Walter Reed National Military Medical Center

Acoustic and Non-Acoustic Factors Influencing Speech Intelligibility in Real-World Environments

Douglas Brungart, PhD

Walter Reed National Military Medical Center, Bethesda, MD, USA

Walter Reed National Military Medical Center

Disclaimer

The views expressed in this talk are those of the authors and do not necessarily reflect the official policy or position of the Department of the Army, the Department of the Navy, the Department of the Air Force, the Department of Defense, nor the U.S. Government.

I have no actual or apparent conflicts to disclose

Walter Reed National Military Medical Center

Acknowledgements

UNIVERSITY OF MARYLAND 18 56

Collaborators:

University of Maryland (Sandra Gordon-Salant)
Jocelyn Songer, Ben Sachtler, GN Resound

Mikayla Abrams

Mary Barrett

Kerrianne Clark

Calli Fodor

Julie Cohen

Walter Reed National Military Medical Center

The challenge of assessing real-world speech communication

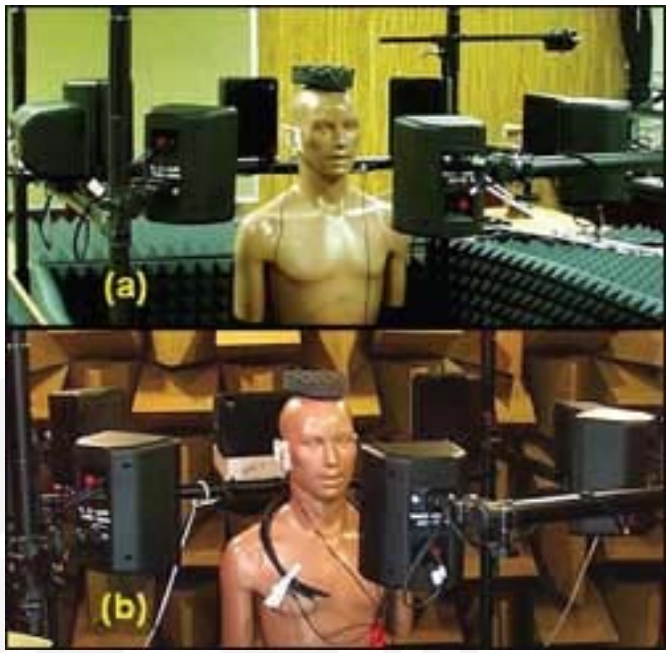
One of the most common problems for hearing impaired listeners is speech communication in noisy public spaces…..

These are also the spaces where hearing aids often fail to deliver the benefits predicted in laboratory or clinical settings

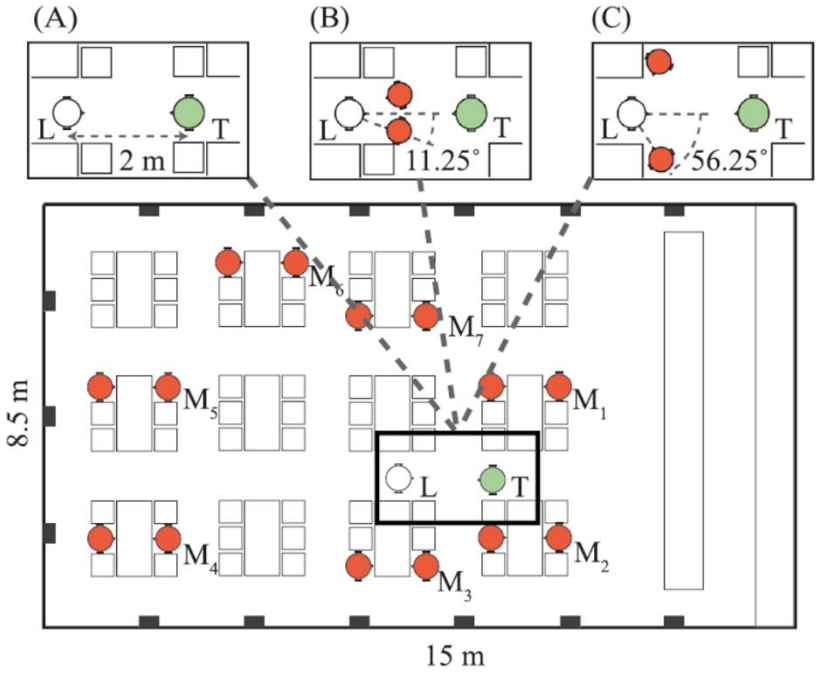
Walter Reed National Military Medical Center

The challenge of assessing real-world speech communication

As researchers, we often respond to this problem by trying to create increasingly realistic auditory environments in lab



R-SPACE, Revitt & Killion, 2007



Westerman and Bucholz, 2017

Walter Reed National Military Medical Center

The challenge of assessing real-world speech communication

However, it is our contention that even the most realistic simulations will fail to capture important aspects of real-world communication

In particular, we are interested in the psychological factors involved in communication in *public* spaces…

Walter Reed National Military Medical Center

The challenge of assessing real-world speech communication

Fundamental to the issue is how SNR is selected in real-world

In the laboratory, we treat SNR as an *arbitrary independent* variable

- **Sometimes we measure performance *as a function* of SNR**
- **Sometimes we adaptively adjust SNR to obtain an SRT**

However, in the real world, SNR is intentionally selected by the talker

Why doesn't the talker just talk loud enough to be understood?

Walter Reed National Military Medical Center

SNR Selection in the Real-World

The answer is that talker speech level is constrained…

Talker expectations and social norms play a powerful role

Consider the case of a NH talker and a HI listener in two situations

Talker will raise voice and talk at 70+ dB,
… but SNR for HI listener is insufficient

a. In a noisy restaurant

b. In a quiet living room

Walter Reed National Military Medical Center

SNR Selection in the Real-World

The answer is that talker speech level is constrained…

Talker expectations and social norms play a powerful role

Consider the case of a NH talker and a HI listener in two situations

Talker will be VERY annoyed if they have to talk at 70+ dB

a. In a noisy restaurant

b. In a quiet conference room

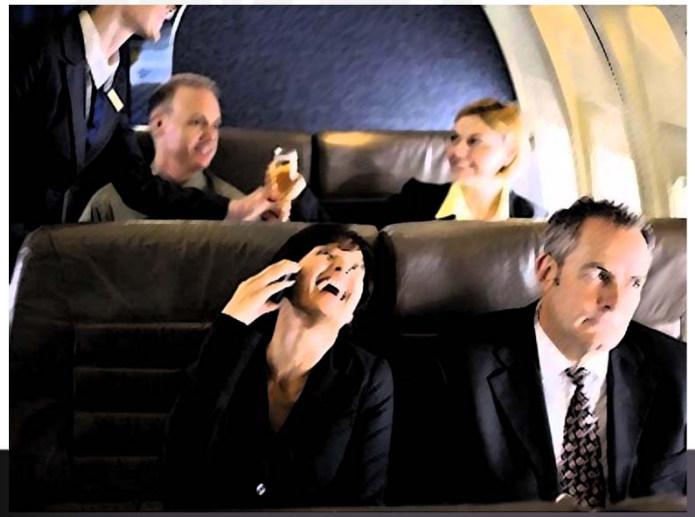
Walter Reed National Military Medical Center

Examples of inappropriate SNR Selection in Real-world

In most cases, talkers choose an appropriate level for environment

However, when that doesn't happen, annoyance and frustration ensues

Talker on Cell Phone



Talker Wearing Hearing Protection

Walter Reed National Military Medical Center

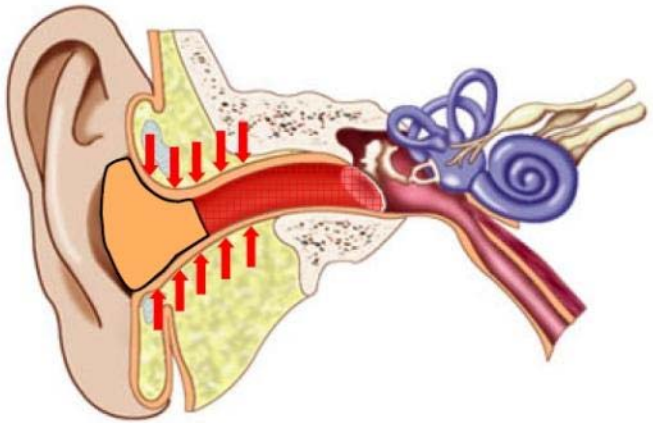
Impact of Occlusion in HPDs

Non-linear HPDs address Detection and Identification… but Face-to-Face Communication may be negatively impacted by occlusion effect

Occlusion is the result of bone-conducted self-generated sound that is trapped in the ear canal, primarily by a shallow, unvented earplug

Occlusion causes your own voice to sound amplified and “boomy”

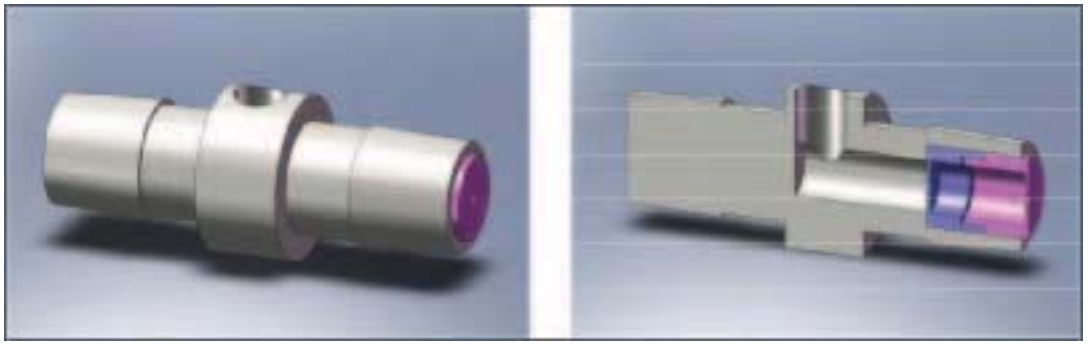
In conjunction with HPD attenuation, which makes room sound quieter, occlusion can cause speakers to talk too quietly in noise to be understood



Combat Arms Earplug
Non-linear Passive Protector

Combat Arms Earplug has small non-linear filter designed by ISL

- **React to gradient of sound pressure wave**
- **Attenuates only very short, very loud impulsive sounds**
- **Provides very little attenuation in continuous noise**

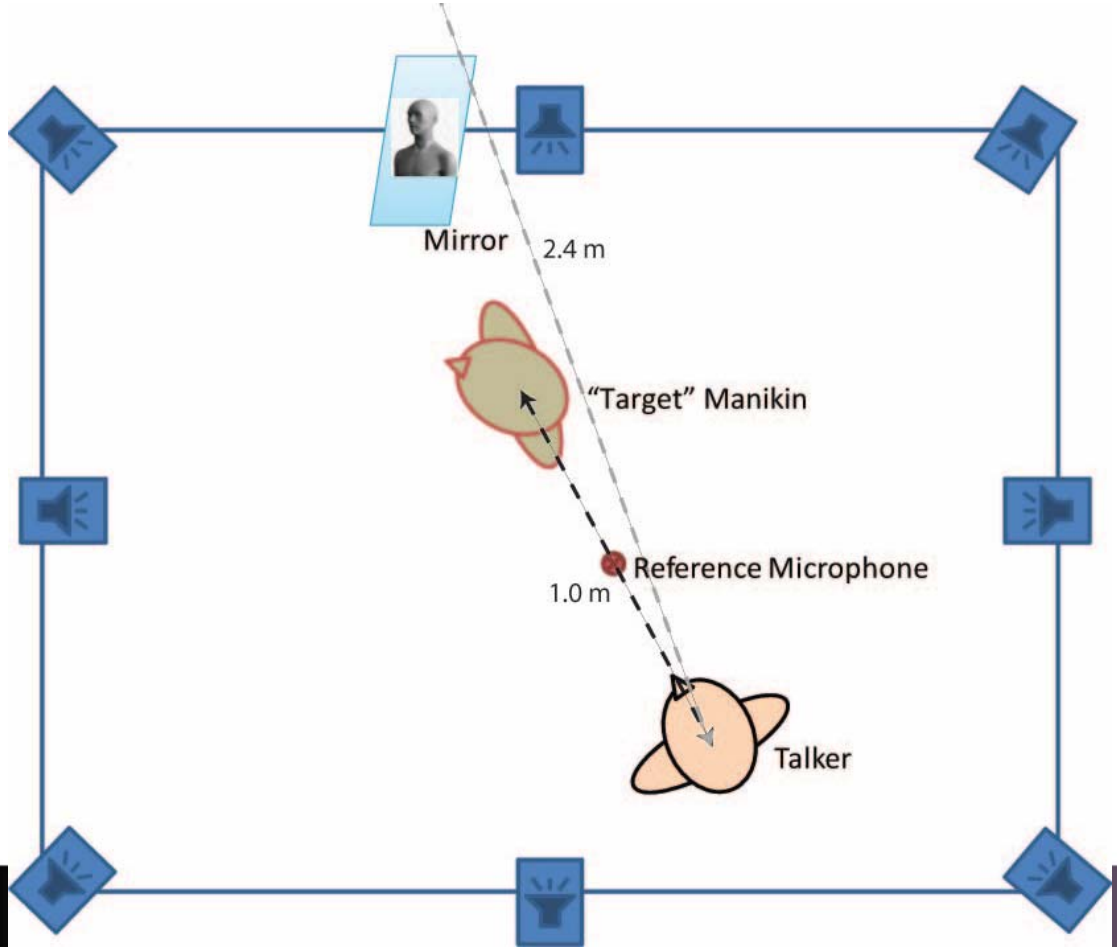


DiSpirito and Binseel, 2008

Walter Reed National Military Medical Center

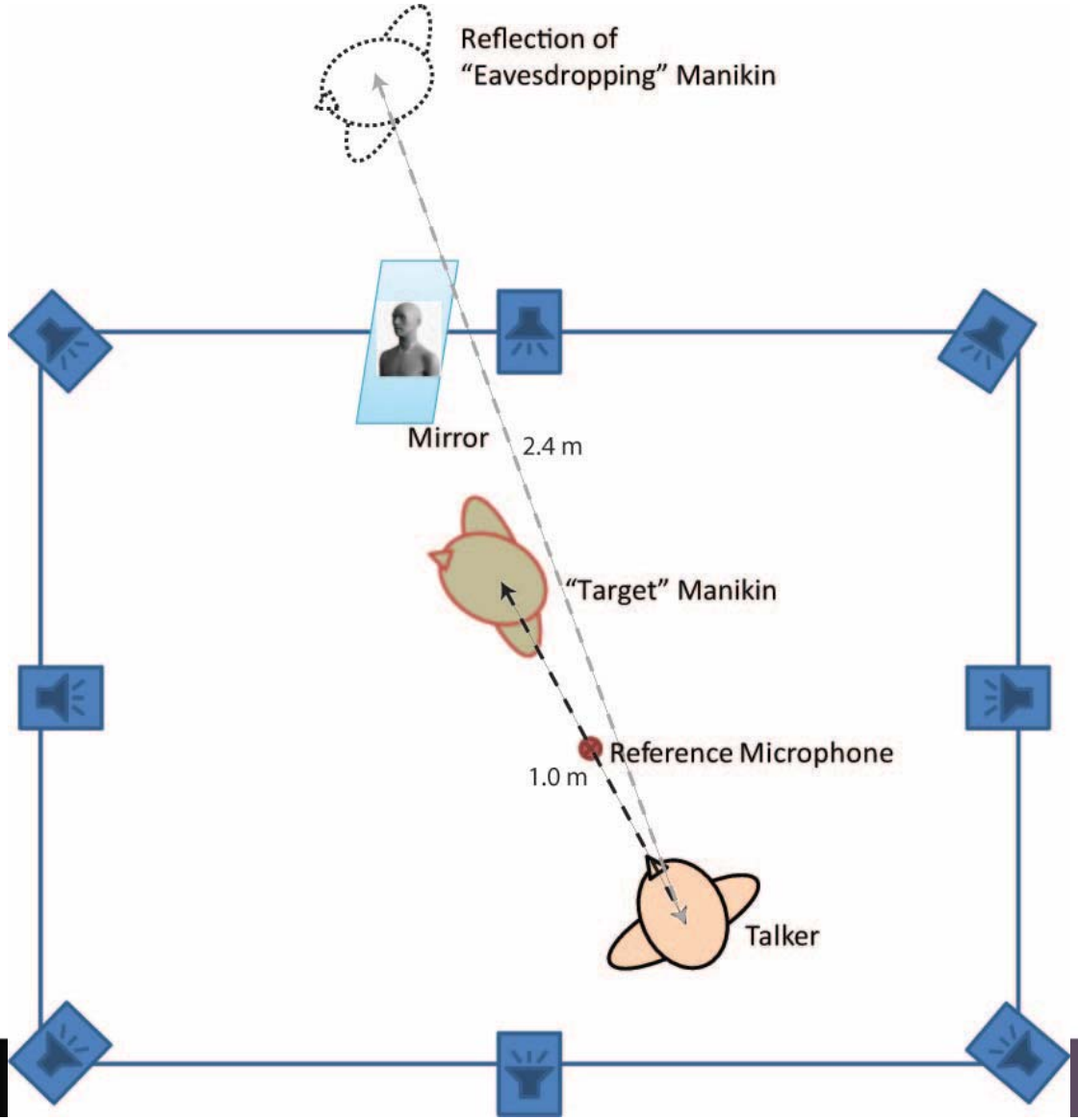
Experiment Examining Occlusion Effect in Nonlinear HPDs

Talkers were placed in a sound booth a with 16-speaker array, 1.0 m from acoustic manikin



Walter Reed National Military Medical Center

Experiment Examining Occlusion Effect in Nonlinear HPDs

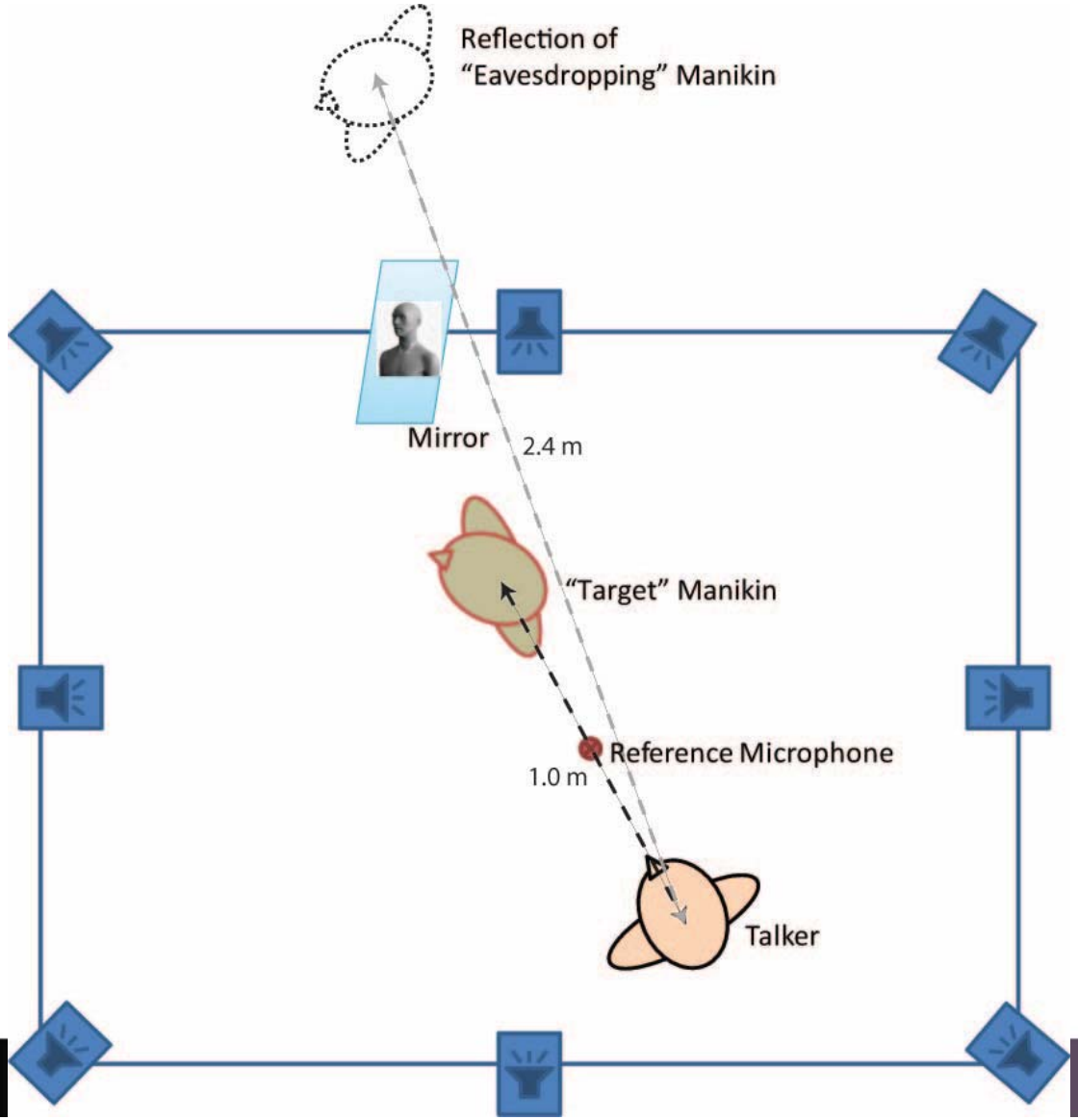


Talkers were placed in a sound booth a with 16-speaker array, 1.0 m from acoustic manikin

Mirror on wall reflected image of manikin at distance of 2.4 m

Walter Reed National Military Medical Center

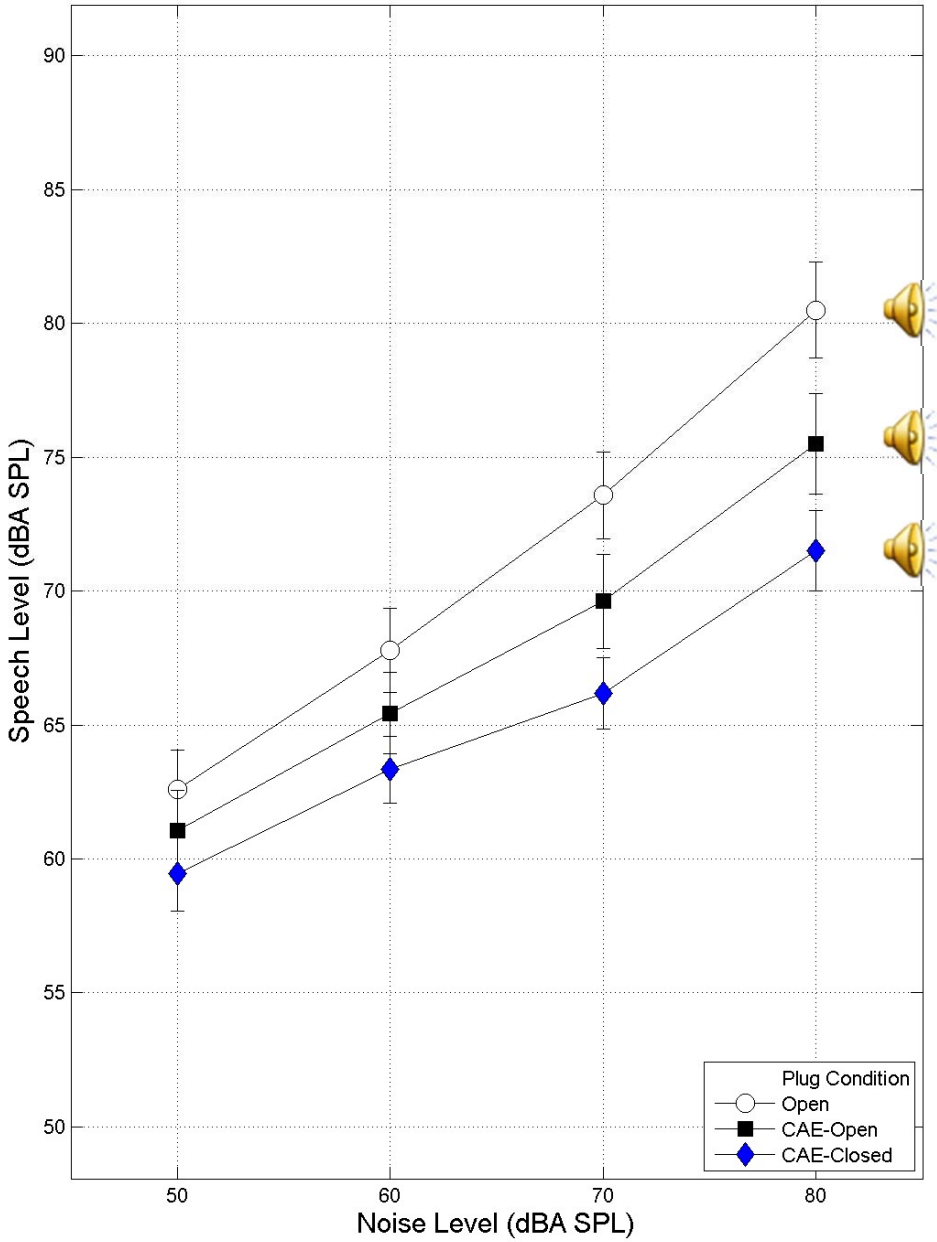
Experiment Examining Occlusion Effect in Nonlinear HPDs



Talkers were placed in a sound booth a with 16-speaker array, 1.0 m from acoustic manikin

Mirror on wall reflected image of manikin at distance of 2.4 m

Talkers asked to speak phrases loudly enough for close manikin, but too softly for far manikin to hear



CAE Earplug reduced voice level by 5.8 dB in CAE Open Condition

And by an additional 4.3 dB in CAE Closed Condition

Plug Condition, Noise Condition, and Interaction all significant at $p<0.001$ level (2-factor within-subject ANOVA)

This could change the talker from completely intelligible to completely unintelligible at manikin location

Walter Reed National Military Medical Center

Conclusion 1a

Intelligibility in Noise depends on the state of the *Talker*,

not just the
Acoustic Environment
or the
Characteristics of the Listener

Walter Reed National Military Medical Center

Question: Would this Happen in the Real World

But wait, you say…

This would never happen in the real world...

People have non-verbal cues to indicate when somebody hears them.

How can we assess intelligibility “in the wild”

Walter Reed National Military Medical Center

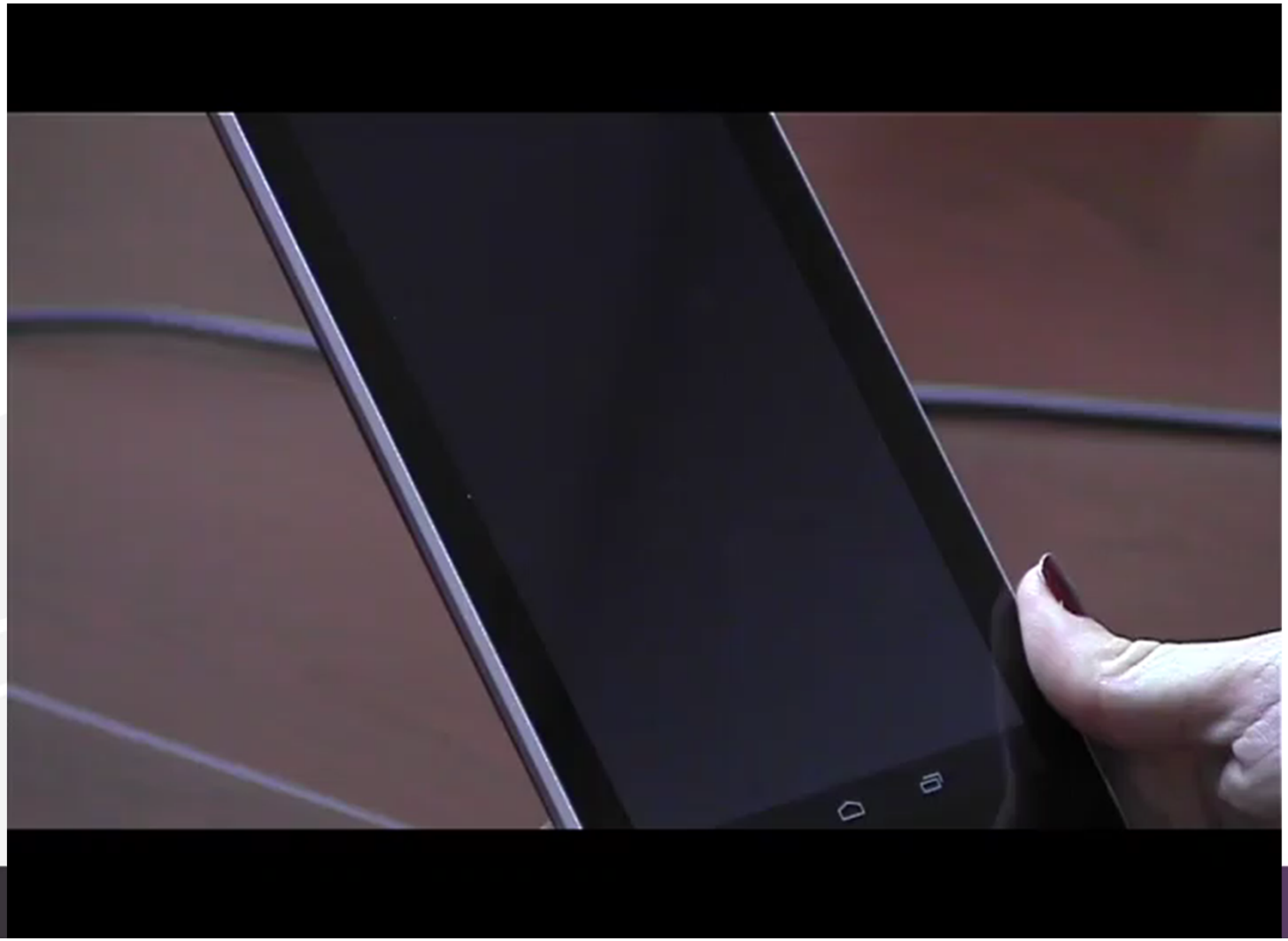
A new model for assessment
"The cafeteria study"

We are now working on a new approach of "in situ" assessment

Walter Reed National Military Medical Center

A new model for assessment
"The cafeteria study"

On each trial, one talker is randomly assigned as Talker



Walter Reed National Military Medical Center

A new model for assessment
"The cafeteria study"

The other three participants have to select from six options

Walter Reed National Military Medical Center

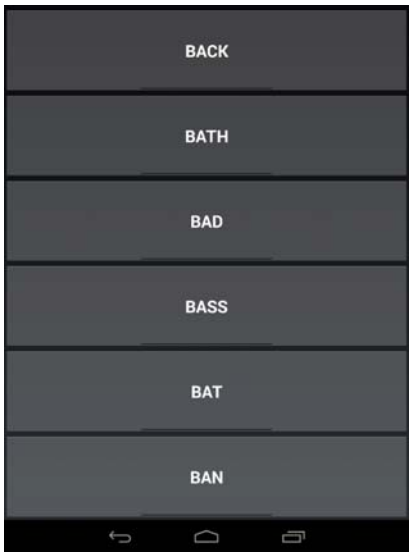
A new model for assessment
"The cafeteria study"

The entire experiment can be moved to cafeteria for data collection

Preliminary results show normal listeners maintain intelligibility in noise, but slow down and start to use visual cues when noise exceeds 80 dBA

Walter Reed National Military Medical Center

Cafeteria Test Paradigm



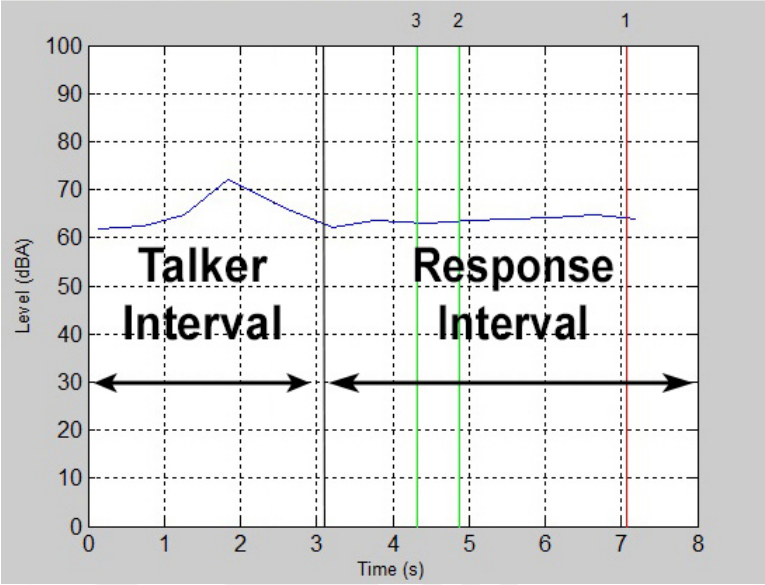
Walter Reed National Military Medical Center

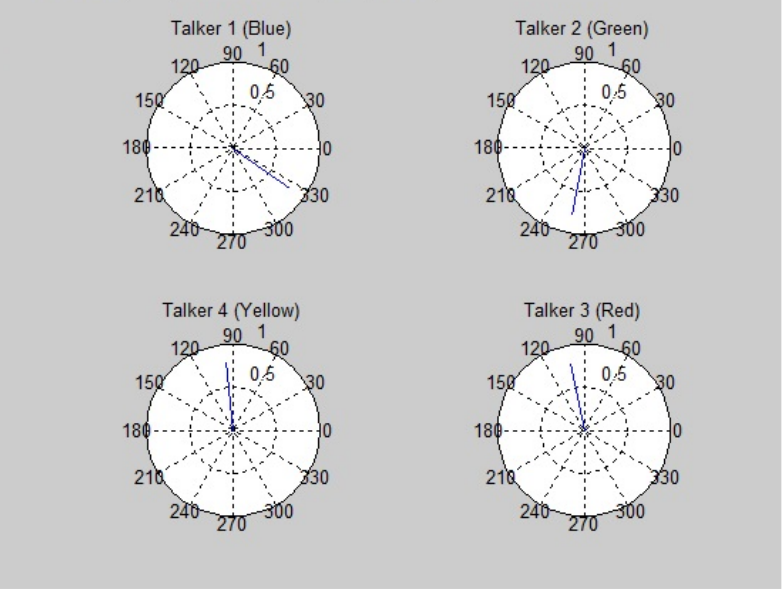
Ratings

- There will be times where you need to rate how easy or difficult it was to hear/understand each talker.
- The higher the number, the more difficult it was to hear/understand the talker.
- You will also rate yourself: How difficult was it for others to understand you?

Walter Reed National Military Medical Center

Cafeteria Test Paradigm





Walter Reed
National Military
Medical Center

Earplug Experiment

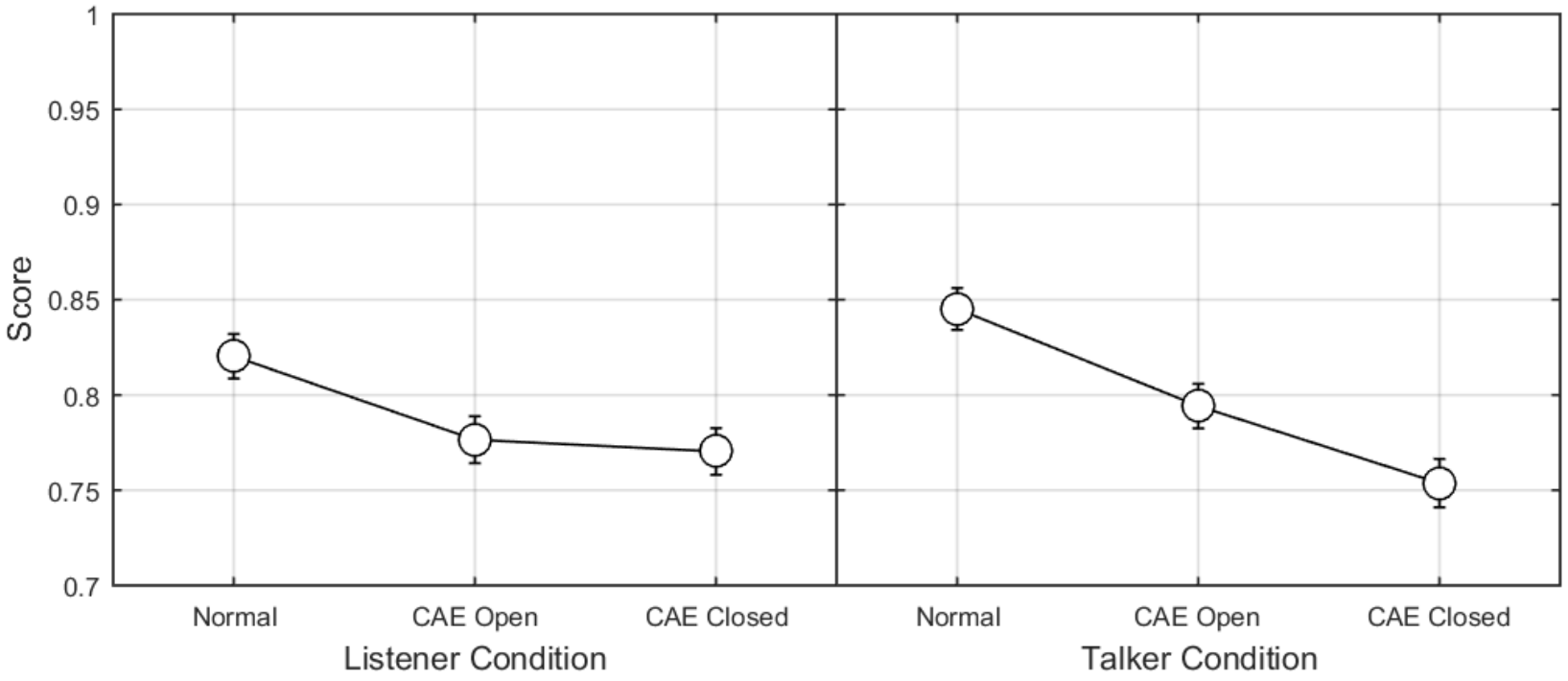
Study Run with

Open Ears
Combat Arms Open
Combat Arms Closed

Walter Reed National Military Medical Center

A new model for assessment
"The cafeteria study"

Talker earplugs matter 2x as much as listener earplugs!

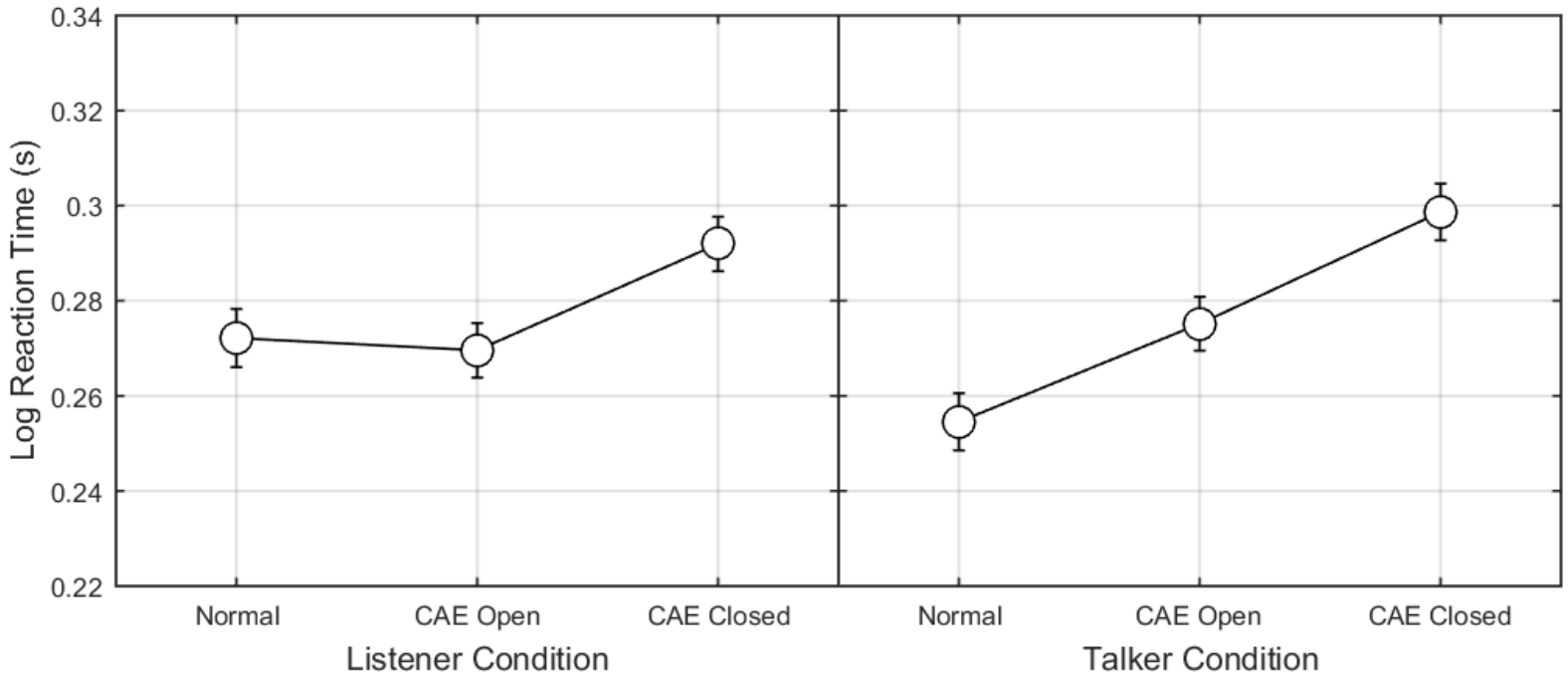


Walter Reed National Military Medical Center

A new model for assessment
"The cafeteria study"

Earplug Experiment: Talker matters more than listener

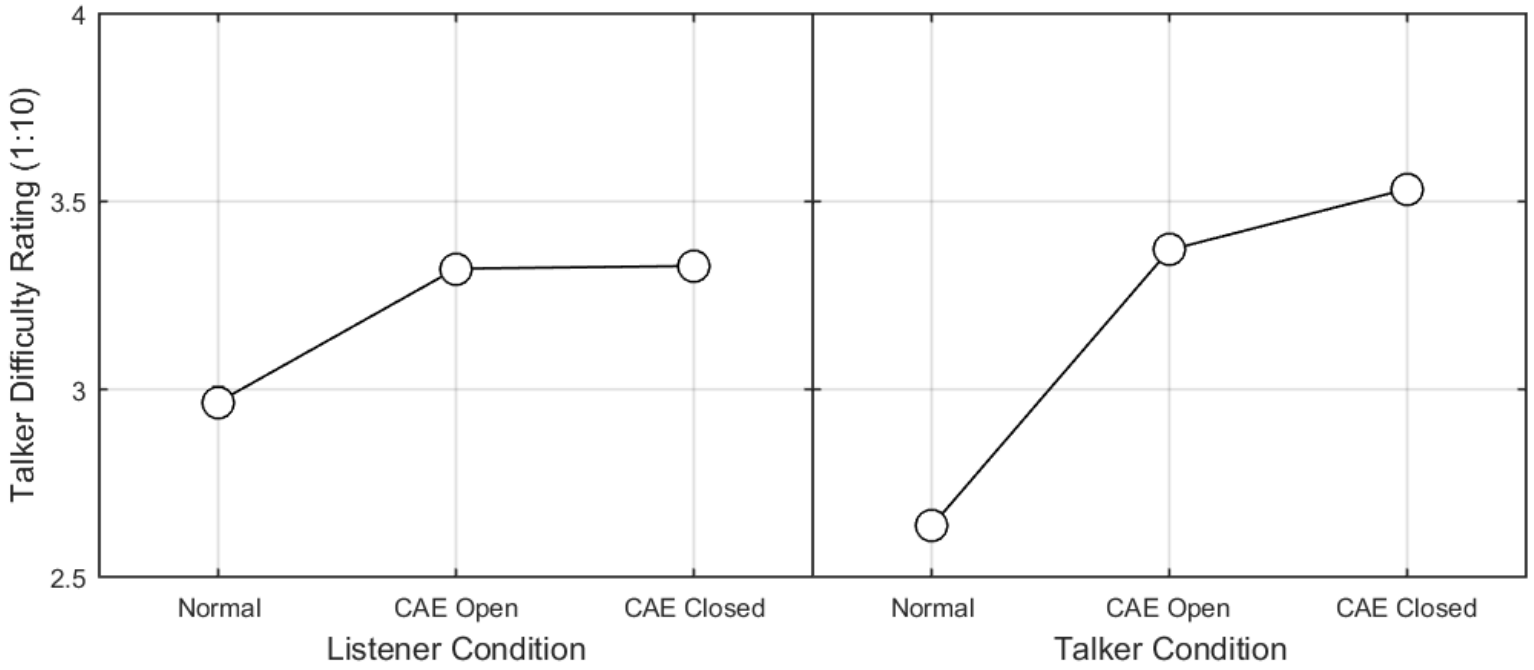
Reaction Time > Score ?



Walter Reed National Military Medical Center

A new model for assessment
"The cafeteria study"

Earplug Experiment: Talker matters more than listener
Reflected in Subjective Rating



Walter Reed National Military Medical Center

Conclusion 1b

Intelligibility in Noise depends

A LOT

on the state of the *Talker*,

not just the

Acoustic Environment

or the

Characteristics of the Listener

Even when talking face-to-face!

Walter Reed
National Military
Medical Center

IMPACT OF DIRECTIONAL HEARING AID PROCESSING ON REAL-WORLD SPEECH PERCEPTION

Walter Reed National Military Medical Center

Evaluate the Effects of Hearing Aids on Real-World Listening Performance

Walter Reed Cafeteria

Normal Hearing (NH): 19 - 60 yrs (n=10); mean 34.3yrs;
Θ ≤ 20 dB HL, 250-4000 Hz

Hearing-Impaired (HI): 53 - 80 yrs (n=26); mean=67.7yrs
Symmetrical sensorineural hearing loss
Experienced bilateral hearing aid users (min 6 months)

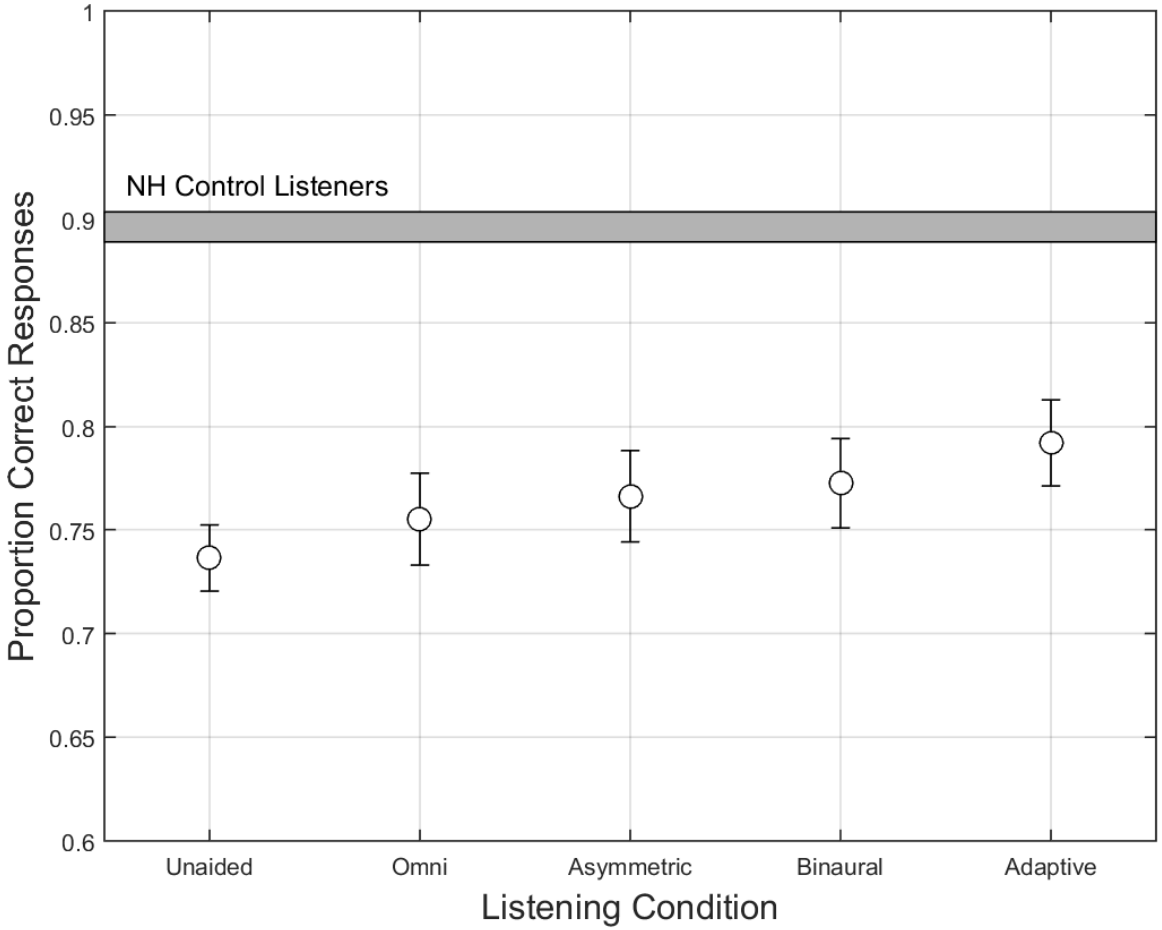
HI participants were fit binaurally with receiver-in-the-canal (RIC) hearing aids

- Omnidirectional
- Asymmetric directionality
- Binaural adaptive directionality
- Adaptive directionality.

Walter Reed National Military Medical Center

Impact of Hearing Aid Directional Processing

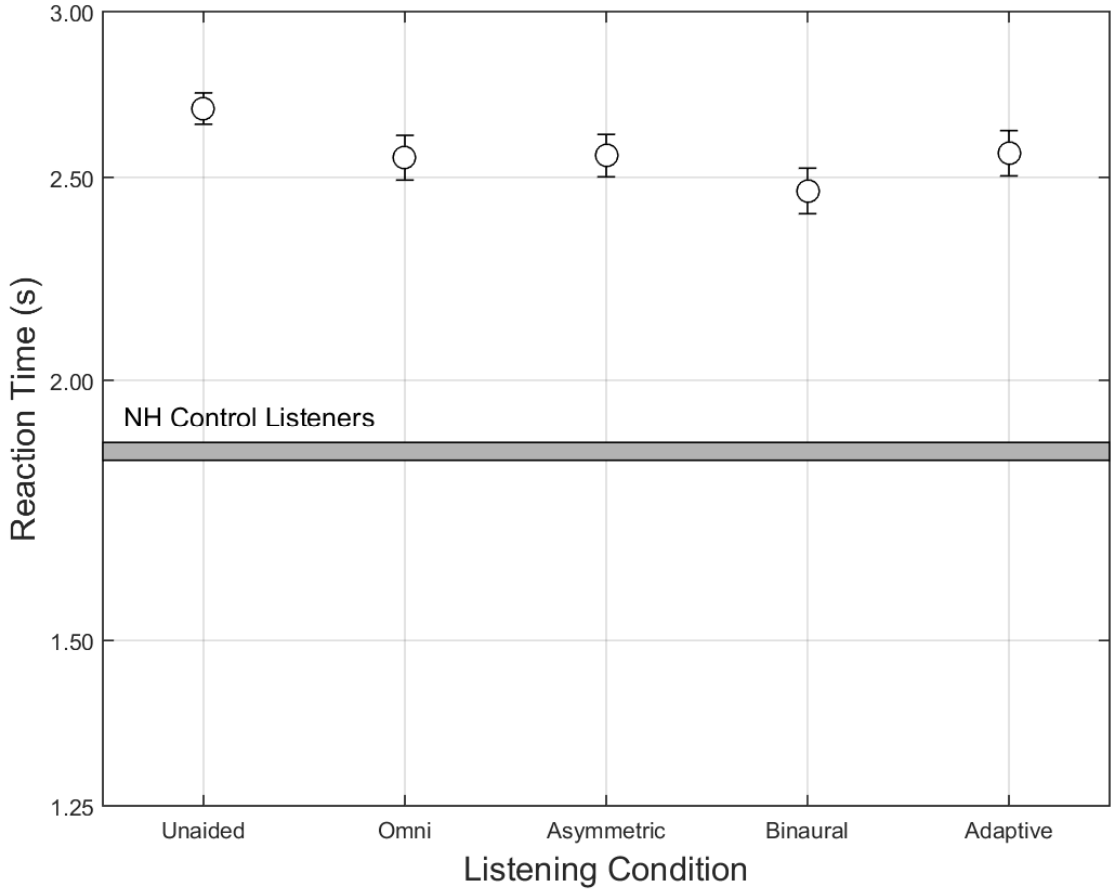
Hearing impaired listeners had more difficulty than young normal, and there was a trend towards better performance with directionality



Walter Reed National Military Medical Center

Impact of Hearing Aid Directional Processing

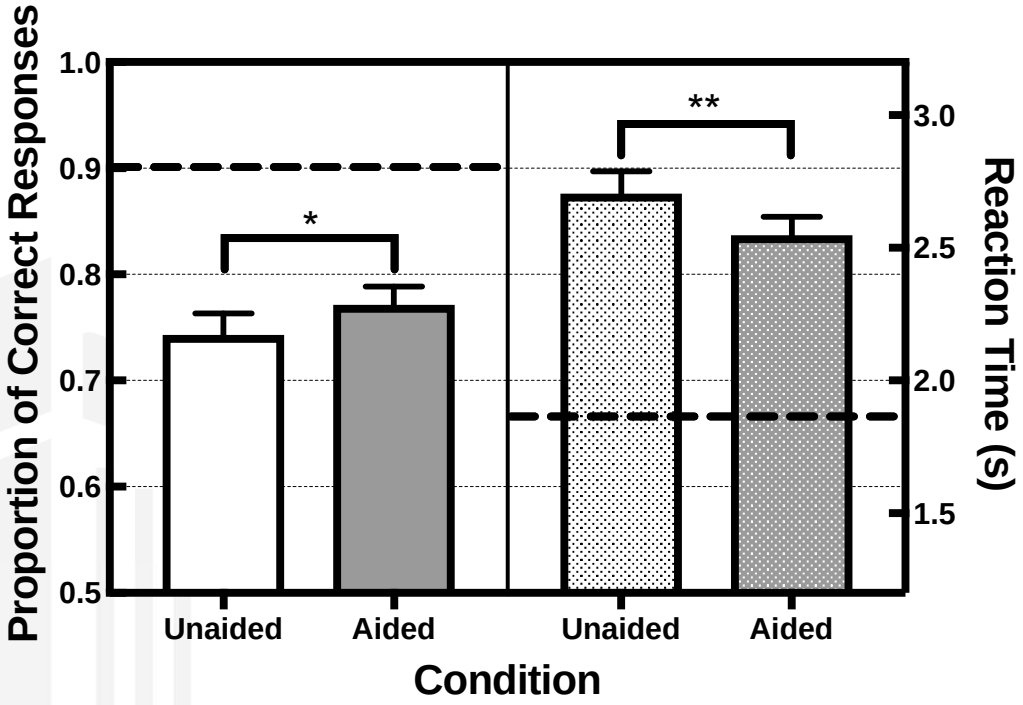
Hearing impaired listeners were also slower than young normals



Walter Reed National Military Medical Center

Impact of Hearing Aid Processing

Overall, there was an effect of aiding, but no difference across directionality

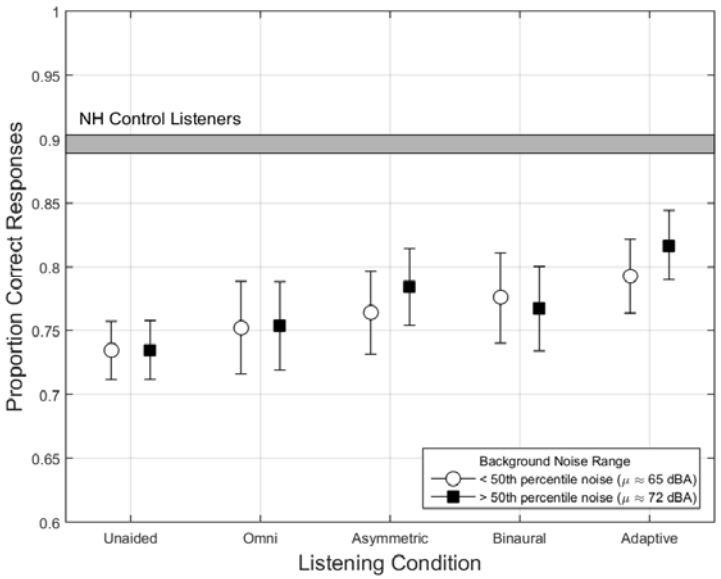


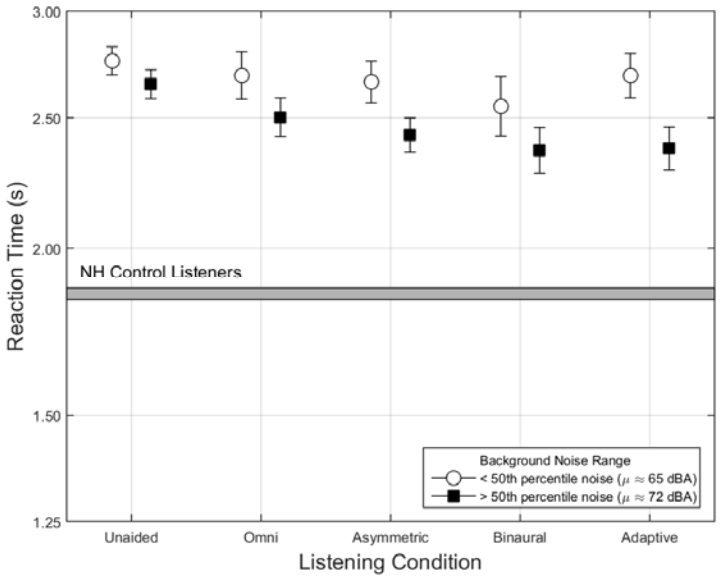
Walter Reed National Military Medical Center

Impact of HA Processing

But the effects of noise were counterintuitive…..

Subjects did slightly better when the noise level was > median





Walter Reed National Military Medical Center

Conclusion 2

In real-world environments, Hearing Impairment

- **Decreases percent correct score**
- **Slows reaction time**
- **Increases perceived effort**

Hearing aids provide a modest improvement in performance

Differences in performance across directional patters are subtle

Walter Reed
National Military
Medical Center

A SURVEY OF LISTENING CONDITIONS IN REAL-WORLD ENVIRONMENTS

Walter Reed National Military Medical Center

Normative Evaluation of Real-World Speech-in-Noise Performance

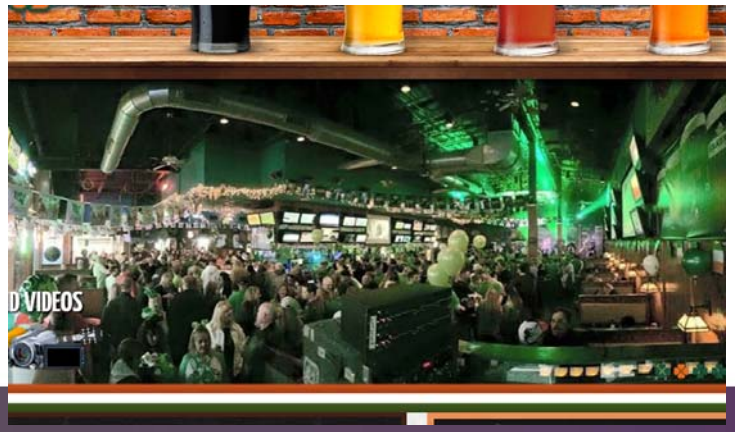
Walter Reed Cafeteria

Restuarant

Library

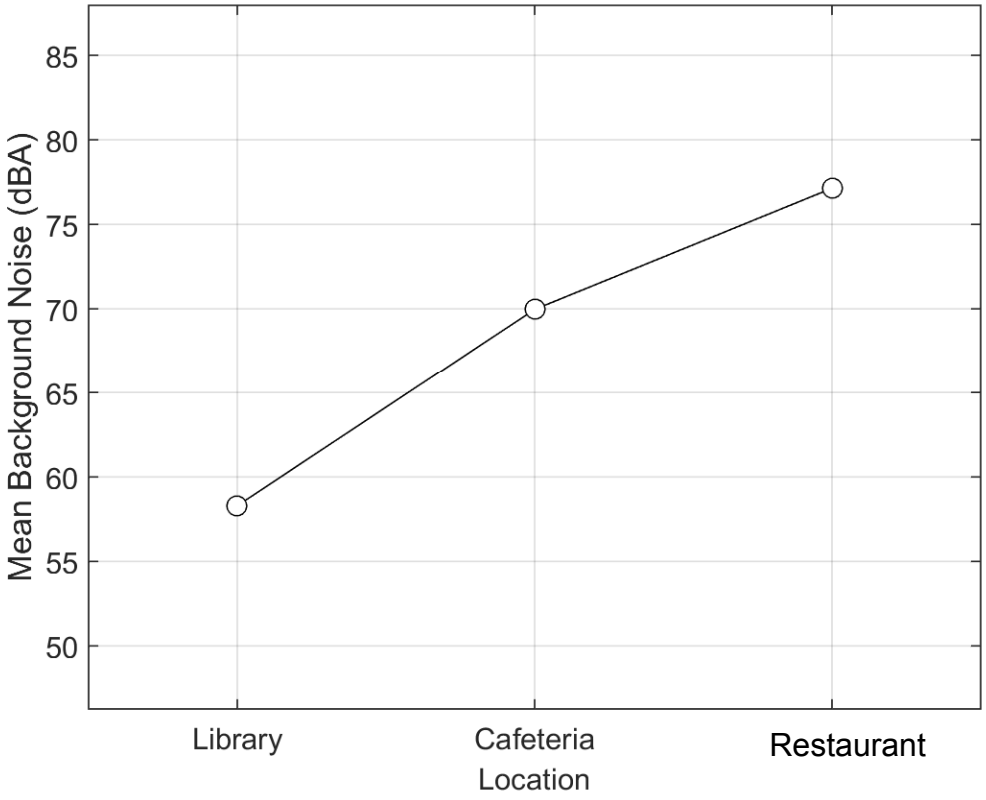
UMD Cafeteria

Pub



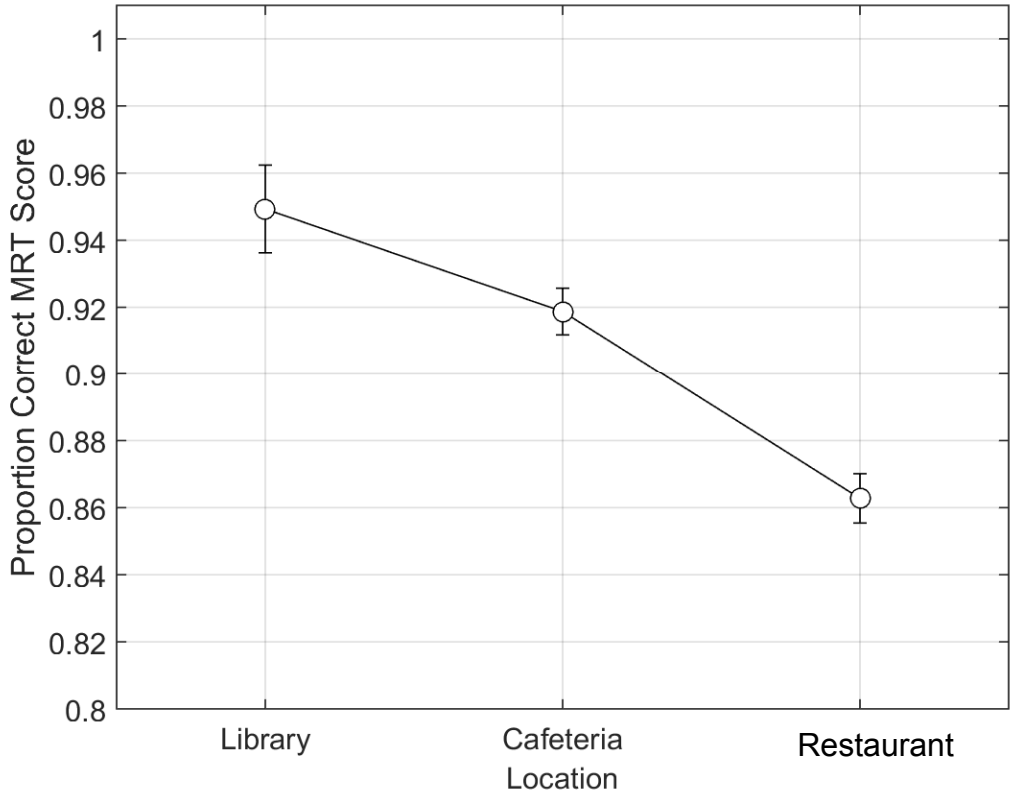
Walter Reed National Military Medical Center

Comparison Across Locations



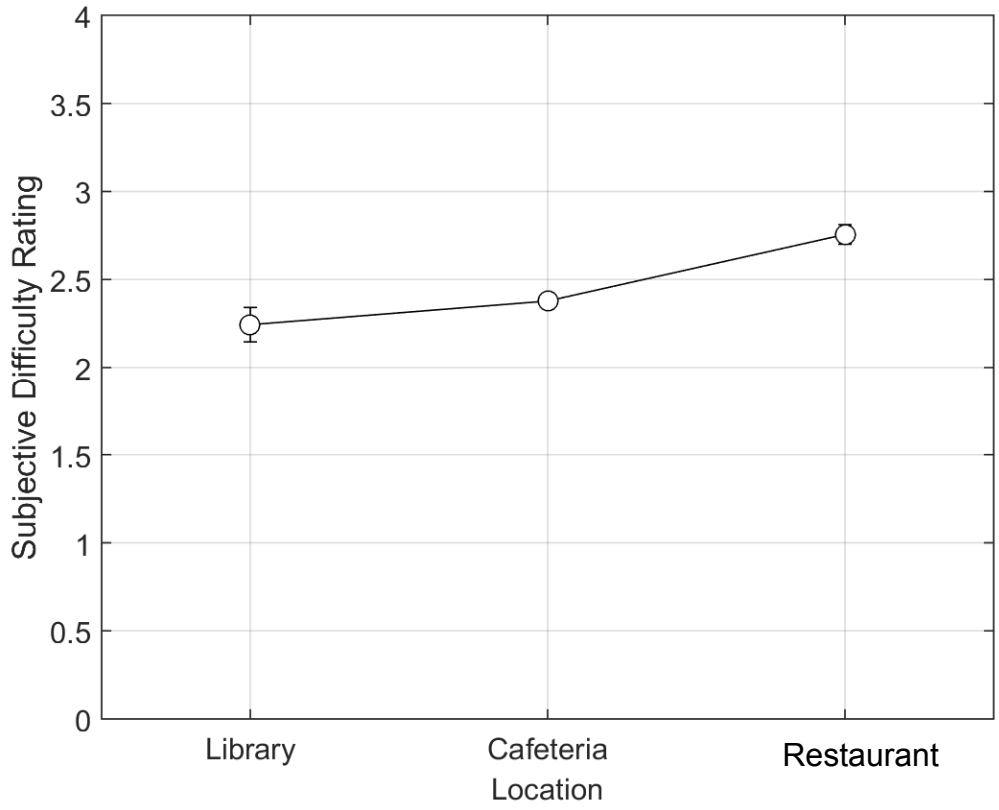
Walter Reed National Military Medical Center

Comparison Across Locations



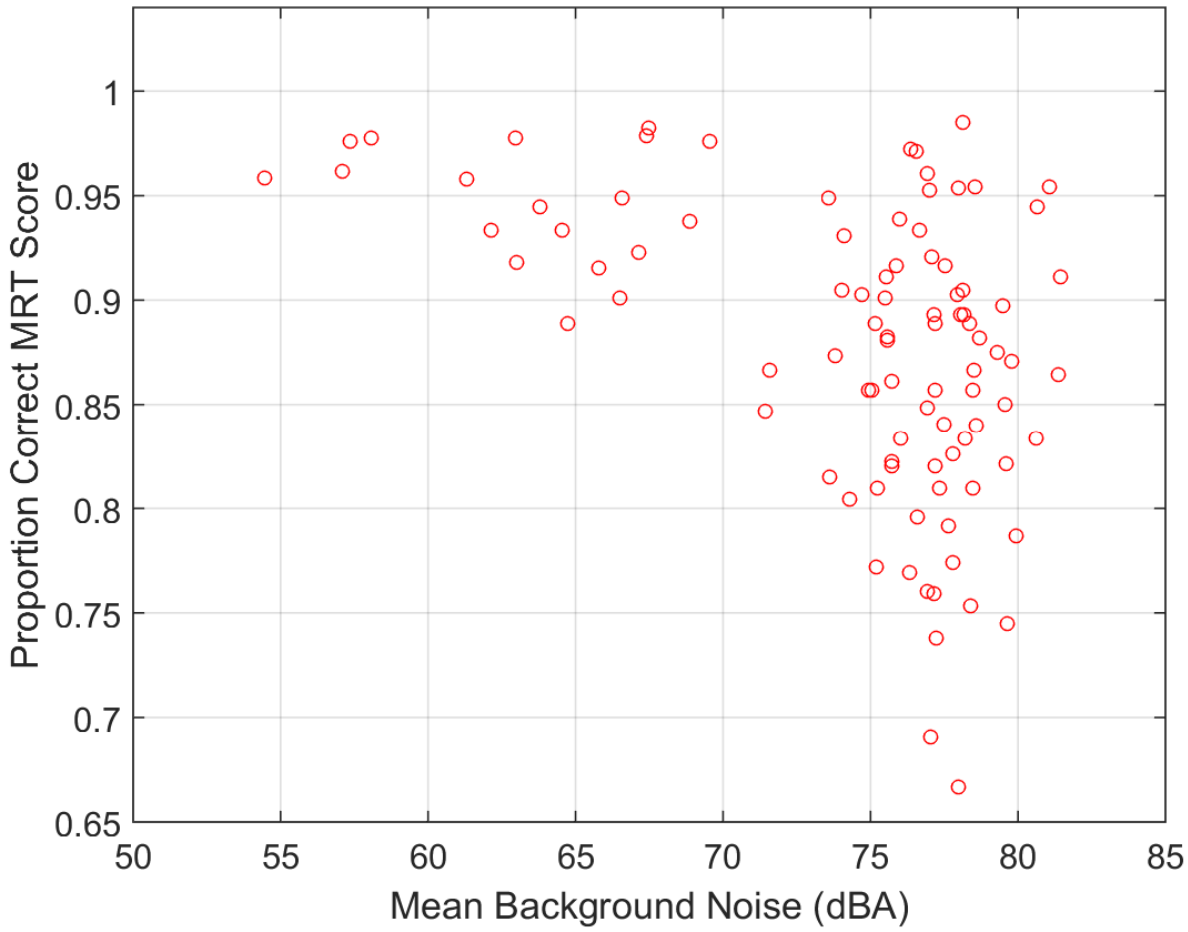
Walter Reed National Military Medical Center

Comparison Across Locations



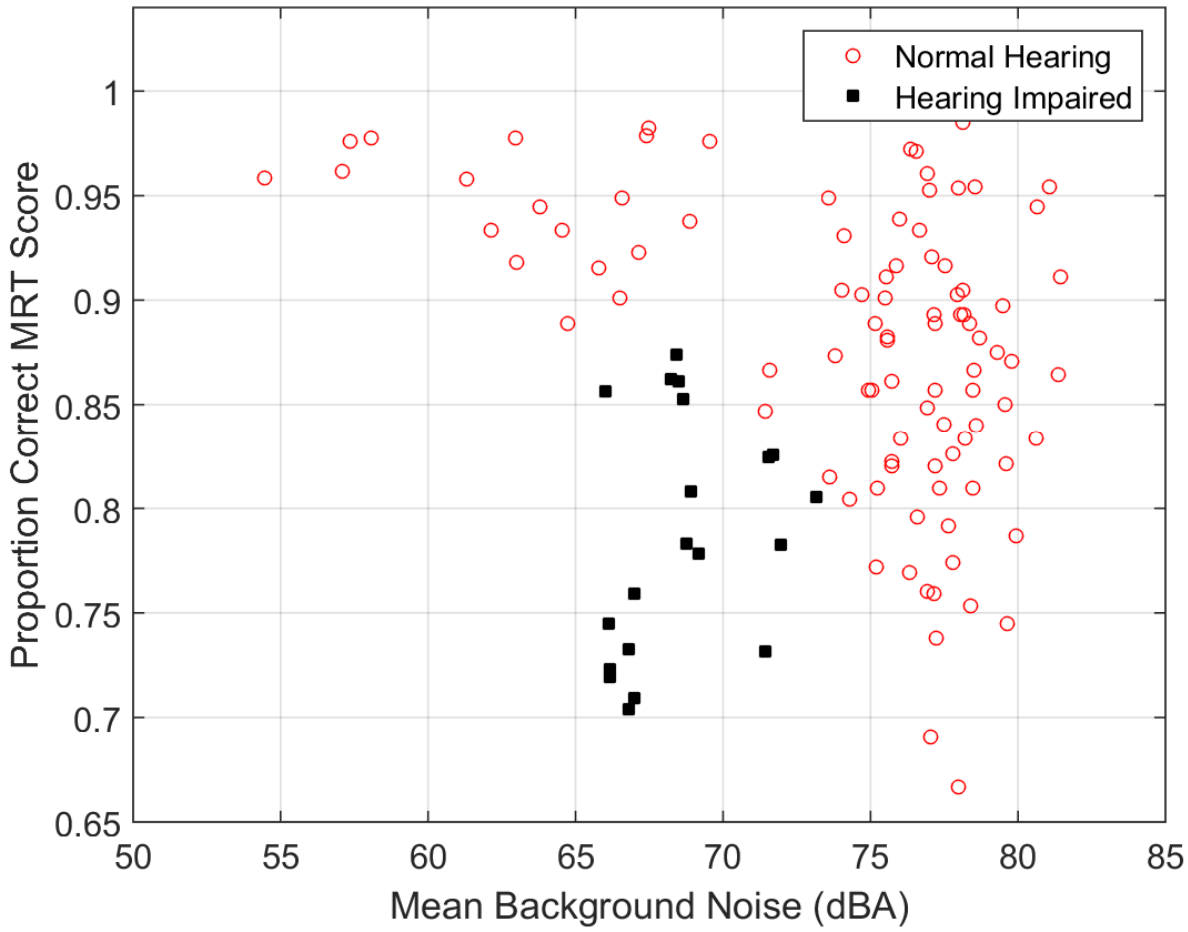
Walter Reed National Military Medical Center

Score vs. Noise Level



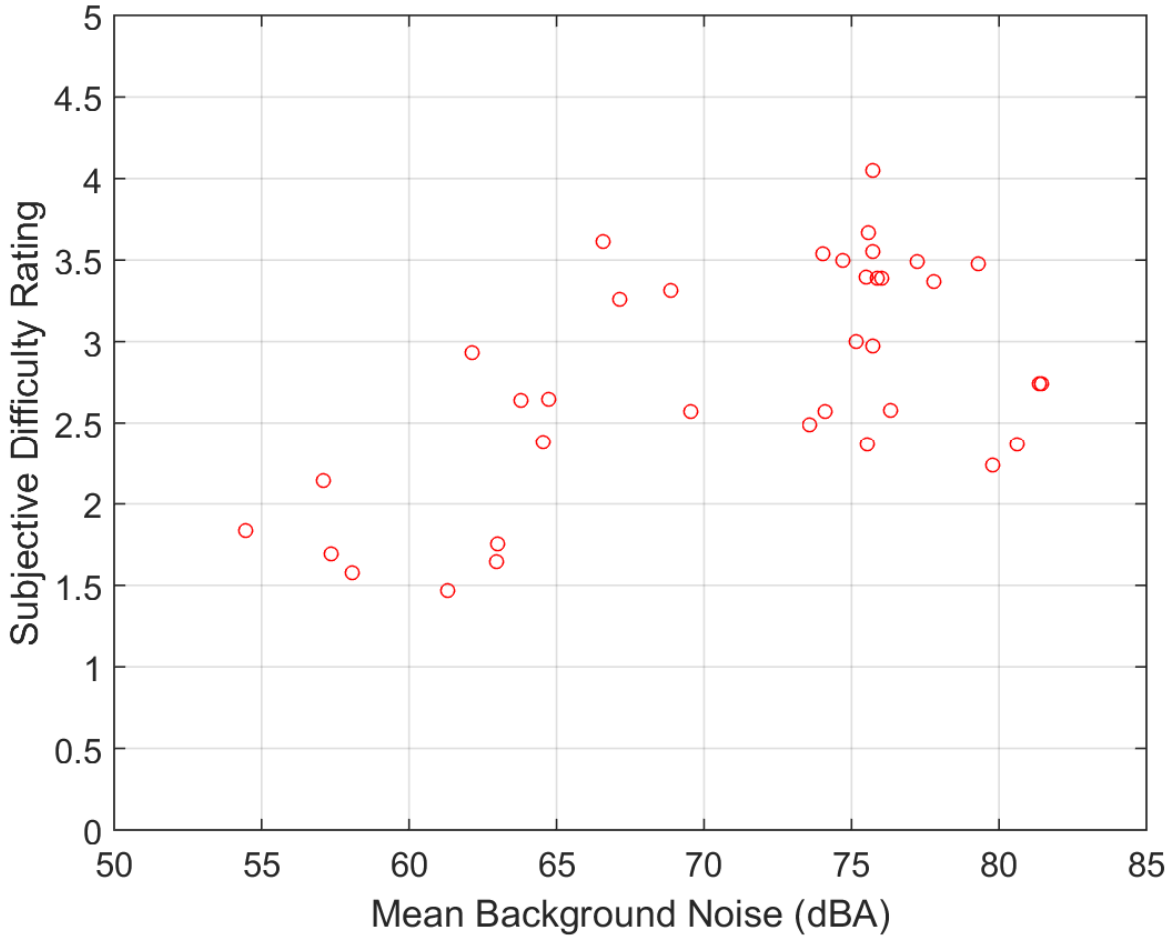
Walter Reed National Military Medical Center

Score vs. Noise Level



Walter Reed National Military Medical Center

Difficulty Rating vs. Noise



Walter Reed National Military Medical Center

Conclusion 3

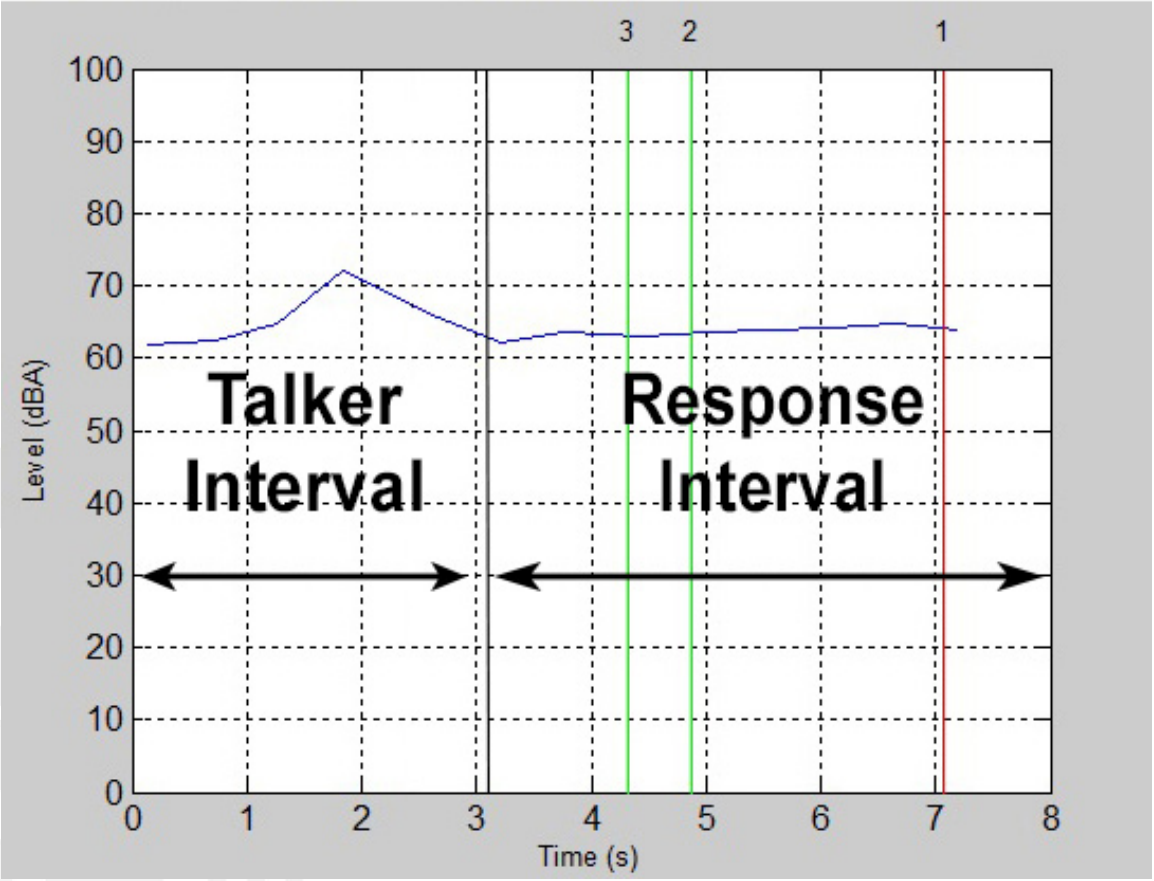
Normal Hearing Listeners are *near ceiling* in real world environments < 70 dB

Performance in variable > 70 dB, but poorer

Hearing impaired listeners are much worse Aided or Unaided

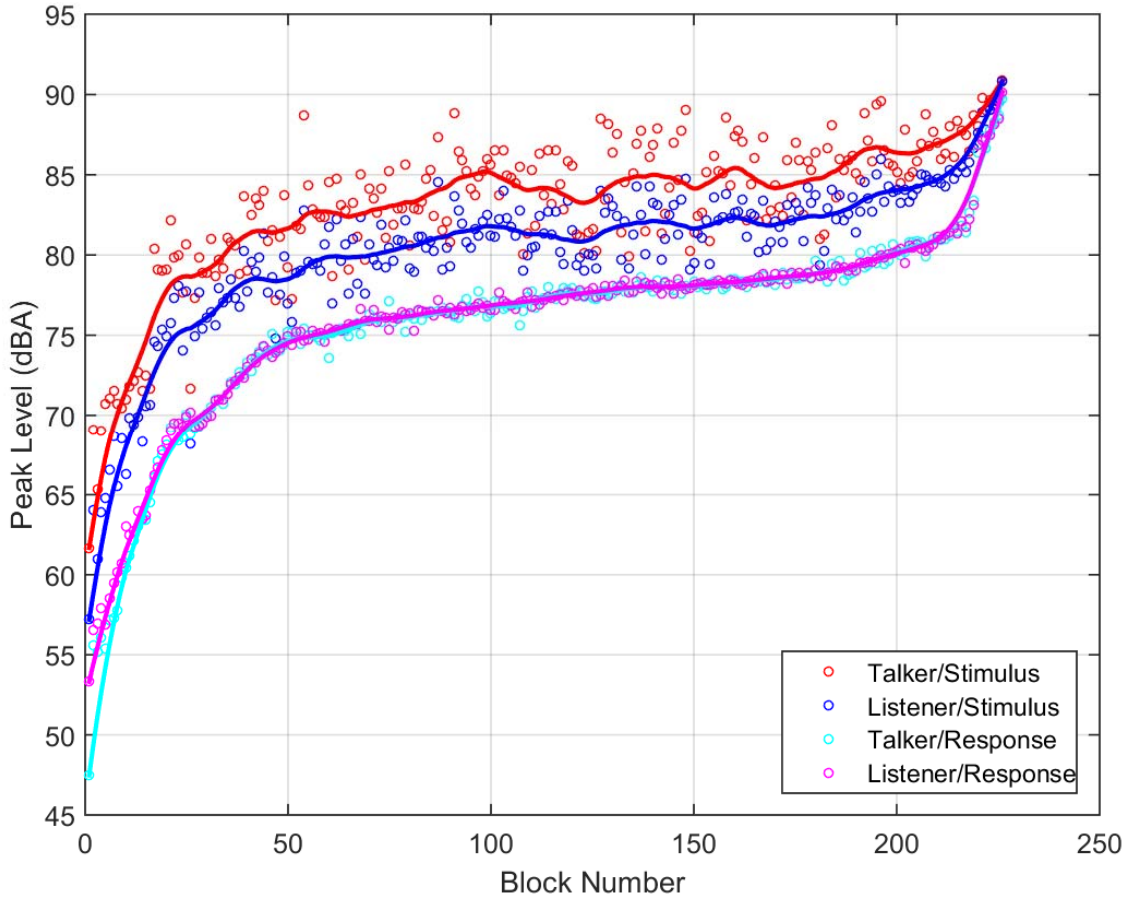
Walter Reed National Military Medical Center

Analysis of SNR



Walter Reed National Military Medical Center

Peak Sound Level



Level at Target Tablet During Target Speech (Target + Masker)

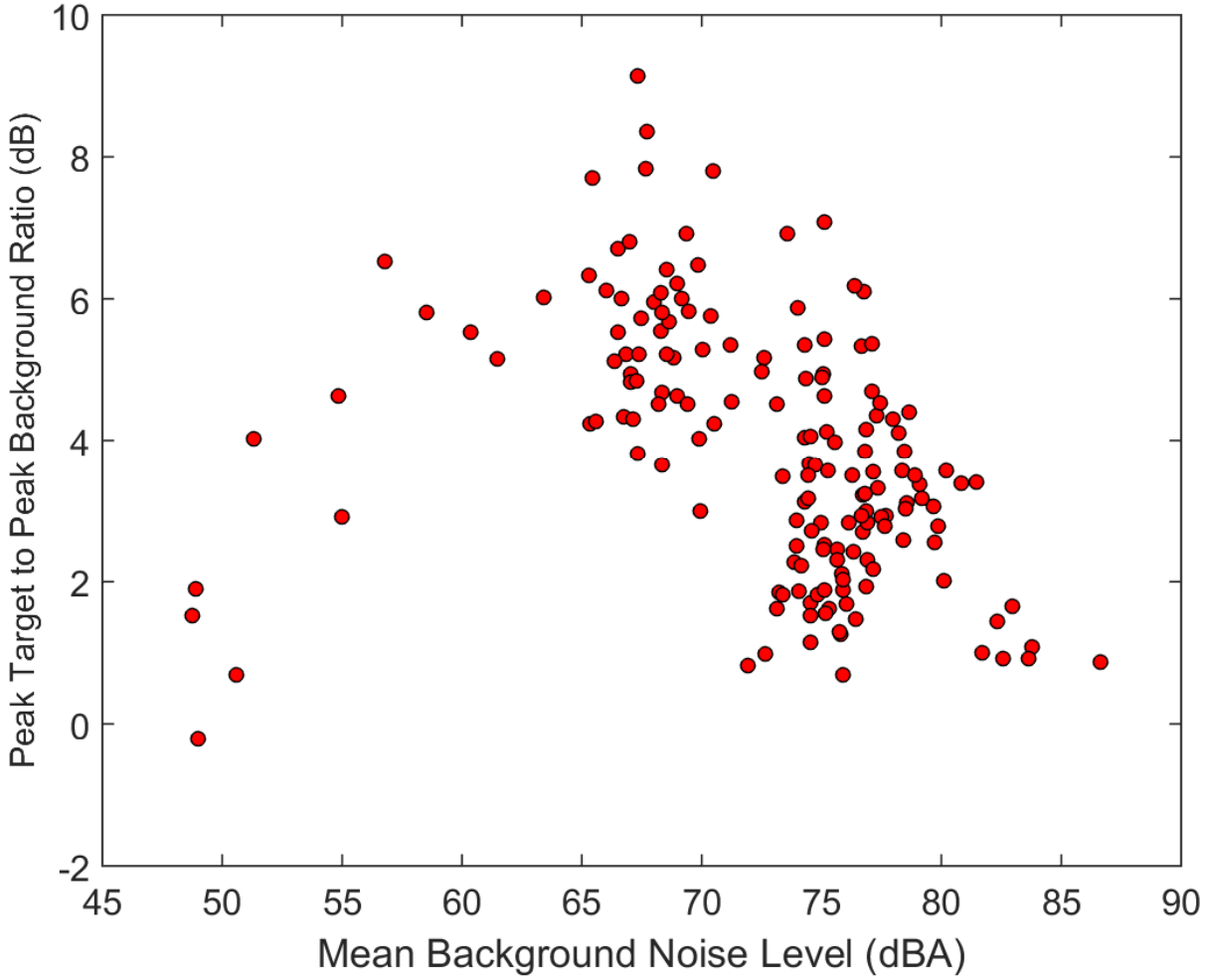
Level at Listener Tablet During Target Speech (Target + Masker)

Level at Target Tablet During Response (Masker Only)

Level at Listener Tablet During Response (Masker Only)

Walter Reed National Military Medical Center

Estimating SNR



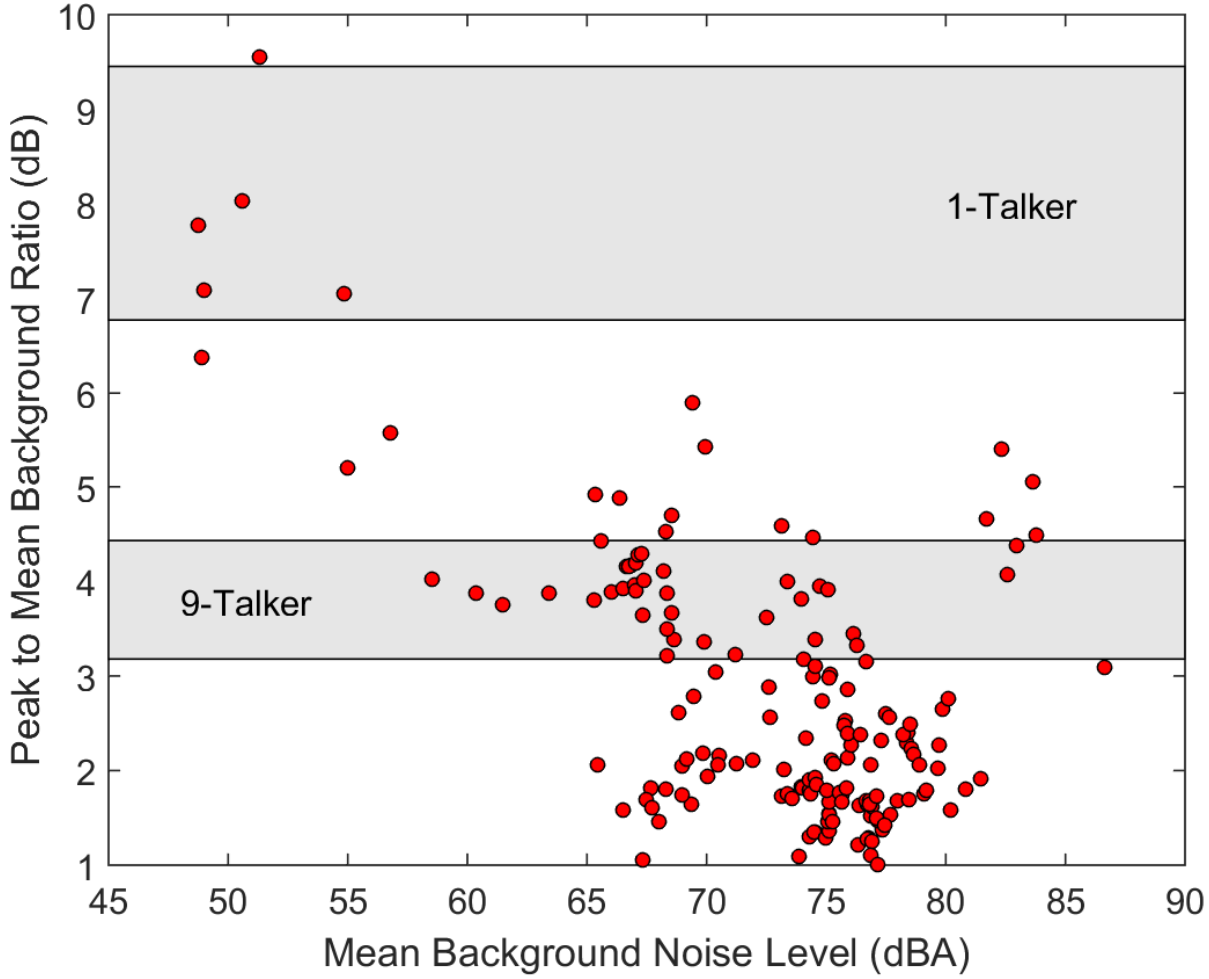
Crude SNR estimate is Peak Target to Mean Response ratio….

i.e.

(Target + Masker)
(Masker Only)

Walter Reed National Military Medical Center

Peak Target to Mean Background SNR Depends on Modulation in Background



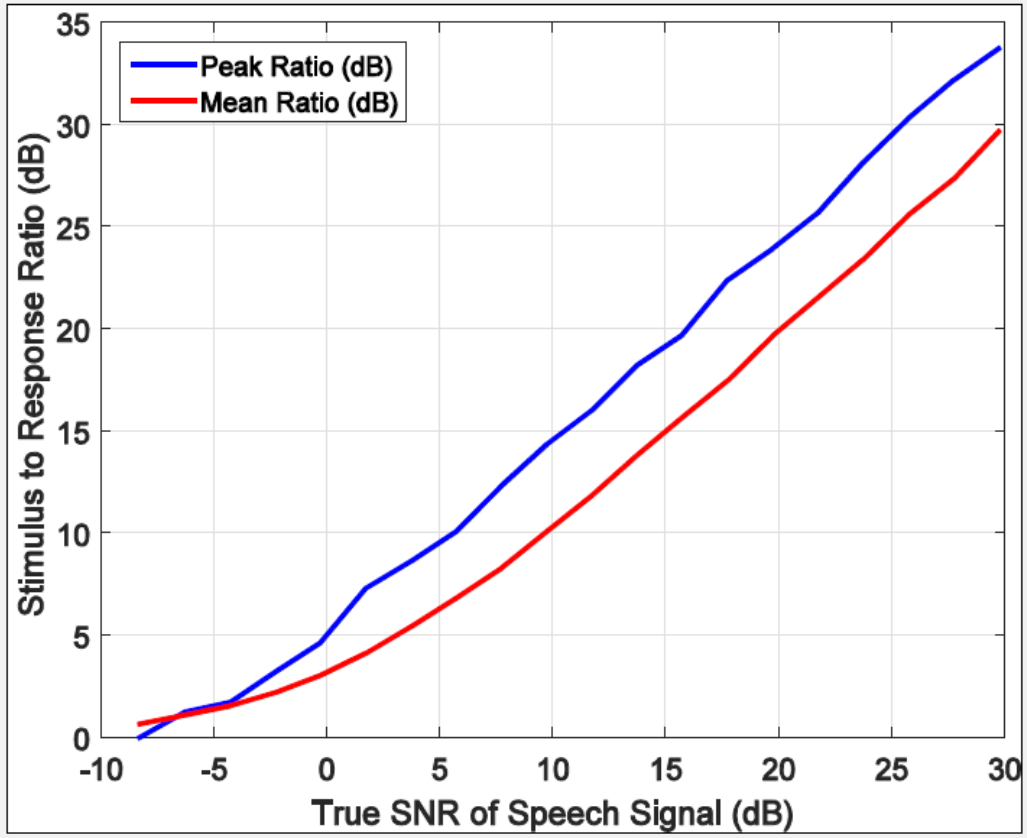
Peak to Mean Ratio of Background depends on Density of noise…

Which relates to overall level….

Simulations show 9-talker babble has P/M ratio of around 4 dB….

Walter Reed National Military Medical Center

Peak Target to Mean Background SNR Depends on Modulation in Background



Transformation from Peak-Target/Peak-Masker

To

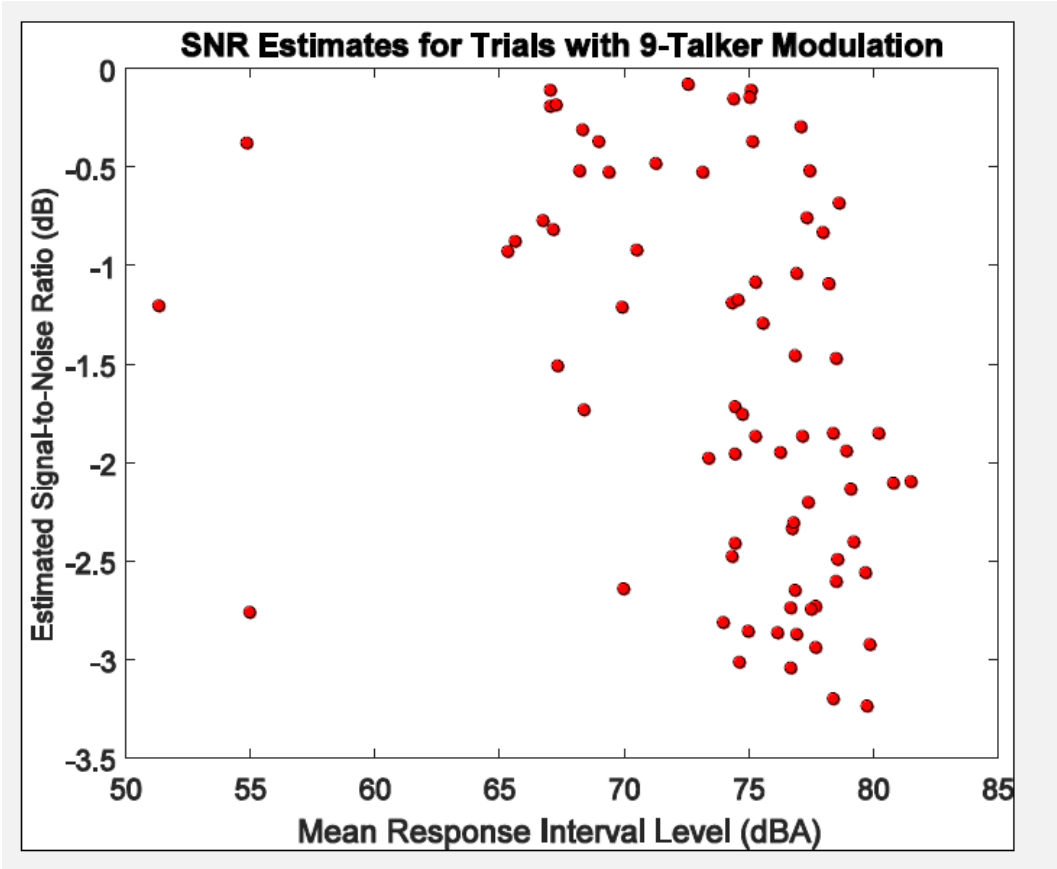
True SNR

Is a nonlinear transformation that requires a simulation and a lookup-table….

With 9-talker masker, 5 dB PT/PM ratio Is 0 dB SNR

Walter Reed National Military Medical Center

Peak Target to Mean Background SNR Depends on Modulation in Background



So..

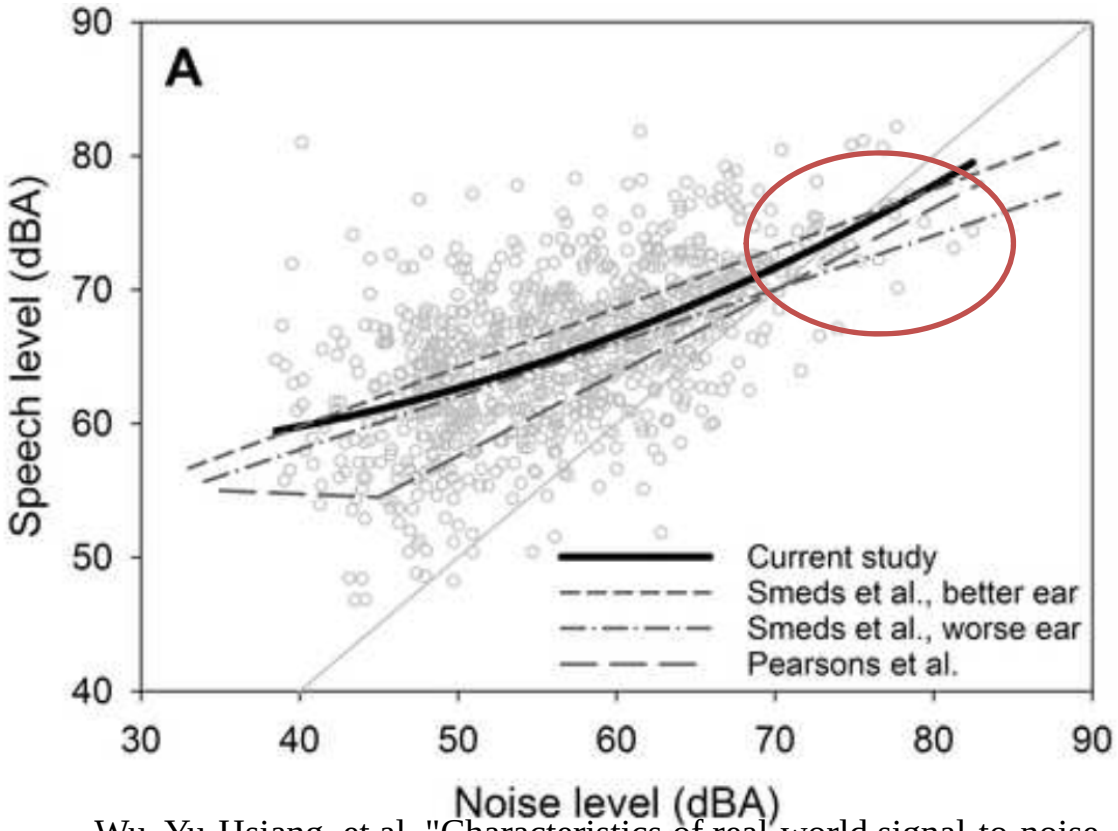
True SNR estimates for trials in this region

(from 70-80 dBA mean Backgound Level)

Range from -3 dB to 0 dB

Walter Reed National Military Medical Center

Comparison to Other Estimates of SNR



Wu, Yu-Hsiang, et al. "Characteristics of real-world signal-to-noise ratios and speech listening situations of older adults with mild-to-moderate hearing loss." *Ear and Hearing* (2017).

Our estimate of -1 to -3 dB
This is not far from
estimates from Smeds,
Wu, Pearson, etc….
at 70-80 dBA

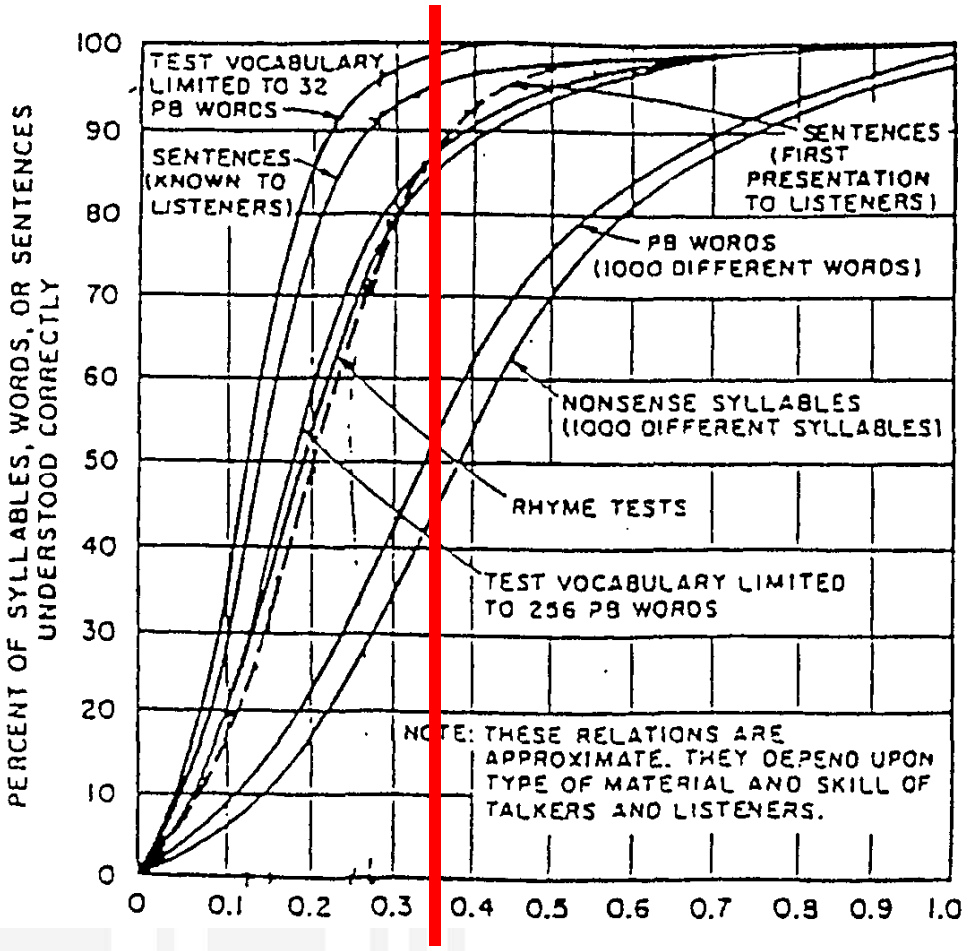
But

Our noise ranges were
at far upper end of range
of what they measured…

That's why they
report typical SNR of
+6-8 dB

Walter Reed National Military Medical Center

Consistent with Reported Values for MRT Score versus SNR



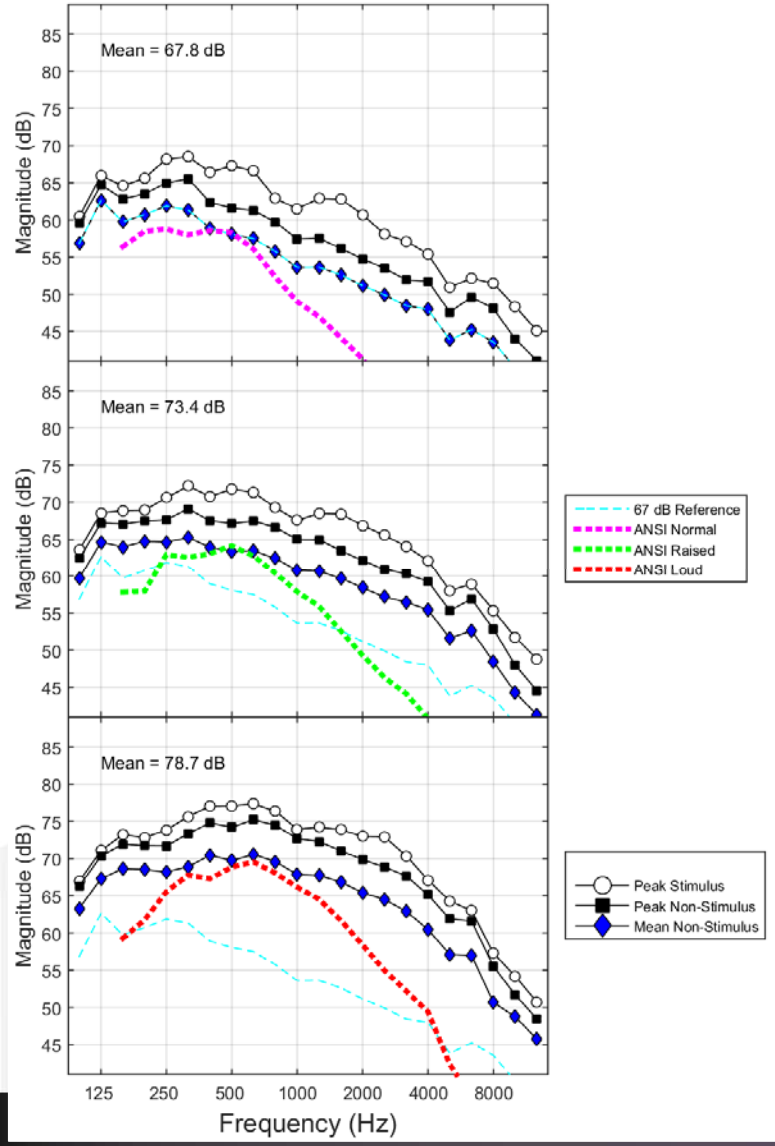
85% Score on MRT
is equal to 0.35 AI

In SSN,
That equals -1.5 dB

So.. close to
what we see in
real world

Walter Reed National Military Medical Center

Speech Spectrum



It is well known that target speech changes spectrum with level of vocal effort….

Less obvious is that the masker changes spectrum to match target

Related to all talkers speaking "same level" to be heard in background noise

Experiments with Babble Noise should Account for This

Walter Reed National Military Medical Center

Conclusion 4

In typical noisy restaurant,

Mean noise level is 75 – 80 dBA SPL

SNR is -3 to -1 dB

MRT Percent Correct is around 85% (AI=0.35)

In lab studies, target and maskers should have "raised" spectrum to reflect Lombard effect

Walter Reed
National Military
Medical Center

DIRECT AND INDIRECT METHODS OF MEASURING VISUAL CONTRIBUTION IN THE REAL-WORLD

Walter Reed National Military Medical Center

Impact of Visual Cues

One important difference between real world and (most) lab studies:

- **Availability of visual cues**

Visual cues are known to substantially improve intelligibility…

- **But how consistently are they used in real-world environments**

Cafeteria study offers opportunity to explore this in detail…

Walter Reed National Military Medical Center

Listener Instructions

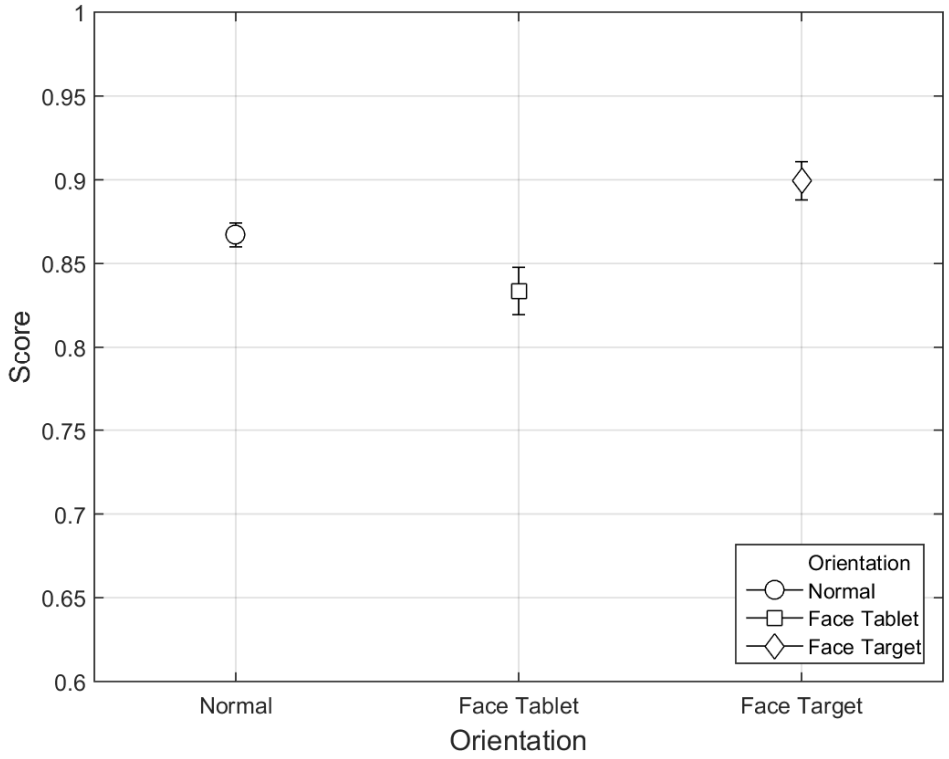
In roughly 20% of trials, Instructions were given on "where to face"

Face the target, or
Face the talker

Walter Reed National Military Medical Center

Impact of Visual Cues

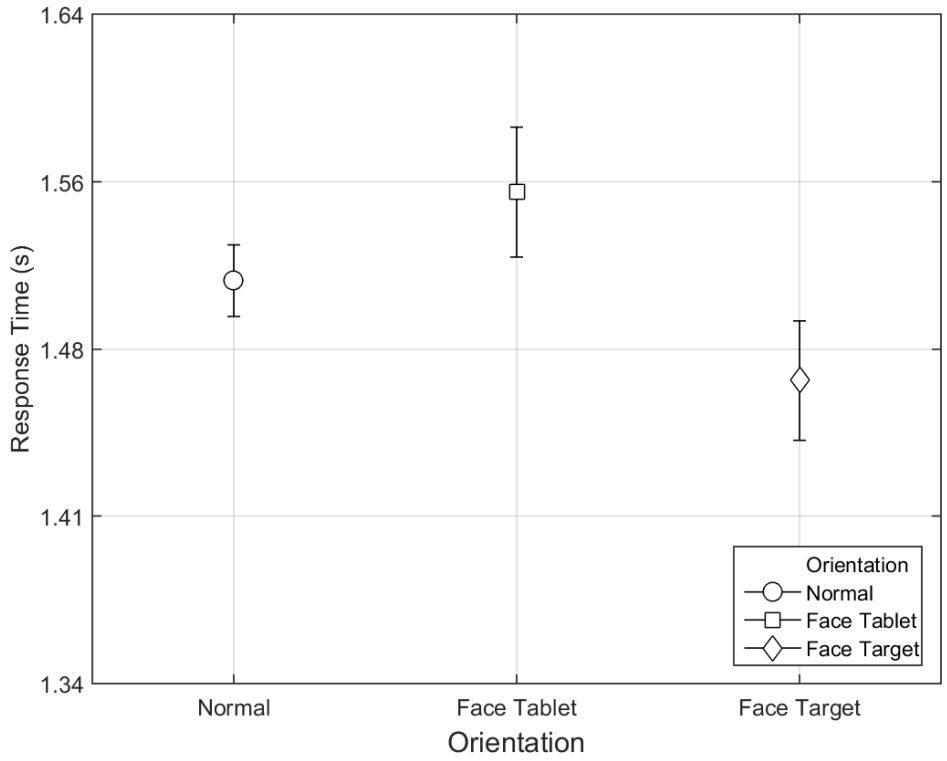
Overall Score was Highest in Face Target Trials lowest in Face Tablet trials



Walter Reed National Military Medical Center

Impact of Visual Cues

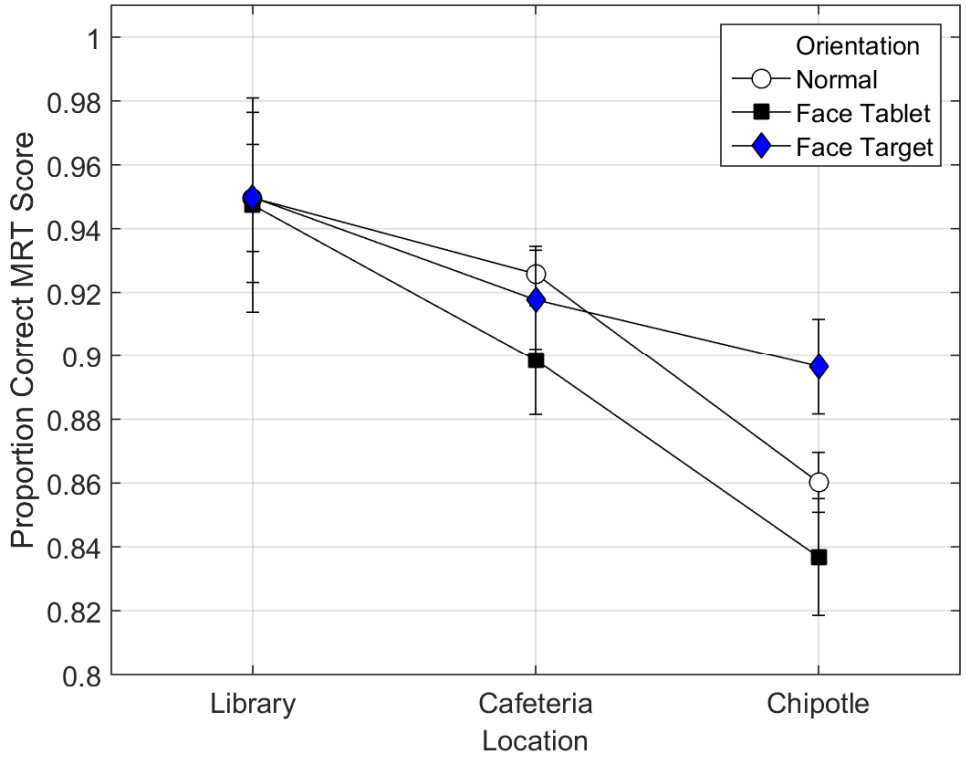
Reverse Function seen for Response Time



Walter Reed National Military Medical Center

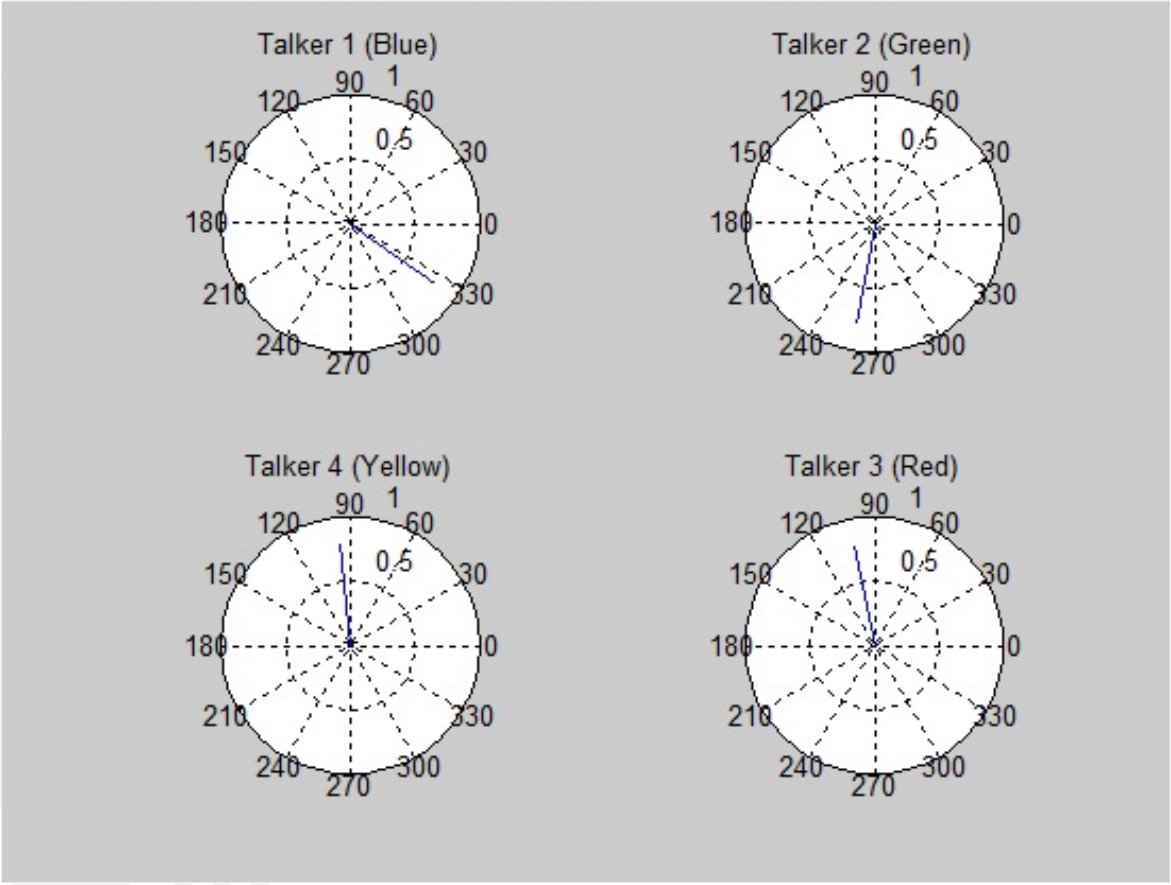
Impact of Visual Cues

Visual Cues matter more in noisy environments



Walter Reed National Military Medical Center

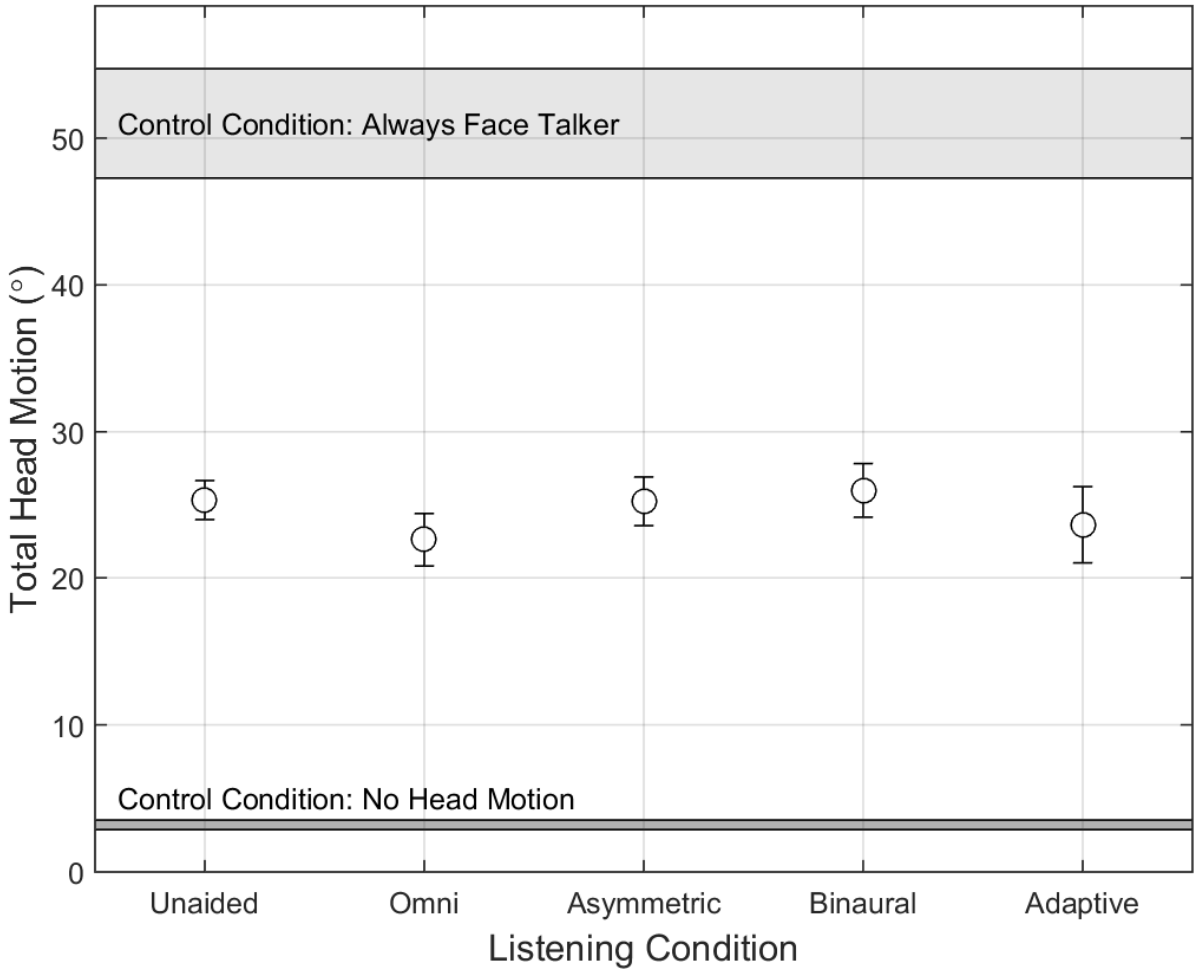
Direct Head Tracking



Walter Reed National Military Medical Center

Total Head Motion

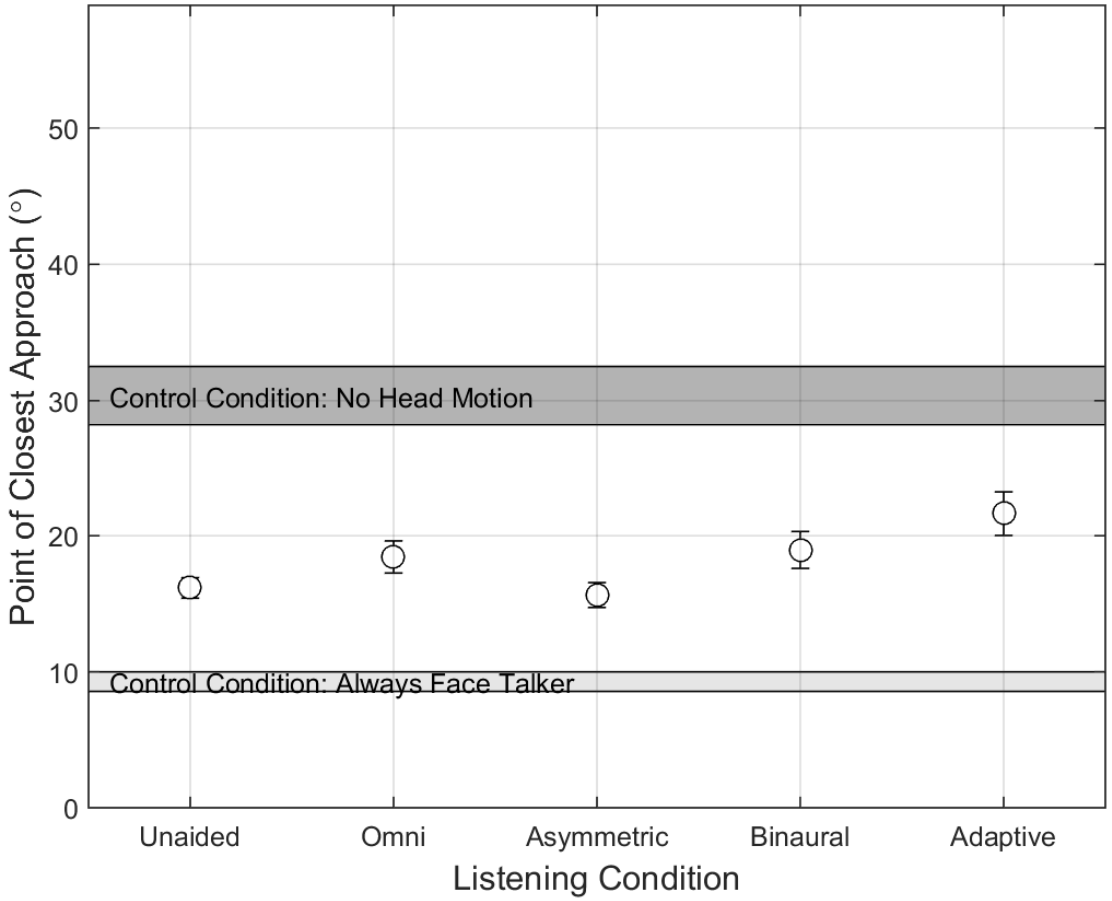
Listeners move their head more to face talker location



Walter Reed National Military Medical Center

Point of Closest Approach

They also turn within 10 degrees of talker *when instructed…*
but only halfway toward talker when no instructions are given



Walter Reed National Military Medical Center

Errors Can Be Scored as Visemes

/k/ as in kit.
/g/ as in got.
/ŋ/ as in thing.

/p/ as in pot.
/b/ as in bet.
/m/ as in mom

/ʃ/ as in shut.
/tʃ/ as in church.
/dʒ/ as in jump.
/ʒ/ as in measure.

/r/ as in rat.

/d/ as in dog.
/t/ as in top.
/n/ as in nod.

/f/ as in fog.
/v/ as in vat.

Walter Reed National Military Medical Center

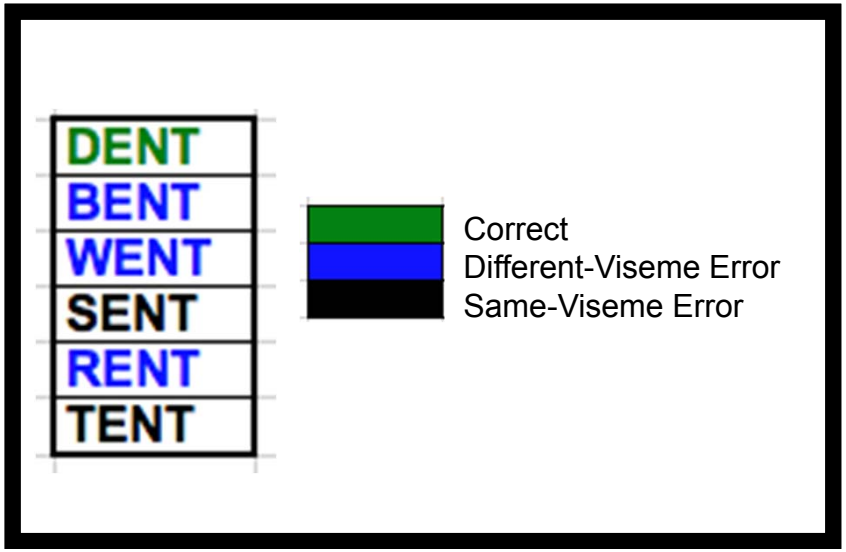
Viseme Analysis

/d/ as in dog.
/t/ as in top.
/n/ as in nod.
/s/ as in sod

/r/ as in rat.
/w/ as in wig

“

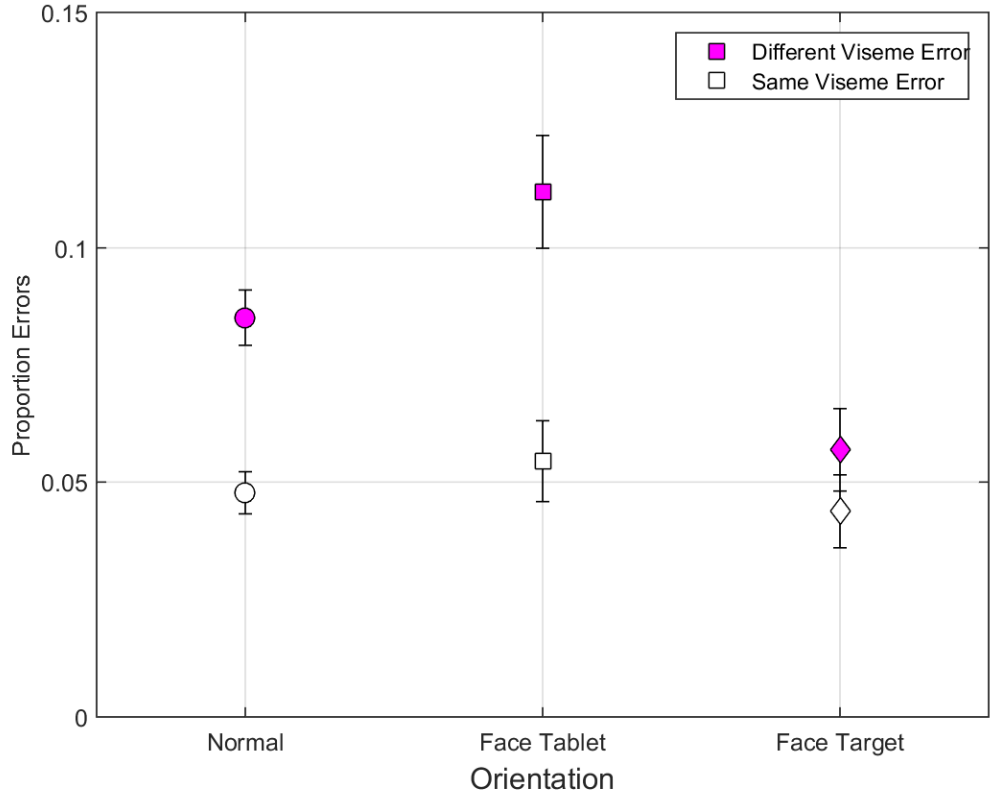
“You will mark **DENT** please”



Walter Reed National Military Medical Center

Viseme Analysis

Percent viseme and non-viseme errors as a function of orientation

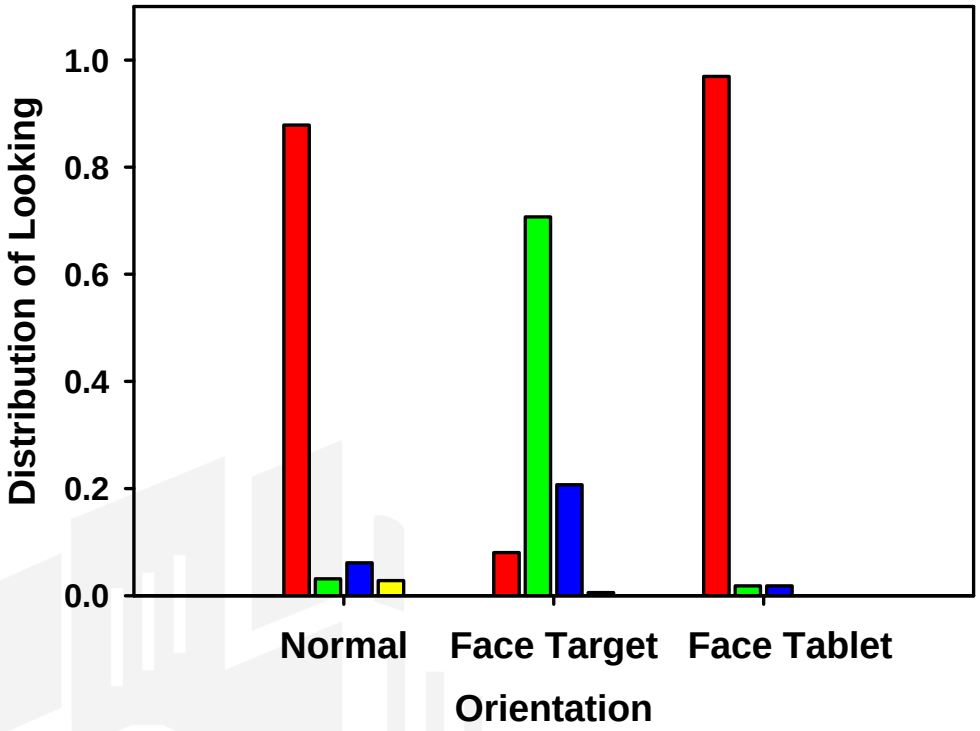


Both Score and
Viseme Analysis
Show “Normal”
Is halfway between
“Face Target” and
“Face Tablet”

Strongly suggests that
Listeners are using visuals
Roughly 50% of the time
In the “Normal” condition

Walter Reed National Military Medical Center

Eyetracking Results



However, limited “eyetracking” data shows listener hardly ever looks at target in normal condition

A puzzle…. Do they know when to look?

Figure 6: Gaze Direction Shown By Green Dot (Center Point of Focus)

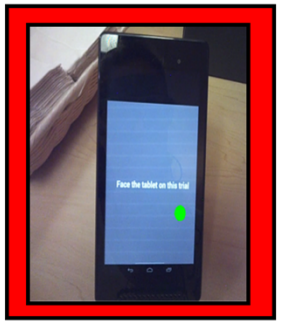


Figure 6a: Looking at Tablet

Figure 6b: Looking at Target

Figure 6c: Looking Close to Target

Figure 6d: Looking Away

Walter Reed National Military Medical Center

Conclusion 5

Visual cues matter in Real-World Environments…

- **They play role when noise level exceeds 70 dB**

- ***Most* evidence suggests unsupervised listeners obtain roughly 50% of possible benefit from visual cues**

- **This ratio can potentially be estimated by visemic error rate**

Walter Reed
National Military
Medical Center

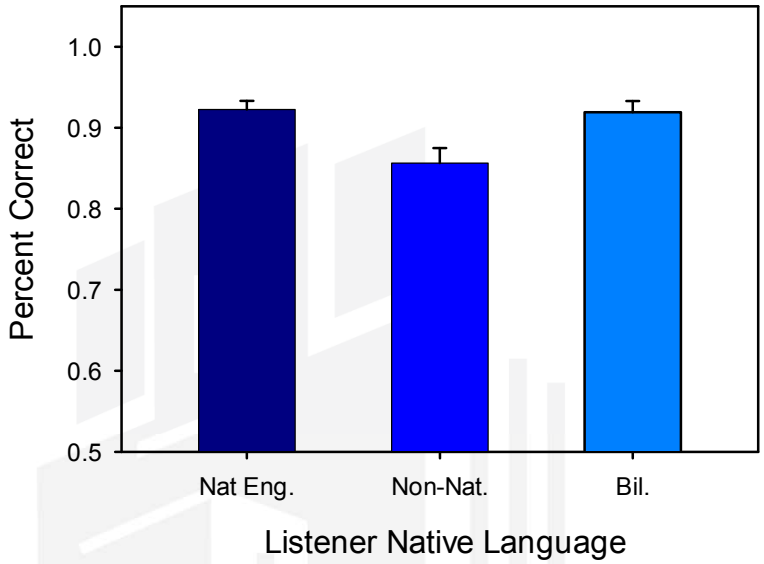
PERFORMANCE OF NATIVE, NON-NATIVE, AND BILINGUAL SPEAKERS OF ENGLISH ON REAL-WORLD SPEECH PERCEPTION

Walter Reed National Military Medical Center

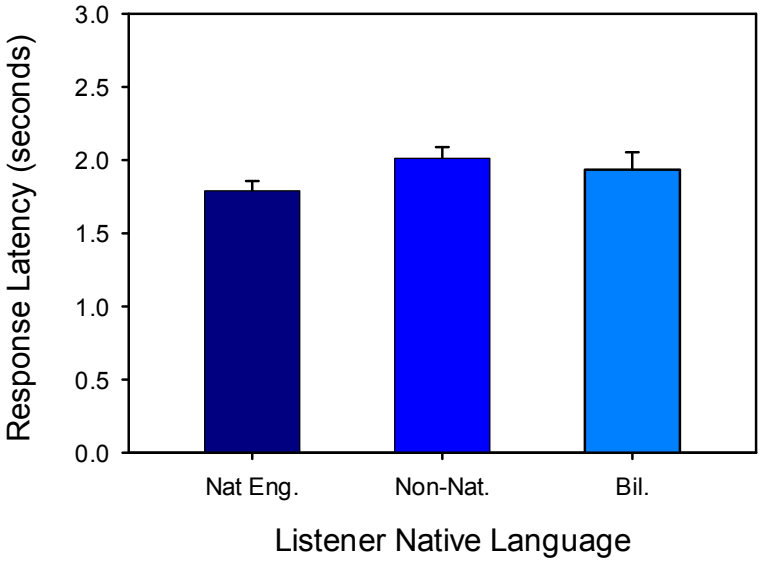
Native English Talker

Non-Native Listeners have roughly a 10% penalty when listening to a Native talker

Recognition of Native English speech



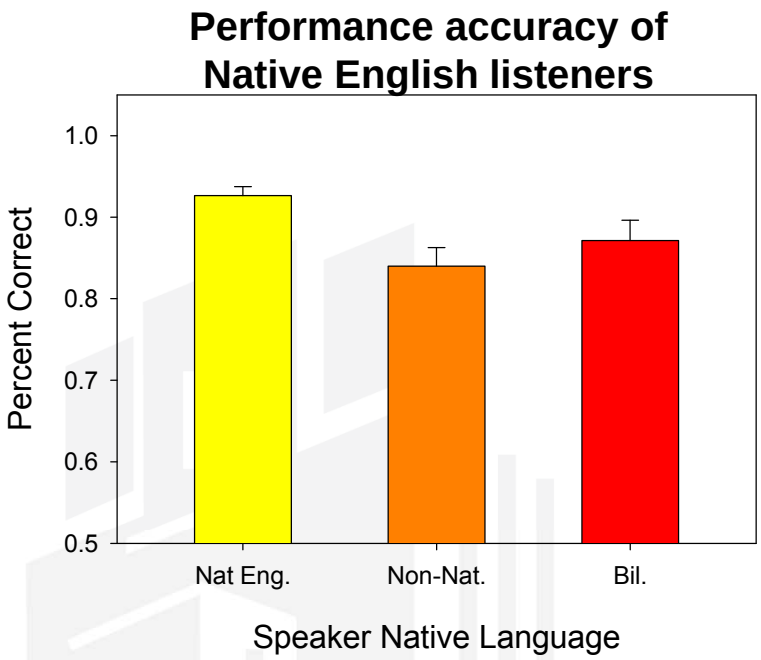
Response Latency for Native English speech

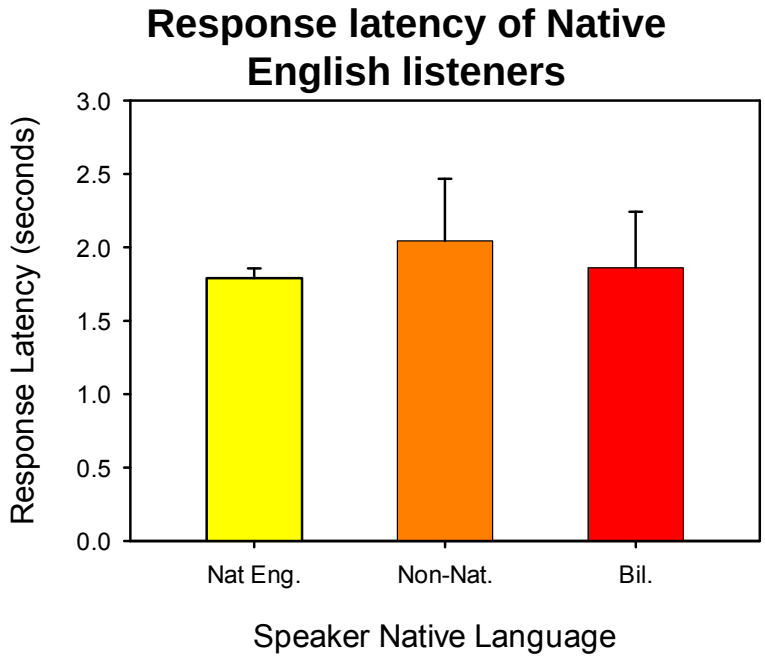


Walter Reed National Military Medical Center

Non-Native Talker

Native Listeners have roughly a 10% penalty when listening to a Non-Native Talker

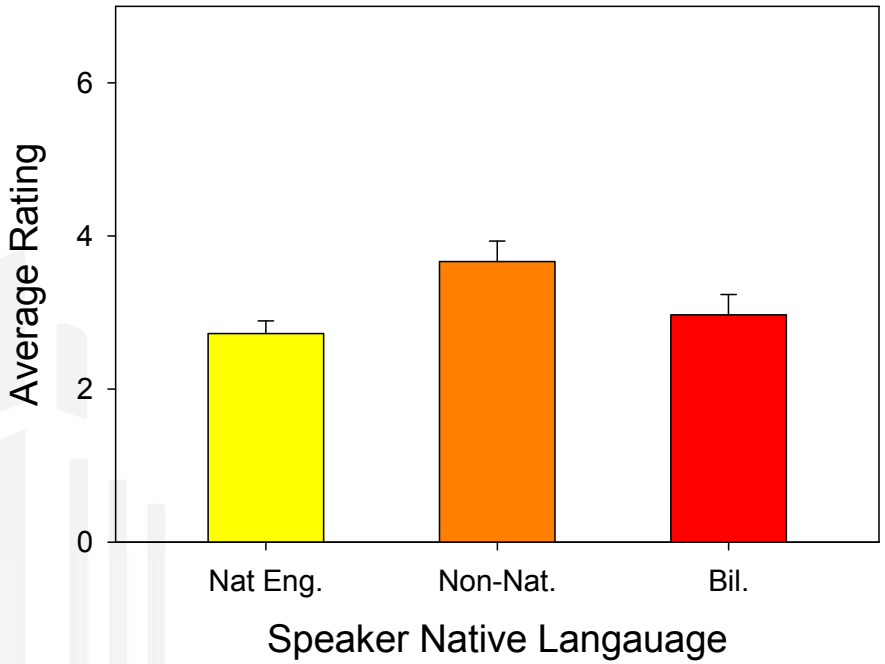




Walter Reed National Military Medical Center

Talker Difficulty Ratings

This is reflected in perceived difficulty ratings



Walter Reed National Military Medical Center

Conclusion 6

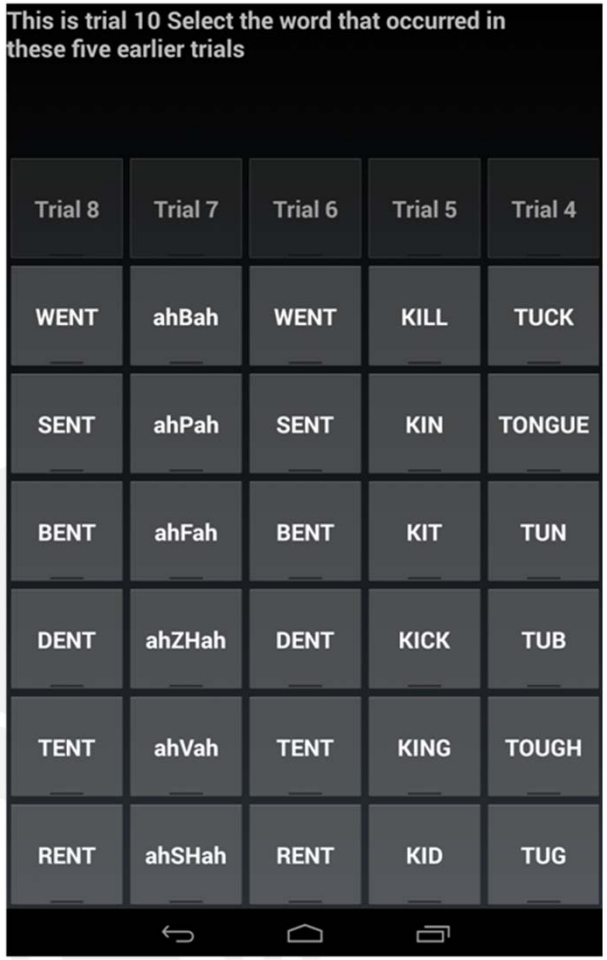
Non-native talkers and listeners are at a significant disadvantage when listening in noisy environments

Walter Reed
National Military
Medical Center

PRELIMINARY EVALUATIONS OF WORKING MEMORY IN REAL-WORLD ENVIRONMENTS

Walter Reed National Military Medical Center

The N-Back Memory Task

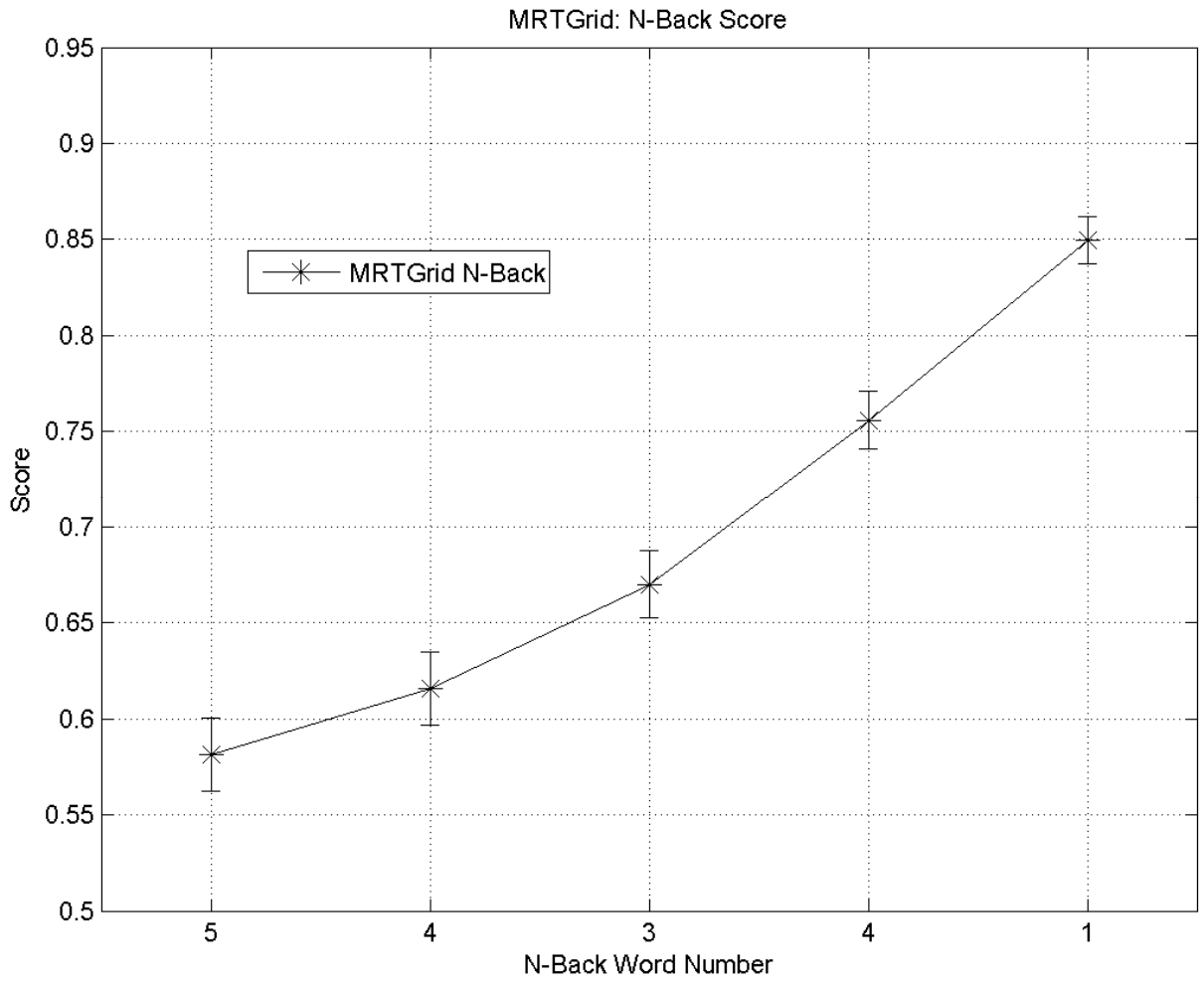


At random intervals,

Talkers are asked to
Identify last 5 MRT words

Walter Reed National Military Medical Center

The N-Back Memory Task

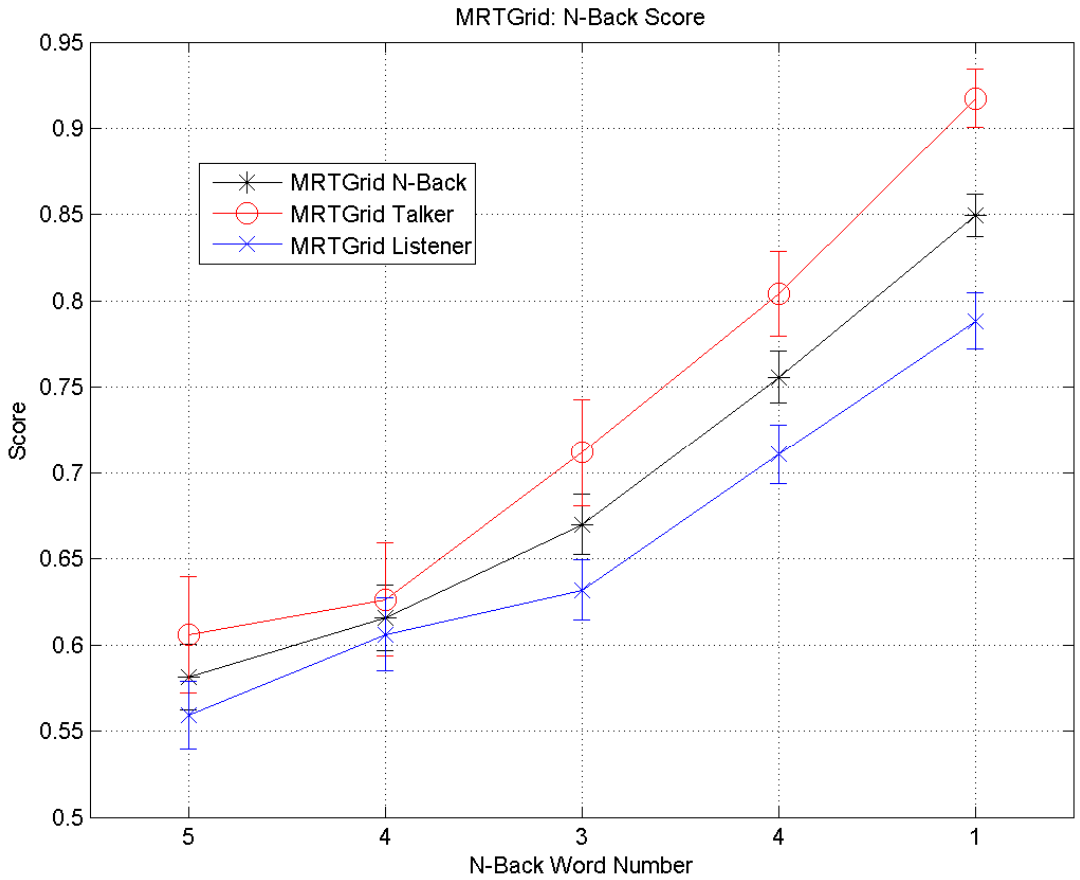


Exponential Decay in Recall Rate…

But well over chance even For 5-back word

Walter Reed National Military Medical Center

The N-Back Memory Task



Recall is much better for words spoken by the responding talker

Walter Reed National Military Medical Center

Overall Conclusions

Relatively little data is available on “real-world” listening

- Particularly in crowded public spaces
- Here we have >18000 trials, >60 subjects

Characteristics of typical “challenging restaurant”

- Noise level of 75-80 dBA SPL
- Overall percent correct on MRT roughly 85-90%
- SNR of roughly -1 dB
- Roughly 50% utilization of visual cues

Performance is stable over wide range of noise levels

- Score drops from 100% to 85% as noise increases from 55 dB to 85 dB
- This shallow slope reflects “adaptability” of talkers and listeners
- Greater noise levels may be necessary to “stress” the system