Combating the Reverberation Problem

Barbara Shinn-Cunningham (Boston University)

- How humans cope in natural settings

Martin Cooke (Sheffield University, U.K.)

- How speech is corrupted by reverberation

DeLiang Wang (Ohio State University)

- Effects of reverberation on pitch, onset/offset, and binaural cues

Joseph Desloge (Sensimetrics Corp.)

- Multi-microphone source separation in reverberant environments

Barbara Shinn-Cunningham (Boston University)

- How humans cope in natural settings

Martin Cooke (Sheffield University, U.K.)

- How speech is corrupted by reverberation

DeLiang Wang (Ohio State University)

- Effects of reverberation on pitch, onset/offset, and binaural cues

Joseph Desloge (Sensimetrics Corp.)

- Multi-microphone source separation in reverberant environments

To understand speech in a "natural" setting, we must

HEAR elements of the target speech (mixed in)

SEPARATE the target speech from other competing sources

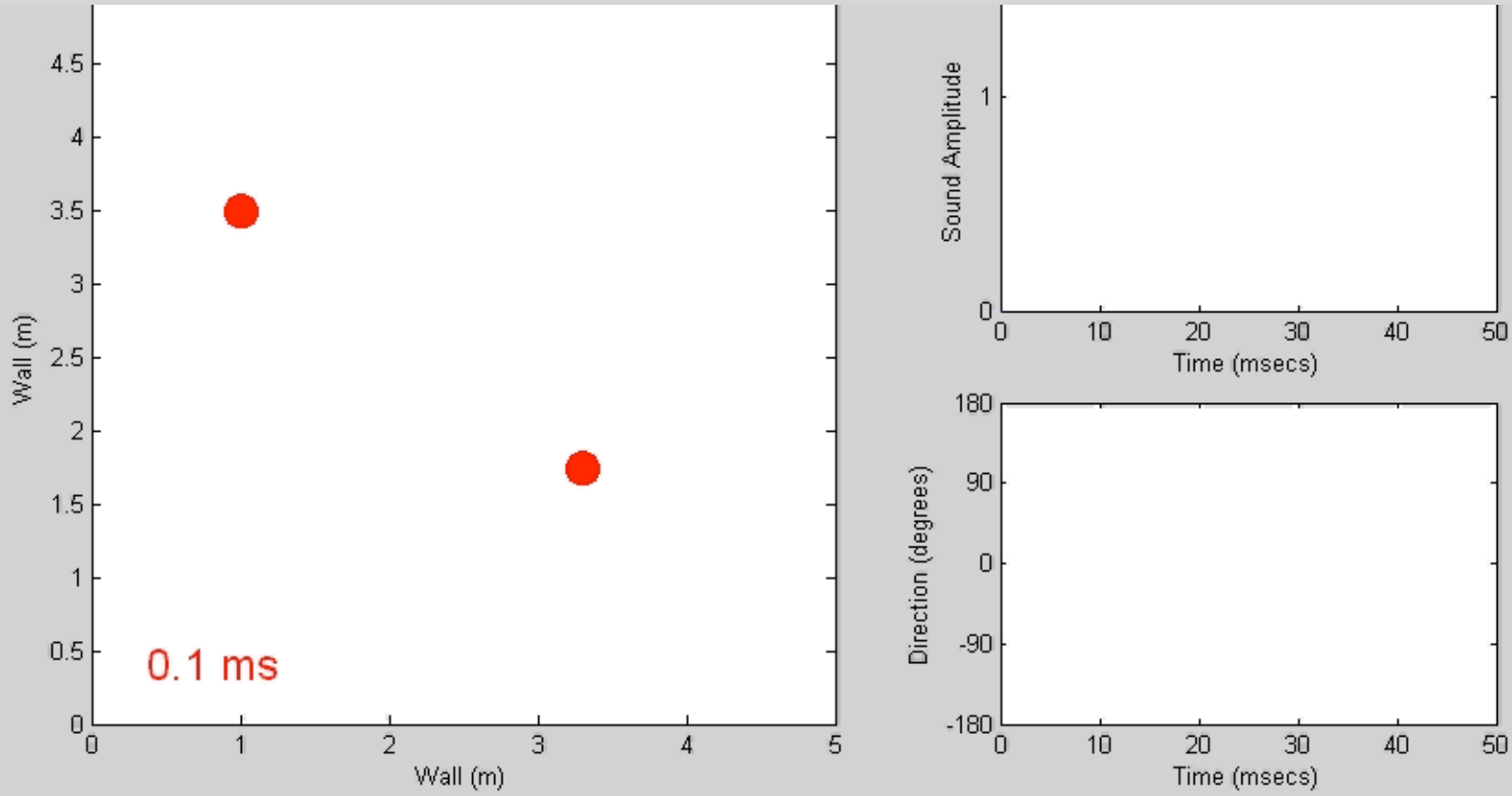
GROUP and STREAM together the pieces of target we hear and have isolated

FILL IN missing information (through statistics at many levels, from acoustics to meaning) sufficiently to comprehend the message

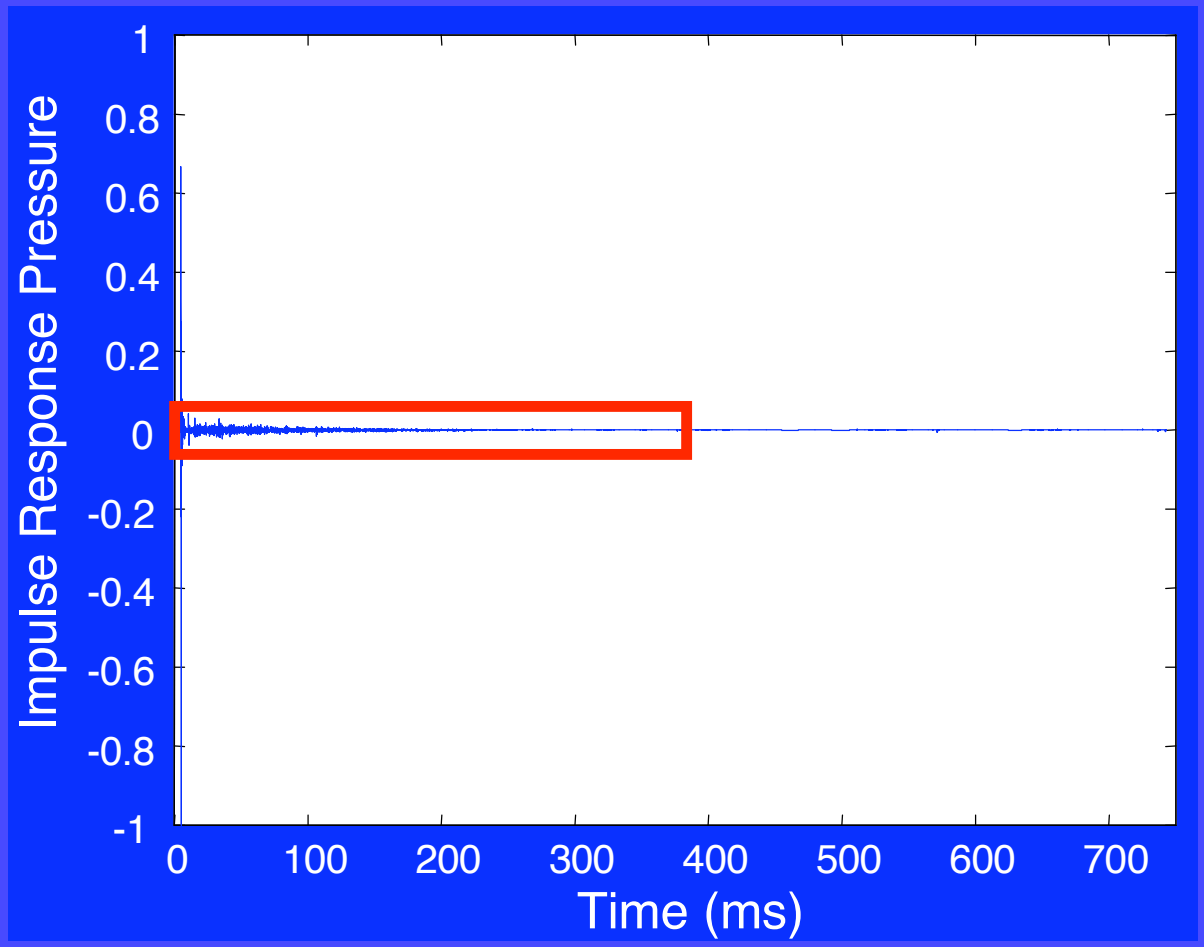
INTERPRET the information

Natural environments have reflections that influence the signals at the ears

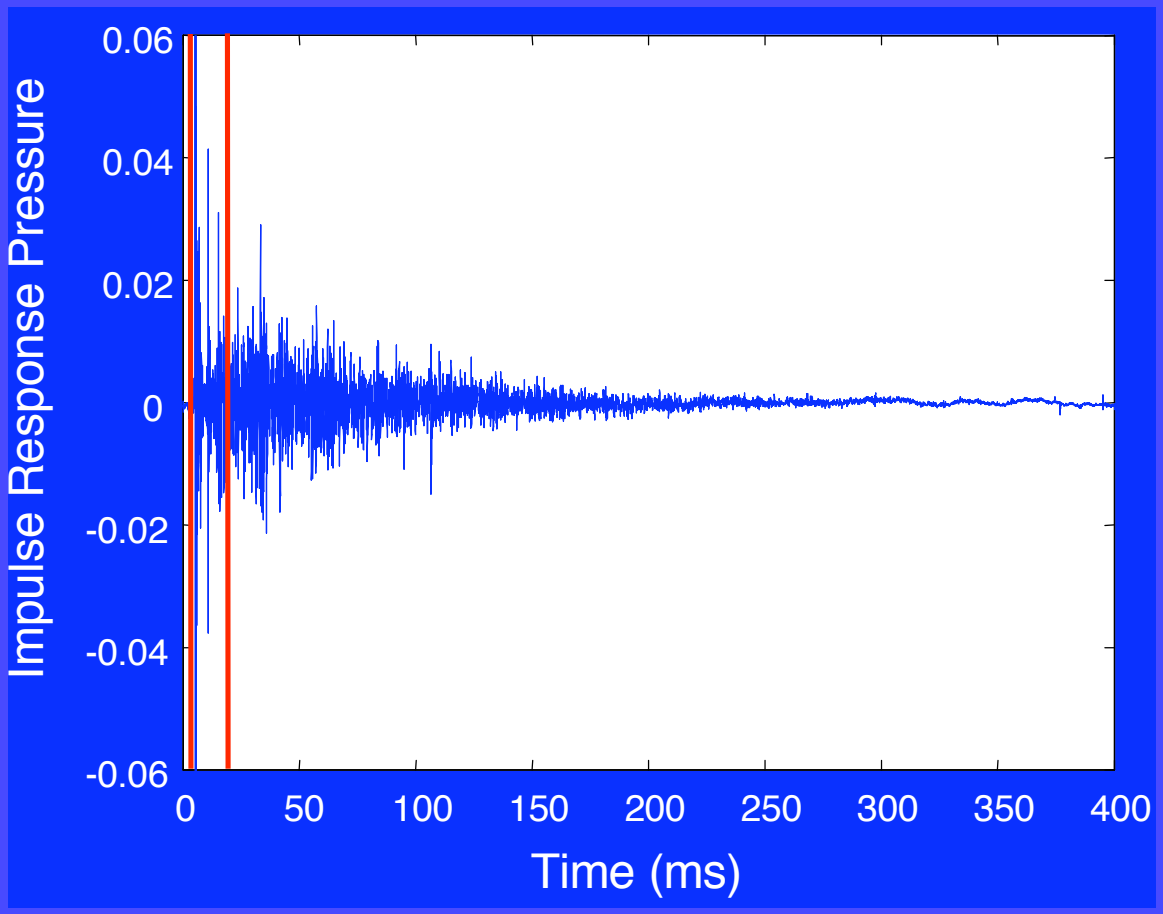
(Courtesy M. Akeroyd, Institute for Hearing Research, Glasgow)



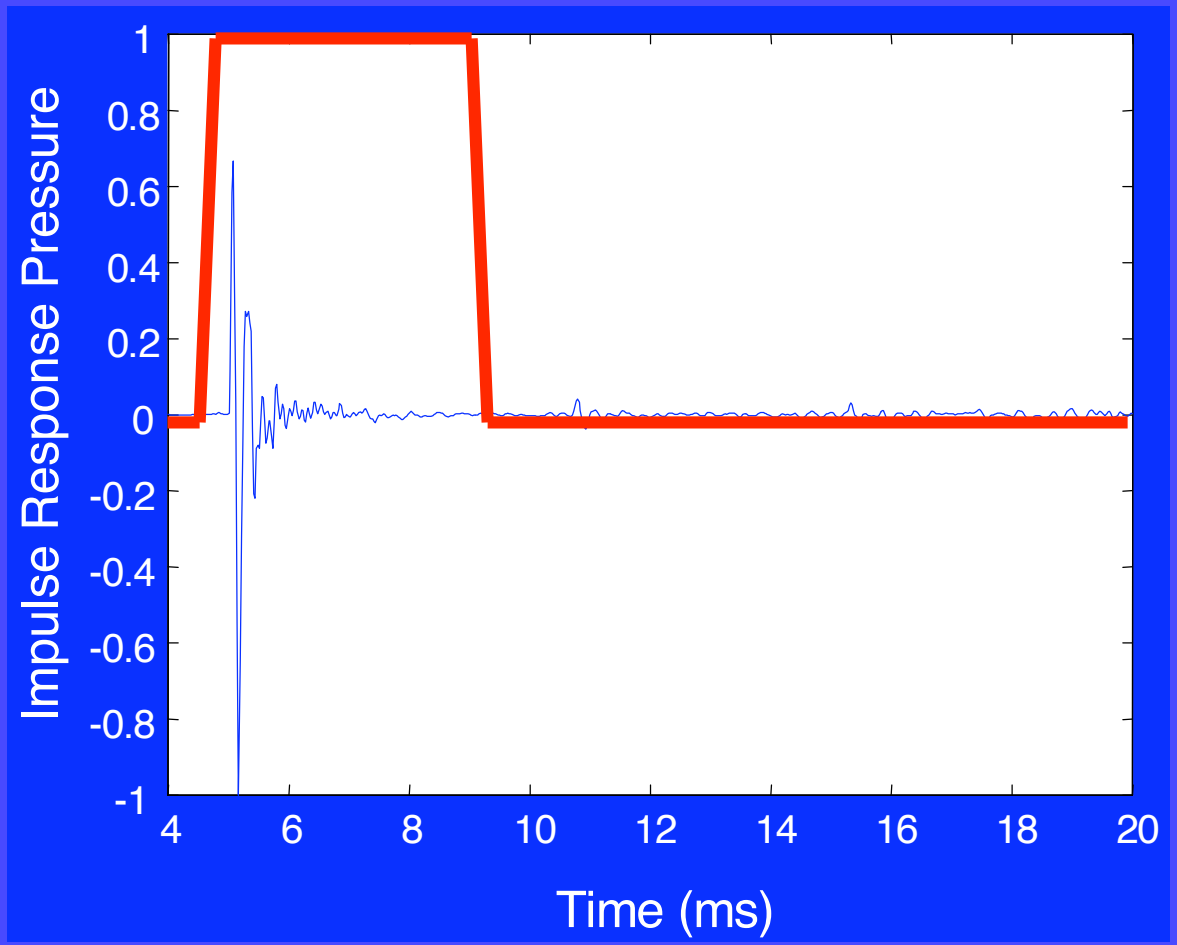
Reverberant Head-Related Impulse Response



Reverberant Head-Related Impulse Response

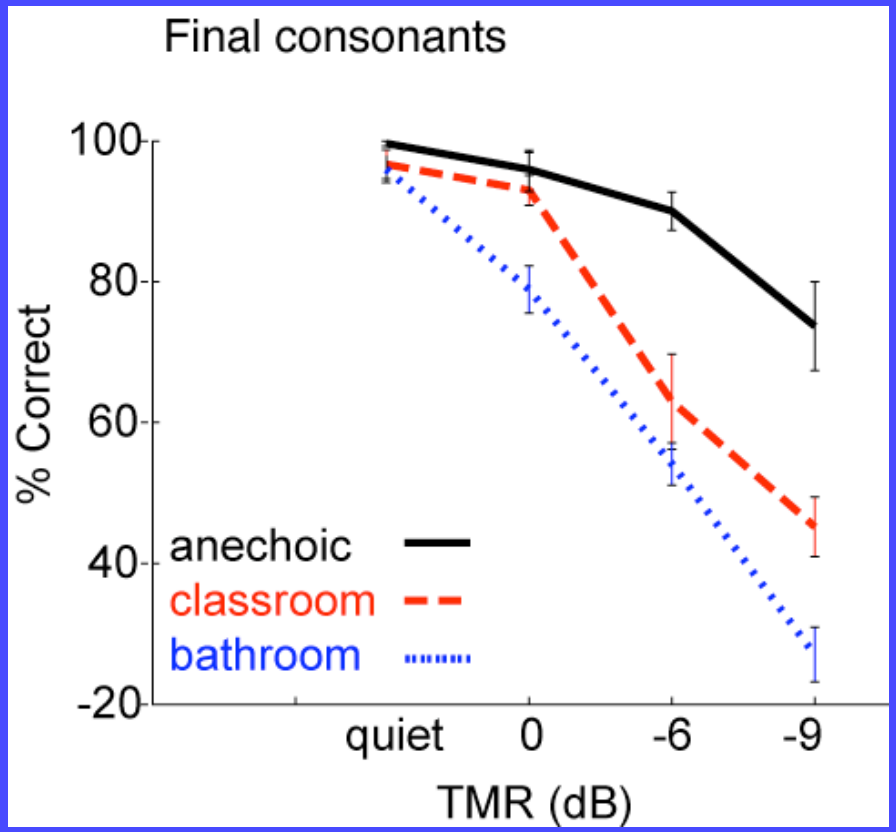


Reverberant Head-Related Impulse Response



Reverberation PLUS noise degrades understanding more than either alone

Devore and Shinn-Cunningham, 2003, Proc ICAD



How does reverberation influence the various processes involved?

HEAR  rget speech (mixed in)

SEPARATE  ech from other competing s

GROUP and STREAM  he pieces of target we hear and hav d

FILL IN  ng information (through statistics at many s, from acoustics to meaning) sufficien o comprehend the message

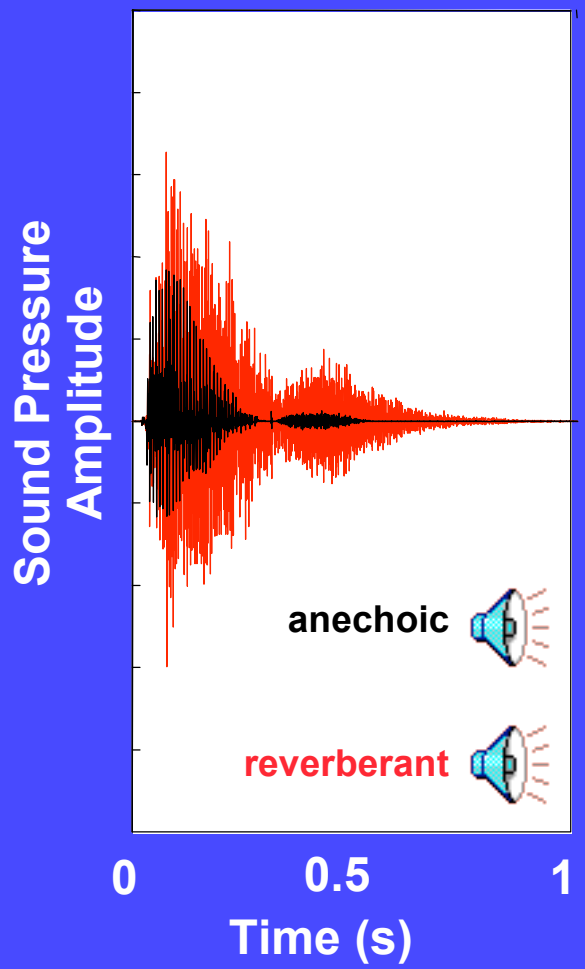
INTERPRET  nformation

How humans cope in natural settings

SEPARATE the target speech from other competing sources

Or, in the case of reverberant settings, from extra reverberant energy

Reverberation smears spectrotemporal features on time scale of 10s-100s of ms



Speech transmission index (STI) predicts perceptual degradation
(e.g., Houtgast and Steeneken, 1985, JASA)

- Early reflections (within ~50 ms) boost energy without distorting modulation / meaning (e.g., Bradley et al., 2003, JASA)
- Late reflections degrade intelligibility

For moderate-sized rooms, degradation not very severe

"Traditional" spatial unmasking is about hearing the source

A signal element just below monaural detection threshold becomes detectible from the interference it can cause (interaural decorrelation)

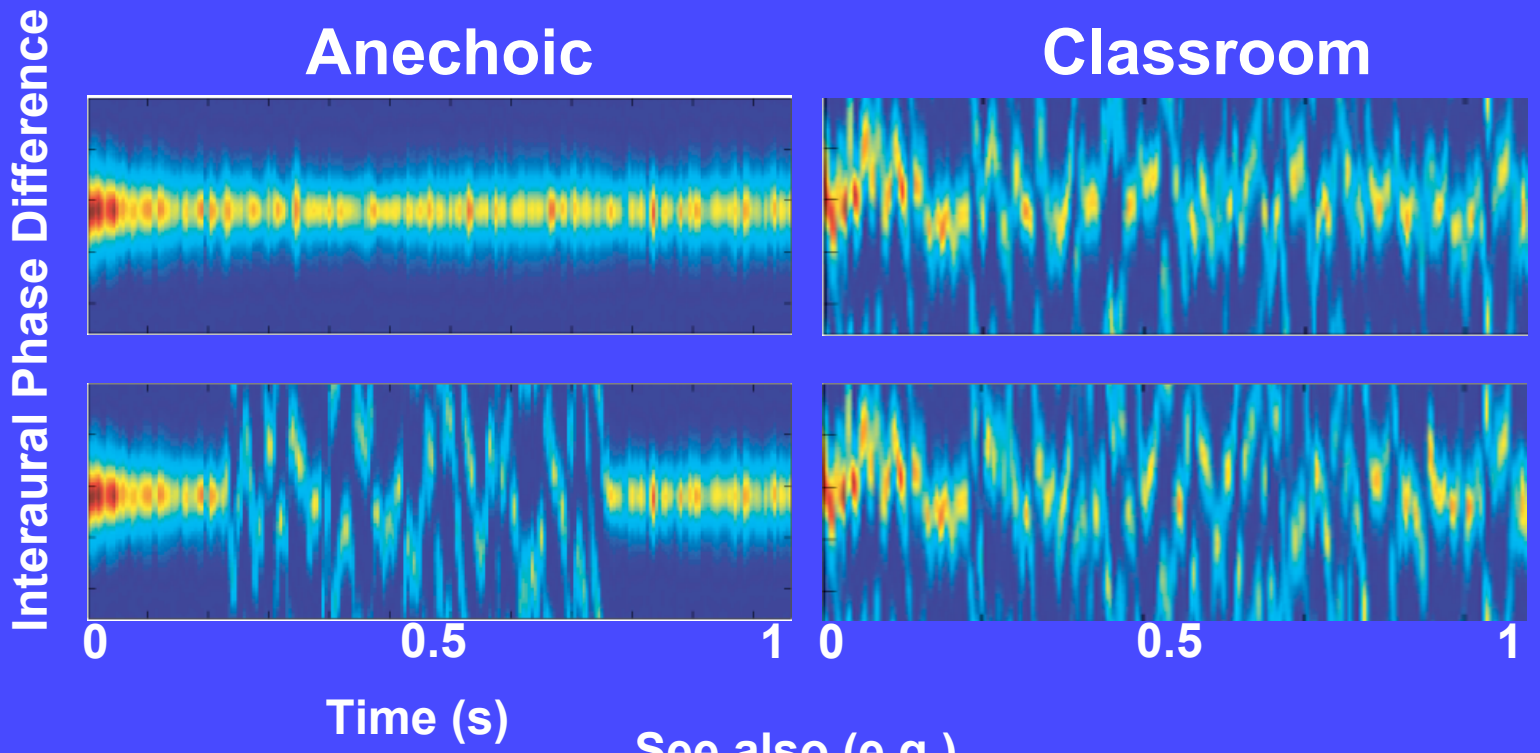
Applied to speech:

Hovel, 1984

Zurek, 1990

Wagener, Brand, others in Oldenburg…

Spatial unmasking due to binaural processing degrades with reverberation



See also (e.g.)
Zurek et al., 2004, JASA
Plomp, 1976, Acustica

GROUP and STREAM together the pieces of target we hear and have isolated

What does reverberation do to grouping and streaming?

Reverberation degrades separation cues, reducing their effectiveness

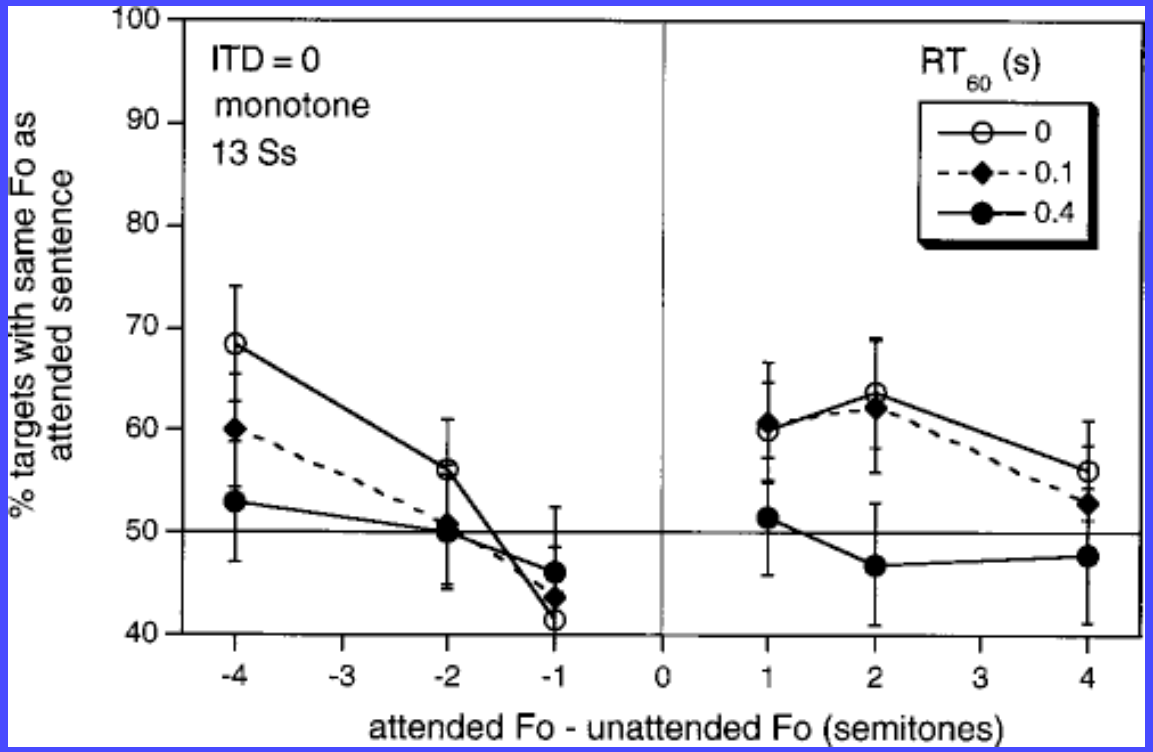
- Onset / offset
- Harmonic structure (time-varying F0)
- Spatial cues

Prosody cues contribute less to separation in reverberant settings

Darwin and Hukin, 2000, JASA

Bigger influence of prosody

No effect of prosody



However, perceived location influences streaming of similar competing signals…

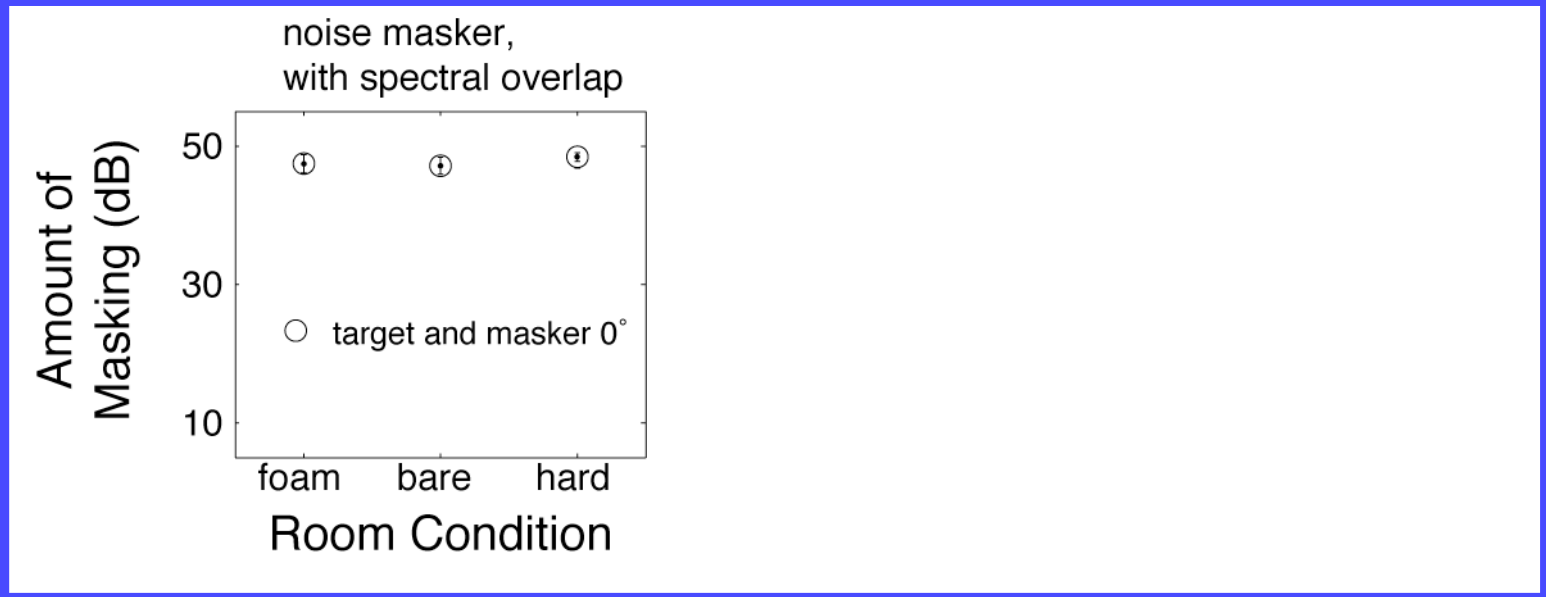
Freyman et al., 1999, JASA

an effect that is robust in rooms

Kidd et al., in press, Acustica

For overlapping spectra, speech-in-noise masking is great

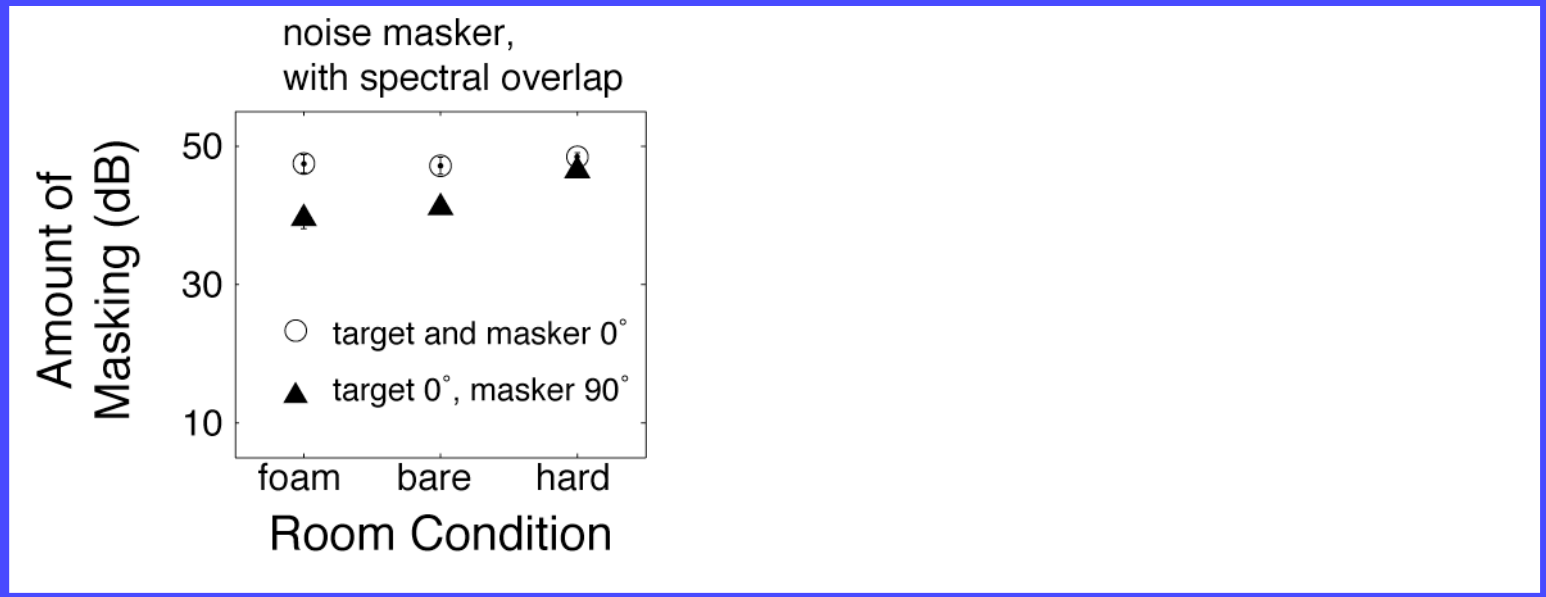
Kidd et al., in press, Acustica



increasing reverberation

Spatial unmasking of speech in noise decreases with increasing reverberation

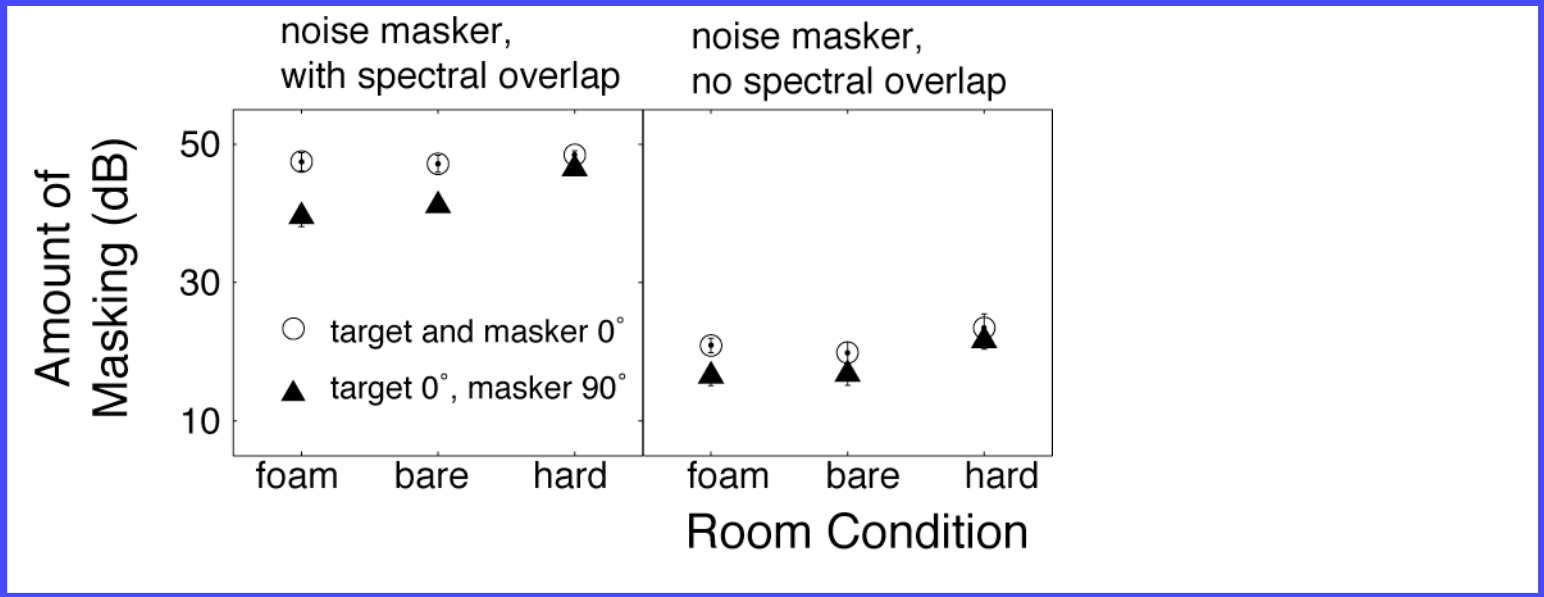
Kidd et al., in press, Acustica



increasing reverberation

For non-overlapping noise, there is little masking and little spatial unmasking

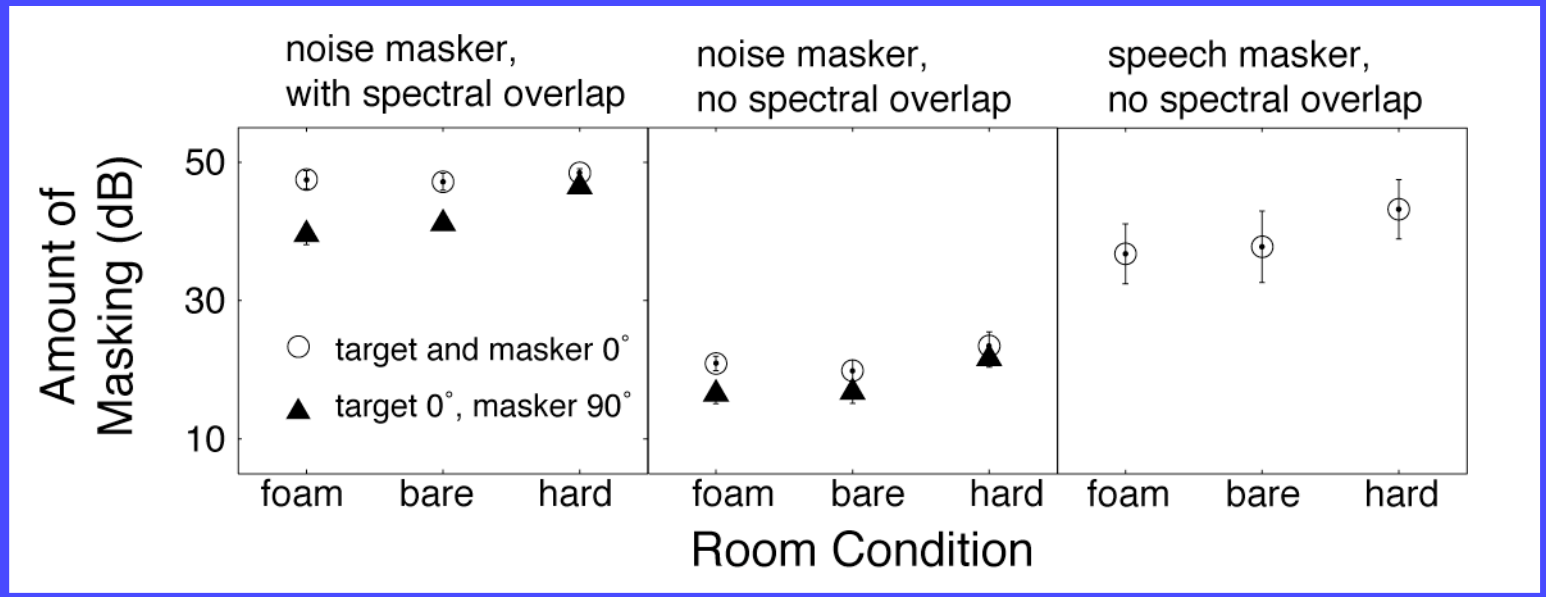
Kidd et al., in press, Acustica



increasing reverberation

Even without spectral overlap, there is a lot of speech-on-speech masking

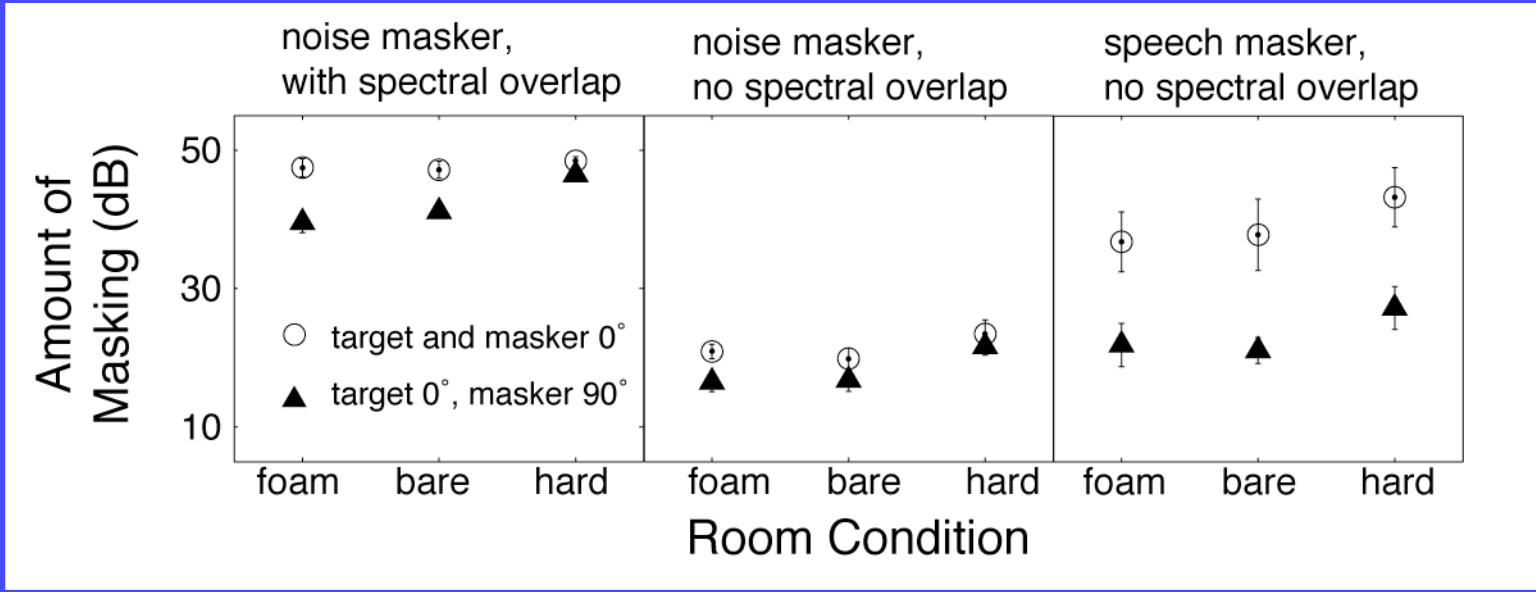
Kidd et al., in press, Acustica



increasing reverberation

Spatial unmasking is large for speech-on-speech masking in all environments

Kidd et al., in press, Acustica



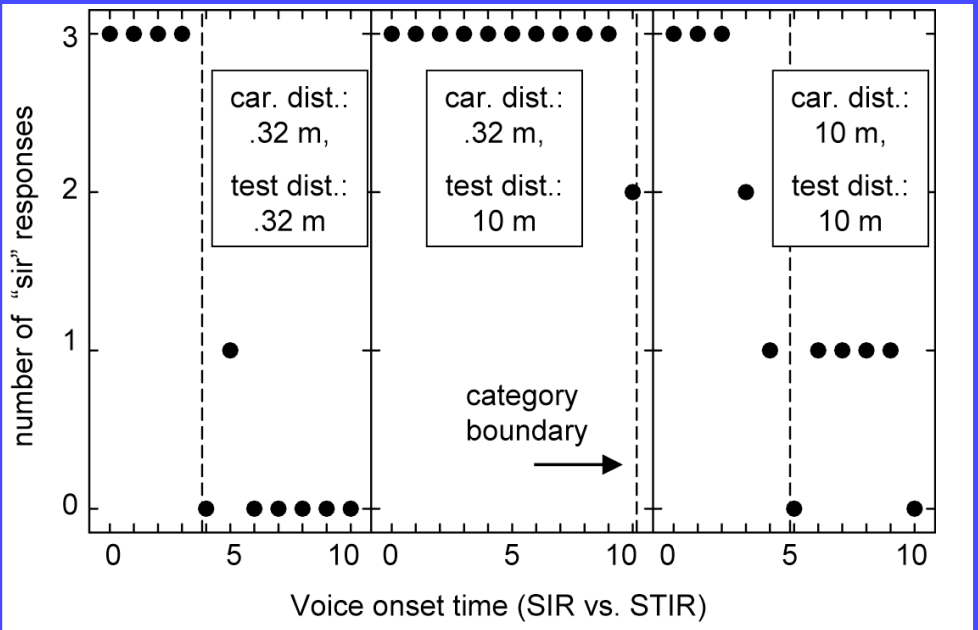
increasing reverberation

INTERPRET the information

Does reverberation affect interpretation?

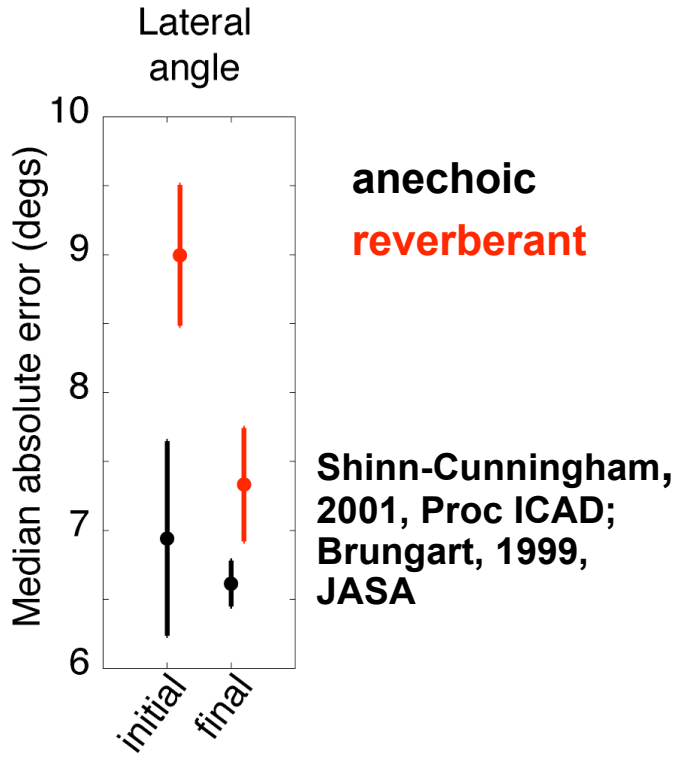
Listeners calibrate to the environment

Listeners change how they interpret spectrotemporal content to compensate for room smearing



Watkins, 2004, Brit Soc Aud

Localization improves with experience in a room



Shinn-Cunningham, 2001, Proc ICAD; Brungart, 1999, JASA

How do humans cope in natural settings?

Reverberation degrades how well we HEAR elements of a source, e.g., binaural unmasking

Reverberation degrades how well we GROUP and STREAM.

In INTERPRETING sources (or locating them), we compensate for the expected effects of reverberation, calibrating to the environment:

Moderate reverberation is not a problem, alone.

Segregation / streaming may be the main issue.

Humans use all cues, accruing information over time, to maintain robust perception.

Acknowledgements

Sasha Devore, Kosuke Kawakyu, Gerald Kidd, Nat Durlach, Steve Colburn, Erick Gallun

Air Force Office of Scientific Research

Office of Naval Research

NIDCD