#61

SPATIAL UNMASKING OF SPEECH IN SIMULATED ANECHOIC AND REVERBERANT ROOMS

Barbara Shinn-Cunningham[1,2,3], Scarlet Constant[1,4], and Norbert Kopčo[1,3]

[1]Boston University Hearing Research Center, Depts of [2]Biomedical Engineering,[3]Cognitive and Neural Systems, and [4]Medical Sciences

ABSTRACT

Masked speech reception thresholds were measured for a speech source in the presence of a speech-shaped noise masker for simulated anechoic and reverberant listening conditions. Both speech and masker sources were simulated using individualized HRTFs. The HRTFs were measured in a moderately reverberant room (T_{60}=550 ms) for sources at different distances (15, 100, and 200 cm) and directions (straight ahead and directly to the right of the subject). Reverberant simulations were generated using the full HRTFs (including reverberation), while anechoic simulations were generated by time windowing the full HRTFs to create pseudo-anechoic HRTFs. Speech and noise sources were then convolved with the appropriate HRTFs to simulate anechoic and reverberant simulations for different speech and noise configurations. For each spatial configuration, subjects were tested binaurally, monaurally with the “better” ear, and monaurally with the “worse” ear. Speech reception thresholds were measured adaptively, varying the target level while keeping the direct portion of the masker constant at the better ear. Results suggest that speech intelligibility improves and spatial unmasking increases when reverberation is included, at least for some of the tested spatial configurations.

Work supported by AFOSR Grant No. F49620-01-1-0005 and the Sloan Foundation.

1. BACKGROUND

Angular separation of target / masker improves
- detectability of target
- speech reception of target (Bronkhurst, 2000)

Two factors often identified:
- head shadow (acoustic "better-ear")
- binaural advantage (Zurek, 1993)

Reverberation may (Drullman et al., 1994; Bradley et al., 1999)
- increase monaural speech audibility
- degrade monaural speech intelligibility
- decrease binaural coherence

Targets and maskers may be within arm's length (Shinn-Cunningham et al., 2000
- large energy effects
- large (x-frequency) interaural differences (head shadow + relative distance to ears)

Current study, measure speech thresholds for
- left, right, binaural listening conditions
- anechoic and reverberant simulations
- different target distances and directions

How does realistic reverberation affect masked speech intelligibility (using characterizable stimuli; e.g., Drullman & Bronkhorst, 2000)?

2. METHODS

Four subjects with normal-hearing

Simulate nearby sources using HRTFs
- Target (T): low-context sentences (e.g., see Payton, Uchanski, Braida; 1994).
- Masker (M): speech-shaped noise

Six spatial configurations
- M always at (0º az, 15 cm distance)
- T at (0º or 90˚) x (15, 100, 200 cm)

For each spatial configuration, equate T level at right (better) ear for anechoic speech stimuli

Adaptively vary T level to threshold (50% words correct (Hawley et al., 1999; Shinn-Cunningham et al., 2001).

Runs blocked by condition (random order):
- Left, Right, Binaural x
- Anechoic, Reverberant

Each cond./config. tested 4x per subject

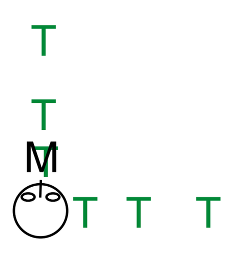
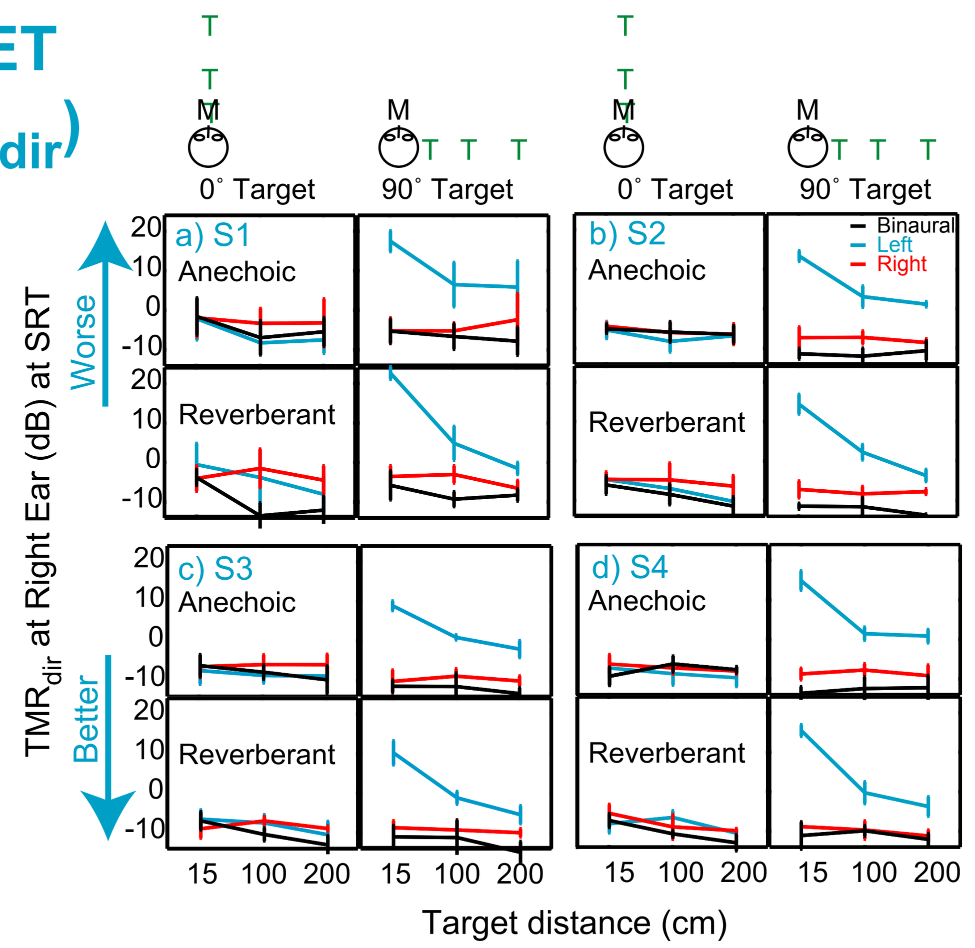


Figure 1: Tested spatial configurations consisted of six target (T) positions and masker (M) at 0˚, 15 cm.

3. DIRECT-SOUND TARGET TO MASKER RATIO (TMR_{dir})

Figure 2. Mean TMR_{dir} in dB RMS at 50% words correct threshold. Error bars show within-subject std. dev. In this analysis, direct-sound TMR is fixed at the right ear to illustrate the better-ear advantage; this analysis ignores any positive contributions of T reverberation.

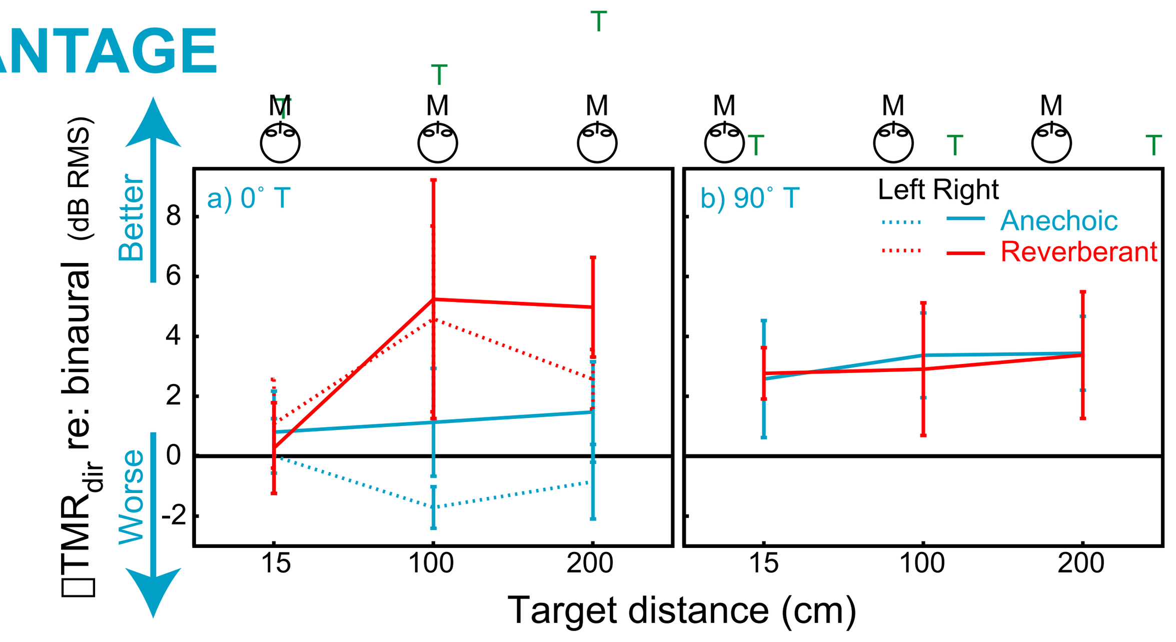
Binaural performance equals or is better than monaural left or right ear performance. For T to the right (right panels), right ear performance is better than left.



4. BINAURAL ADVANTAGE

Figure 3: Mean (x-subject) improvement in TMR_{dir} when listening binaurally (re: better-ear monaural). Error bars show x-subject std. dev.

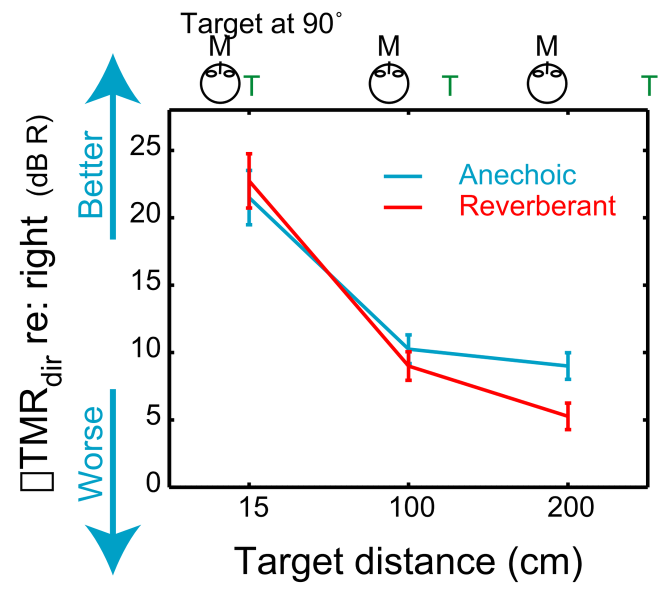
A significant binaural advantage arises when M & T are 1) in different directions (panel b) in anechoic or reverberant space and 2) different distances in reverberant space (panel a).



5. BETTER EAR ADVANTAGE

Figure 4: Mean (x-subject) improvement in TMR_{dir} when listening with right (acoustically-better) ear compared to left (acoustically-worse) ear when T is at 90˚. Error bars show x-subject std. dev.

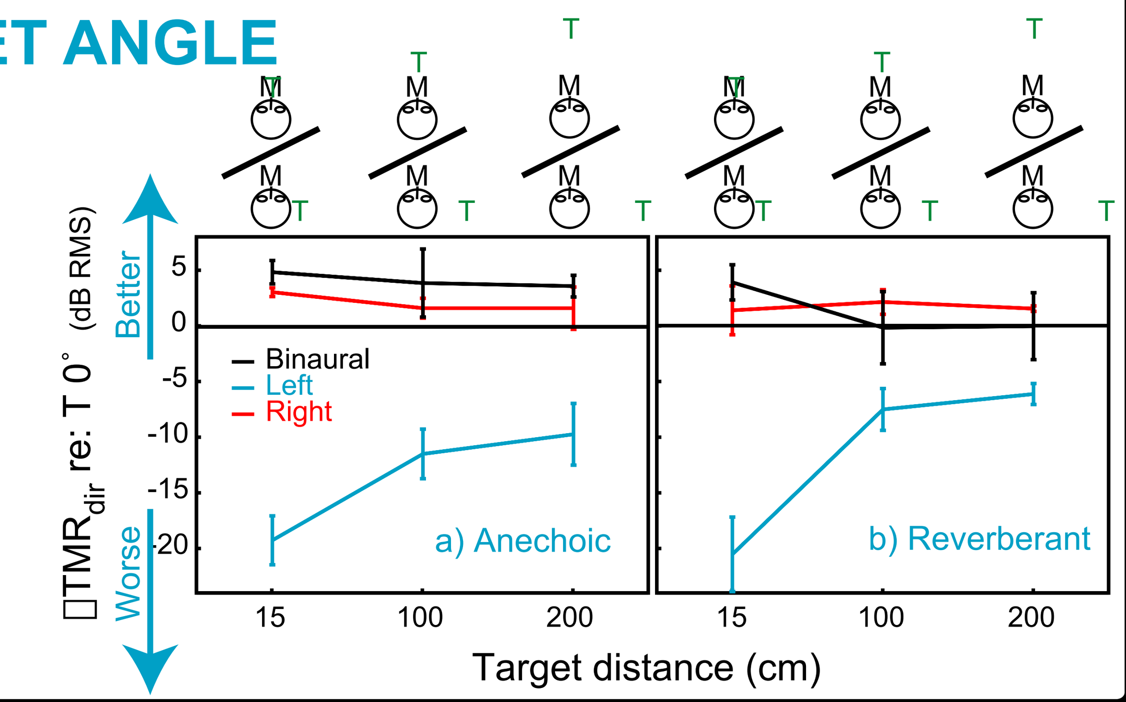
The better ear advantage decreases with distance as the direct sound interaural level difference decreases. The better ear advantage is smaller in reverberant than anechoic space (recall that signals are normalized to direct-sound energy at right ear).



6. EFFECT OF TARGET ANGLE

Figure 5: Mean (x-subject) improvement in TMR_{dir} for 90˚ T (re: T at 0˚). Error bars show x-subject std. dev.

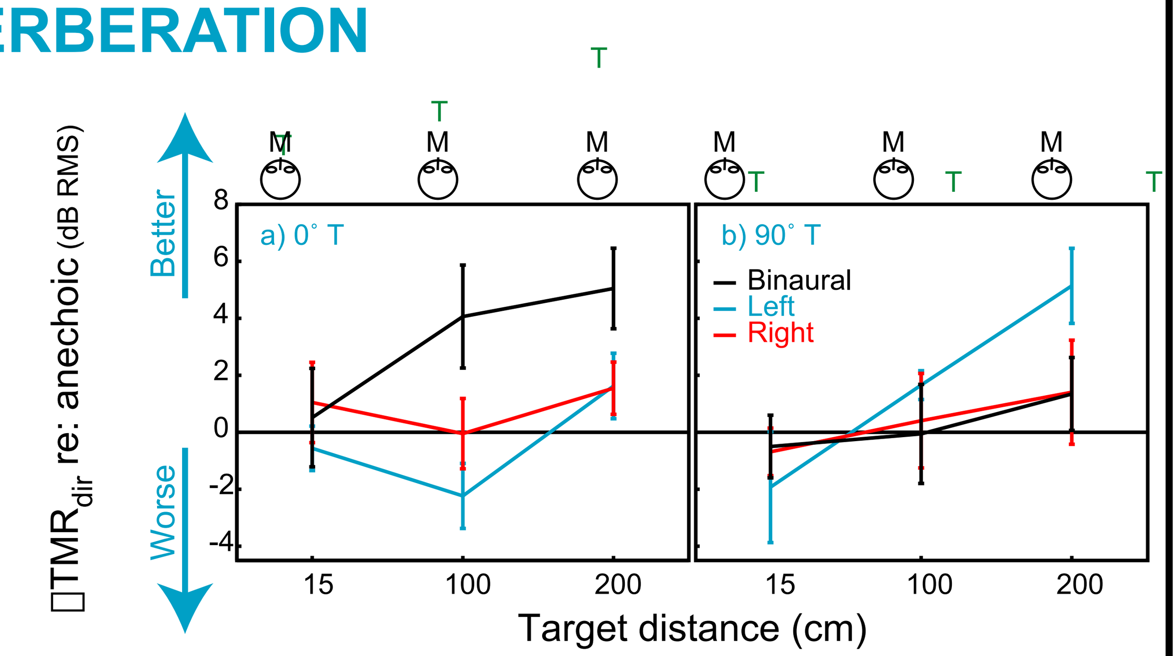
Right ear performance improves and left ear performance degrades for T at 90˚ (re: T at 0˚). **Binaural** performance improves in anechoic space (panel a) when T moves to the side, but only improves in reverberant space (panel b) when T & M are at the same distance.



7. EFFECT OF REVERBERATION

Figure 6: Mean (x-subject) improvement in TMR_{dir} for reverberant conditions (re: anechoic). Error bars show x-subject std. dev.

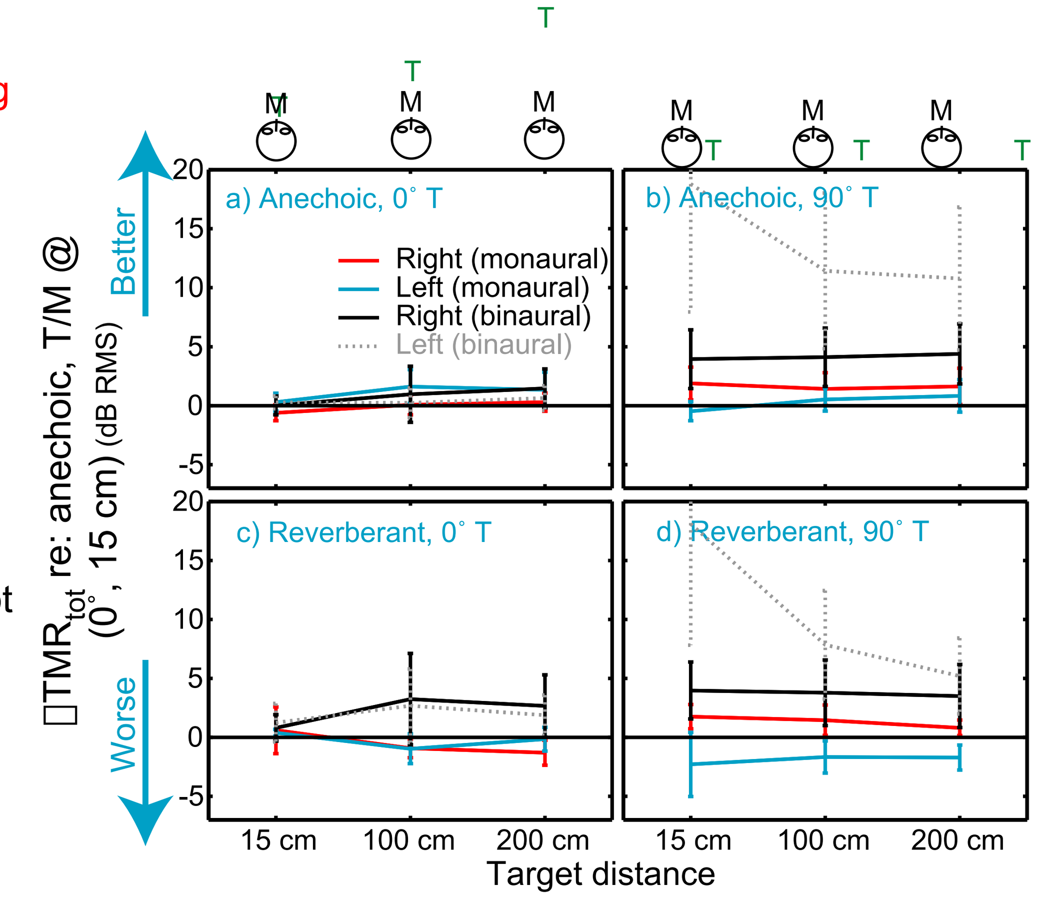
When T & M are both at 0˚ (panel a), reverberation can improve **binaural** performance. When T is at 90˚ (panel b), worse-ear performance improves. Effects of reverberation generally increase with distance.



8. TOTAL TARGET TO MASKER ENERGY RATIO (TMR_{tot})

Figure 7: Mean (x-subject) improvement in TMR_{tot} (total-energy TMR including reverberant energy) re: anechoic listening with T & M at (0˚, 15 cm). This analysis treats direct and reverberant energy equally (TMR_{tot}=TMR_{dir} in anechoic cases). Error bars show x-subject std. dev. [Note in panels b & d, the left ear binaural signal is essentially unintelligible if played alone]

Distance has little effect on threshold TMR_{tot}. In anechoic conditions, there is a **binaural** gain when T is at 90˚ (panel b). In reverberant conditions (panels c, d), there is a **binaural** gain if T and M are not at the same location. When T is at 90˚ in reverberant conditions (panel d), TMR_{dir} must be greater for the left ear condition, suggesting that reverberant energy improves speech intelligibility less than direct-sound energy.



9. CONCLUSIONS

OVERALL

The largest effects of spatial configuration of T and M are simple acoustic effects, i.e., changes in TMR_{tot} (and TMR_{dir}) at the ears.

Binaural advantages of 3-5 dB arise if T and M directions differ by 90˚ in anechoic and reverberant conditions.

Moderate reverberation does not degrade any binaural advantage, and can even cause one.

ANECHOIC CONDITIONS

Separation of T and M in distance can change TMR_{dir}, but otherwise has little influence on speech intelligibility.

The acoustic better-ear advantage can be greater than 20 dB for some configurations.

REVERBERANT CONDITIONS

Reverberation can lead to a binaural gain when T & M are in the same direction but different distances, probably by decorrelating T.

Reverberation can increase intelligibility by increasing the effective TMR (TMR_{tot}).

Increasing TMR_{tot} improves performance, but less than increasing TMR_{dir}.

Reverberation tends to decrease large better-ear advantages.

FUTURE WORK

Analysis of speech transmission index and effect of reverberation on speech waveform modulations (e.g., Houtgast & Steeneken, 1985).

Extensions of binaural models (e.g., Zurek, 1988).

10. REFERENCES

Bradley, JS, RD Reich & SG Norcross (1999). "On the combined effects of signal-to-noise ratio and room acoustics on speech intelligibility." J Acoust Soc Am, 106(4): 1820-1828.

Bronkhorst, AW (2000). "The cocktail party phenomenon: A review of research on speech intelligibility in multiple-talker conditions." Acustica 86: 117-128.

Drullman, R & AW Bronkhorst (2000). "Multichannel speech intelligibility and talker recognition using monaural, binaural, and three-dimensional auditory presentation." J Acoust Soc Am, 107: 2224-2235.

Drullman, R, JM Festen & R Plomp (1994). "Effect of reducing slow temporal modulations on speech reception." J Acoust Soc Am, 95(5): 2670-2680.

Hawley, ML, RY Litovsky & HS Colburn (1999). "Speech intelligibility and localization in a multi-source environment."J Acoust Soc Am, 105: 3436-3448.

Houtgast, T & HJM Steeneken (1985). "A review of the MTF concept in room acoustics and its use for estimating speech intelligibility in auditoria." J Acoust Soc Am, 77(3): 1069-1077.

Payton, KL, RM Uchanski & LD Braida (1994). "Intelligibility of conversational and clear speech in noise and reverberation for listeners with normal and impaired hearing." J Acoust Soc Am, 95(3): 1581-1592.

Shinn-Cunningham, BG, J Schickler, N Kopco and RY Litovsky (2001). "Spatial unmasking of nearby speech sources in a simulated anechoic environment." J Acoust Soc Am, 110(2): 1118-1129.

Shinn-Cunningham, BG, S Santarelli and N Kopco (2000). "Tori of confusion: Binaural localization cues for sources within reach of a listener." J Acoust Soc Am, 107(3): 1627-1636.

Zurek, PM (1993). Binaural advantages and directional effects in speech intelligibility. Acoustical Factors Affecting Hearing Aid Performance. G Studebaker and I Hochberg. Boston, MA, College-Hill Press.