRT ≤ .84
- 4 types of speech coding, 7 acoustic environments

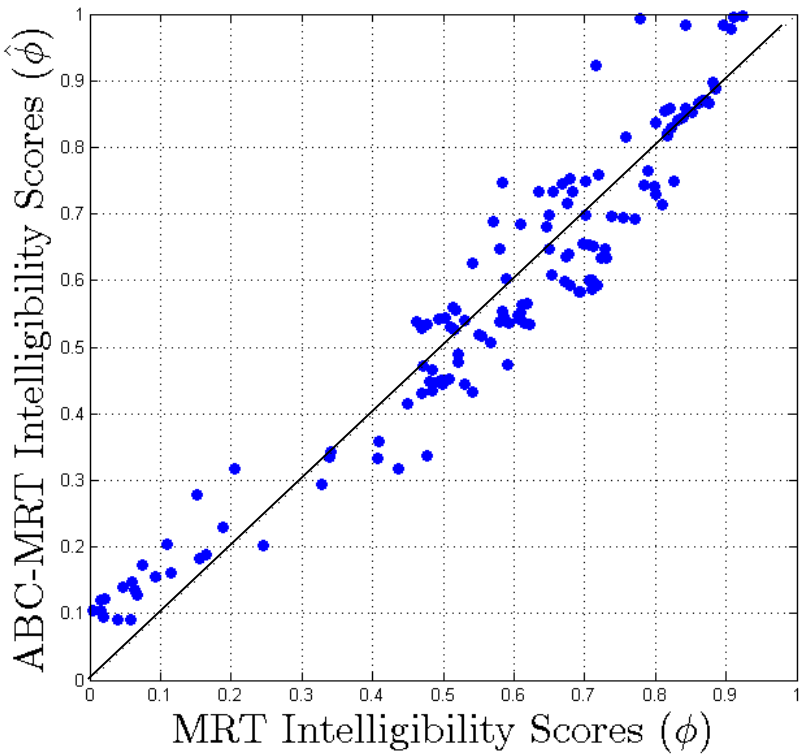


- Pearson Correlation 0.950
- RMSE 0.084
- No optimized parameters

PSCR

How good are the Estimates?

- Evaluation using all available data
- Total of 139 scenarios, 71% of data unseen
- Unique acoustic conditions, speech coding, amplifier overload, radio channel impairments with analog and digital radios
- Apply adjustment: $\hat{\phi} = 1.109 \cdot c' + 0.050$



- Pearson Correlation 0.955
- RMSE 0.073
- Only 2 optimized parameters
- ABC-MRT Paper and Software at http://go.usa.gov/BvRQ

PSCR

What does it Mean?

- The MRT scoring step can be approximated by ABC-MRT
  - Replace 15 to 45 persons and playback chamber with one computer
  - Reduce MRT scoring time from months to hours
  - Reduce MRT scoring costs
- We expect RMS estimation error of 7.3% of full scale for 139 scenarios observed so far and for similar ones
  - Current systems (P25, Analog FM)
  - Potential future systems (AMR based)
  - In quiet, with masks, alarms, and other relevant noises
- ABC-MRT values are estimates
  - Provide preliminary intelligibility values for large number of scenarios at relatively low cost and high speed
  - Follow up a smaller number of promising results with actual MRT to get precise values
- PSCR is doing this for the stakeholders

PSCR

Looking Forward - Wideband Speech

- MRT and ABC-MRT results to date are for narrowband speech
  (nominal bandwidth to about 4 kHz, 17 AI bands)
- Need to test and extend ABC-MRT for wideband use
  (nominal bandwidth to about 7 kHz, 20 AI bands)
- Motivation: LTE can support wideband speech

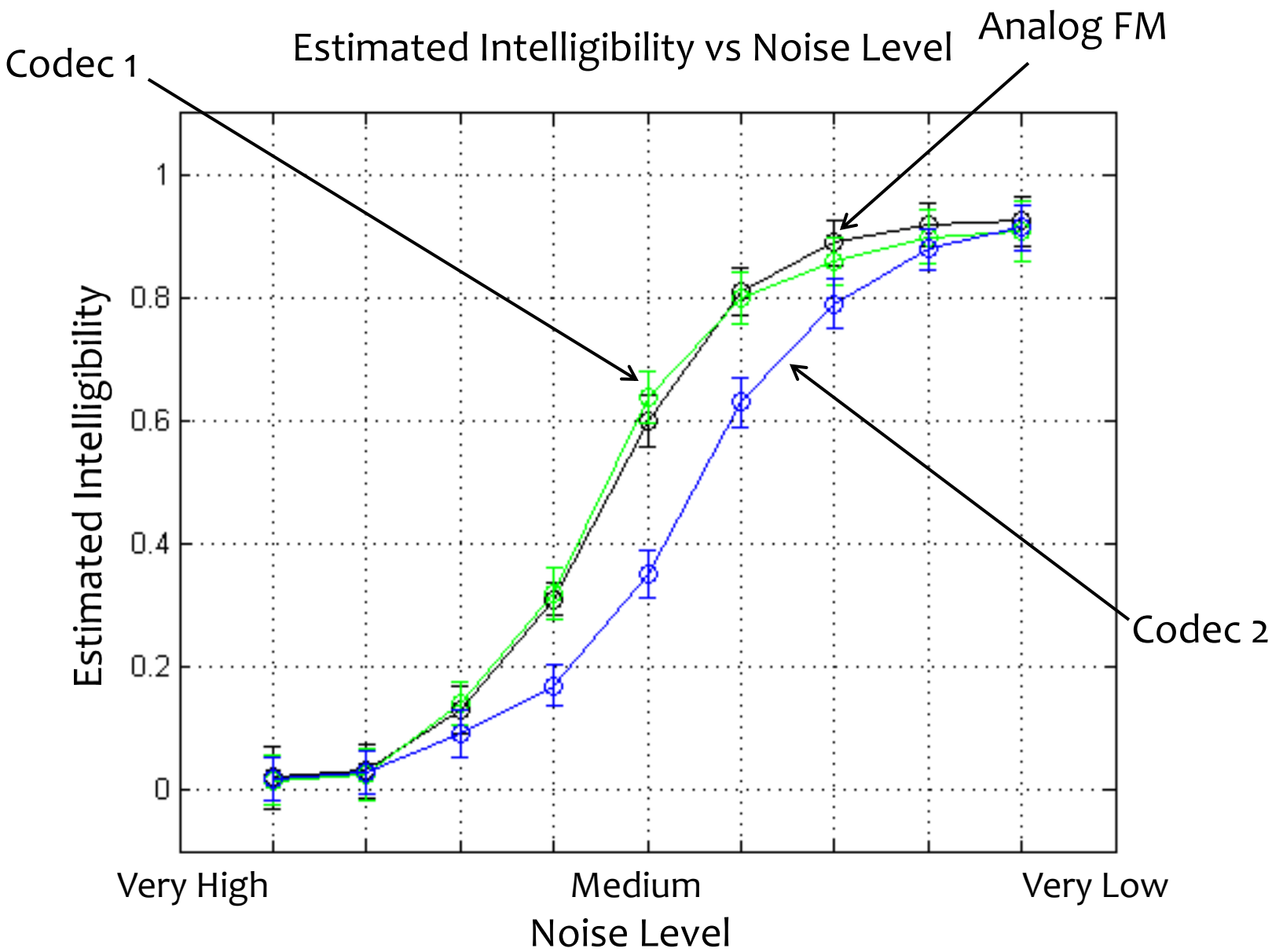
PSCR

Looking Forward - Mission Critical Voice Applications for LTE

- Original digital radio systems required low data rate for digital speech (2.4 – 4.8 kbps) → high levels of speech compression
- LTE can support higher rates → less speech compression is required
- Codecs for Commercial VoLTE: AMR, AMR-WB, EVS
- Codecs for Mission-Critical Voice applications over LTE?
  - One key driver: intelligibility in high noise environments
  - One key reference point: intelligibility of analog FM LMR
- PSCR has plan in place to evaluate codecs using
  - first ABC-MRT to get estimates and narrow the search space
  - then MRT to get precise values

PSCR

Hypothetical Example Result



PSCR

In Summary

- PSCR has
  - MRT experience
  - Awareness of the issues surrounding digital speech coding in noisy conditions
  - Developed ABC-MRT tools
  - Plan to apply ABC-MRT and MRT to investigate codec options for Mission-Critical Voice applications over LTE

pscr.gov

Stephen Voran
svoran@its.bldrdoc.gov

Andrew Catellier
acatellier@its.bldrdoc.gov

PSCR