Predicting speech intelligibility in noisy rooms.

John F. Culling[1], Mathieu Lavandier[2] and Sam Jelfs[3]

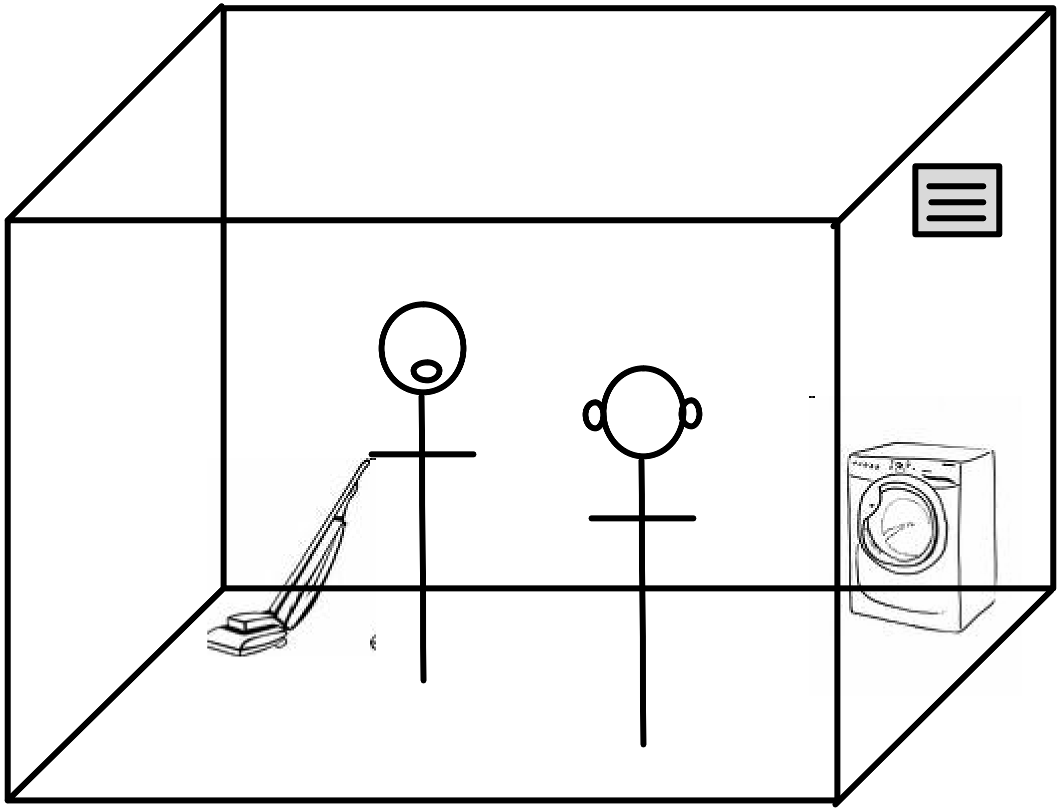
[1]School of Psychology, Cardiff University, Tower Building, Park Place, Cardiff, CF10 3YG. U.K.

[2] Université de Lyon, Ecole Nationale des Travaux Publics de l'Etat, Département Génie Civil et Bâtiment, Unité CNRS 1652, Rue M. Audin, 69518 Vaulx-en-Velin Cedex, France

[3] Welsh School of Architecture, Bute Building, King Edward VII Avenue, Cardiff, CF10 3NB, U.K.

Acknowledgement:
Work supported by UK EPSRC

Speech intelligibility in noise



Two factors: Binaural unmasking and better-ear listening

A general equation for generating E-C predictions.

$$BMLD = \frac{k - \cos(f_s - f_n)}{k - r_n}$$

$$k = (1 + s_e^2)\exp(w_0^2 s_d^2)$$

$$s_e^2 = 0.25$$

$$s_d^2 = 105ms$$

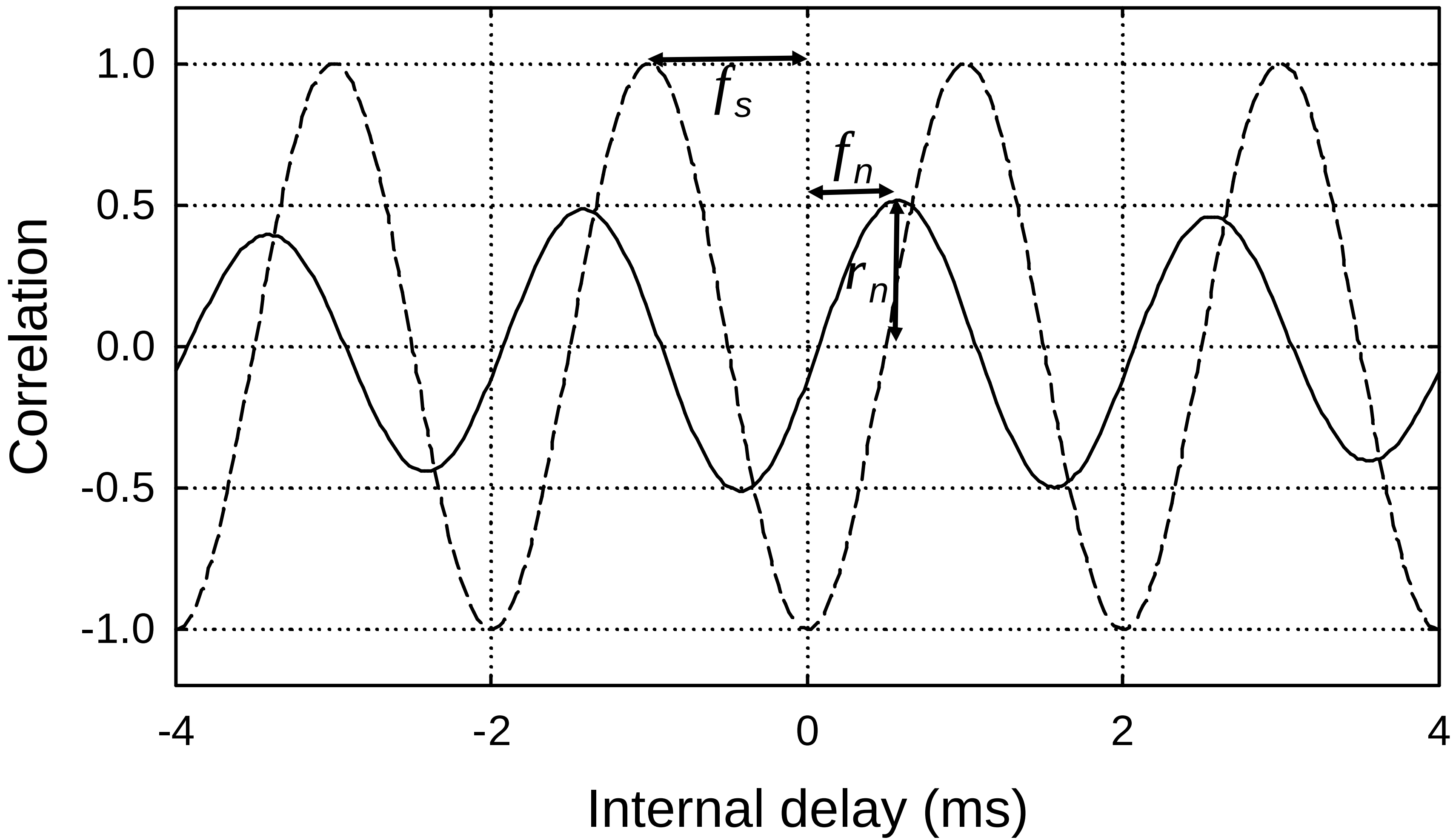
$f_s =$ interaural phase of signal

$f_n =$ interaural phase of noise

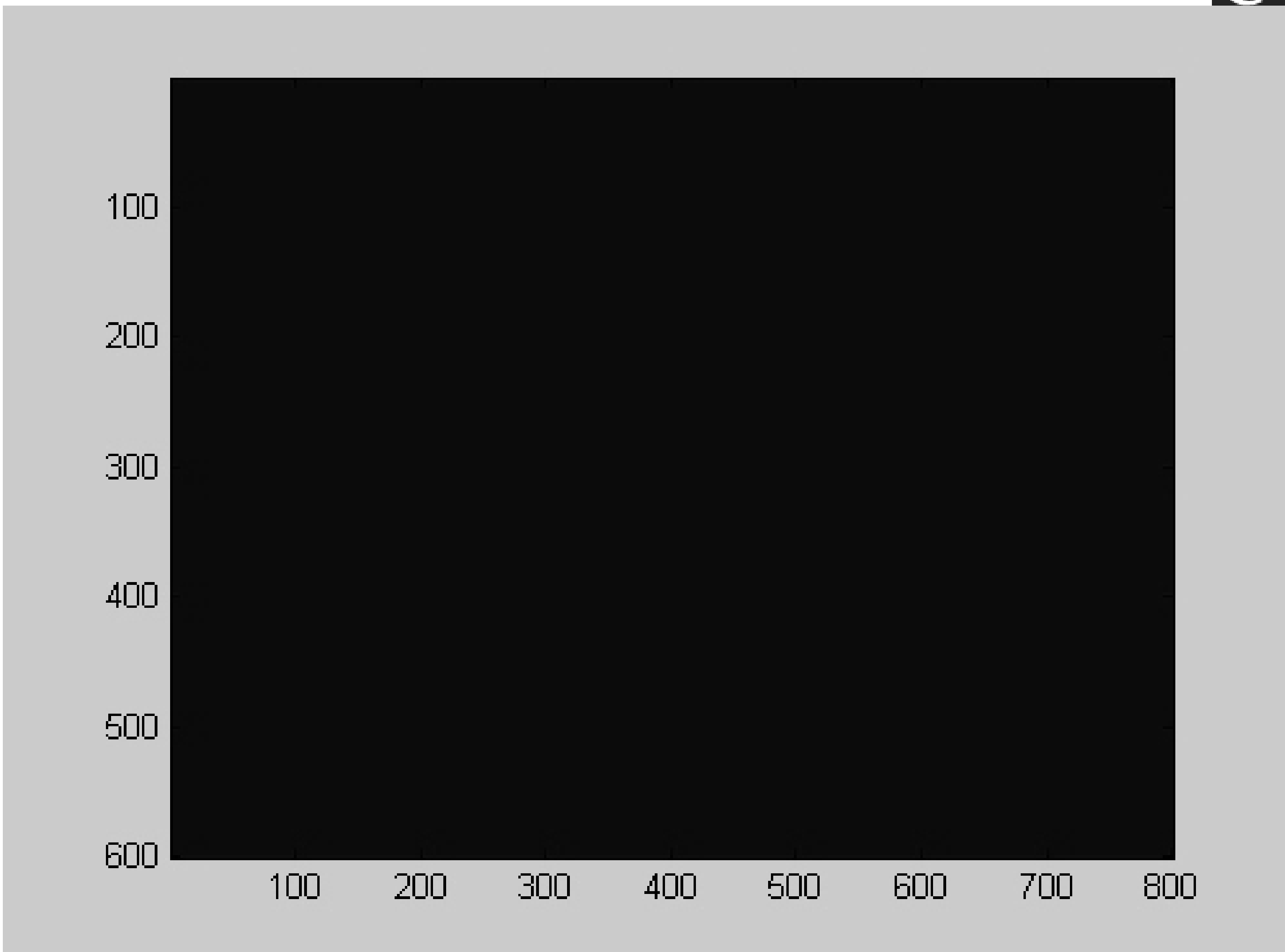
$r_n =$ interaural coherence of noise

This single equation can replace equations 55-66 from Durlach (1972), but requires the determination of f_s, f_n and r_n from cross-correlations.

Extraction of parameters from target (dashed) and masker (solid) cross-correlations

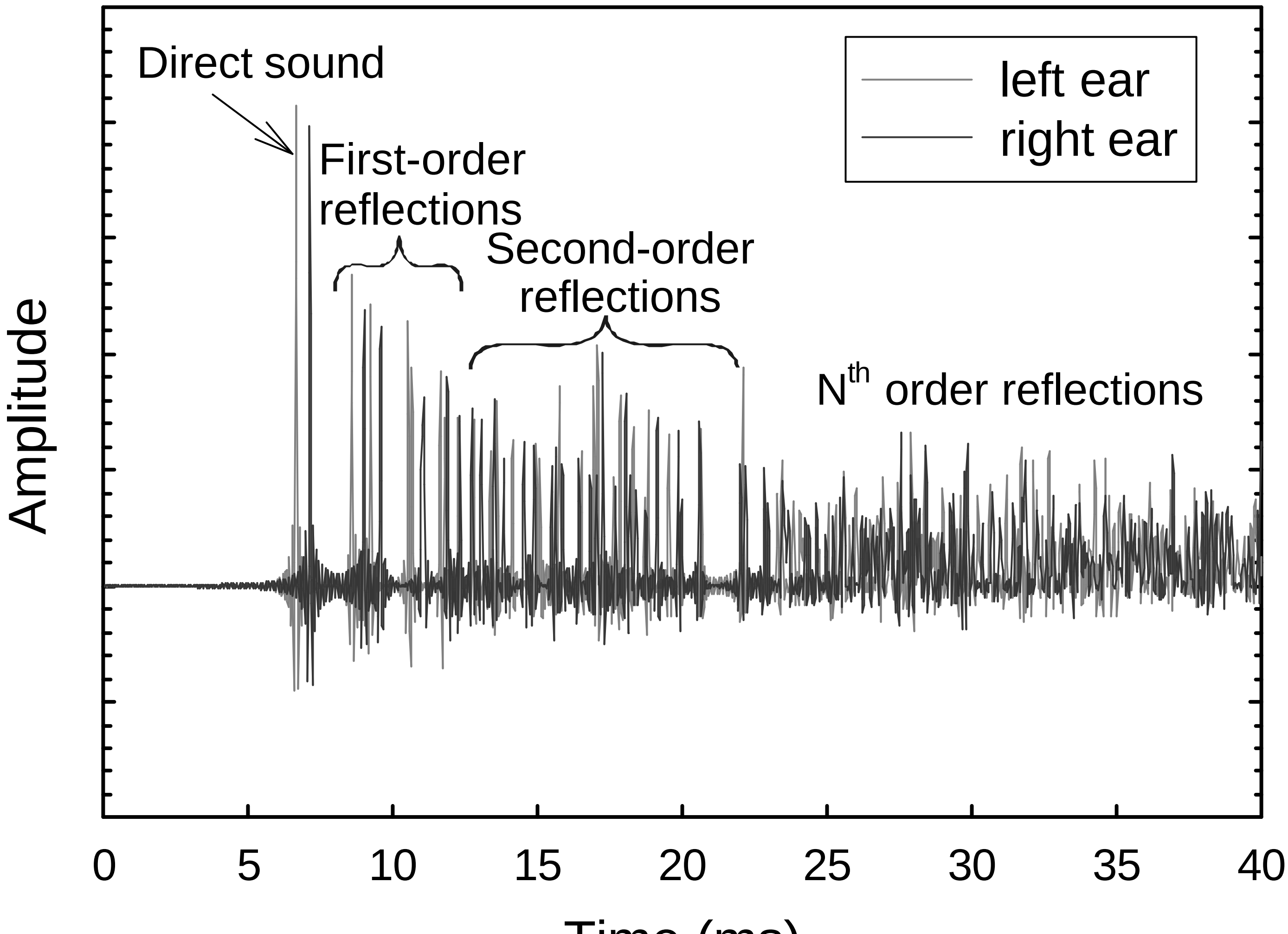


Propagation of sound in a room…



Video created by Sam Jelfs

A Binaural Room Impulse Response (BRIR)



Experiment probing the effects of reverberation on masker cancellation.

r_n was manipulated by varying:-

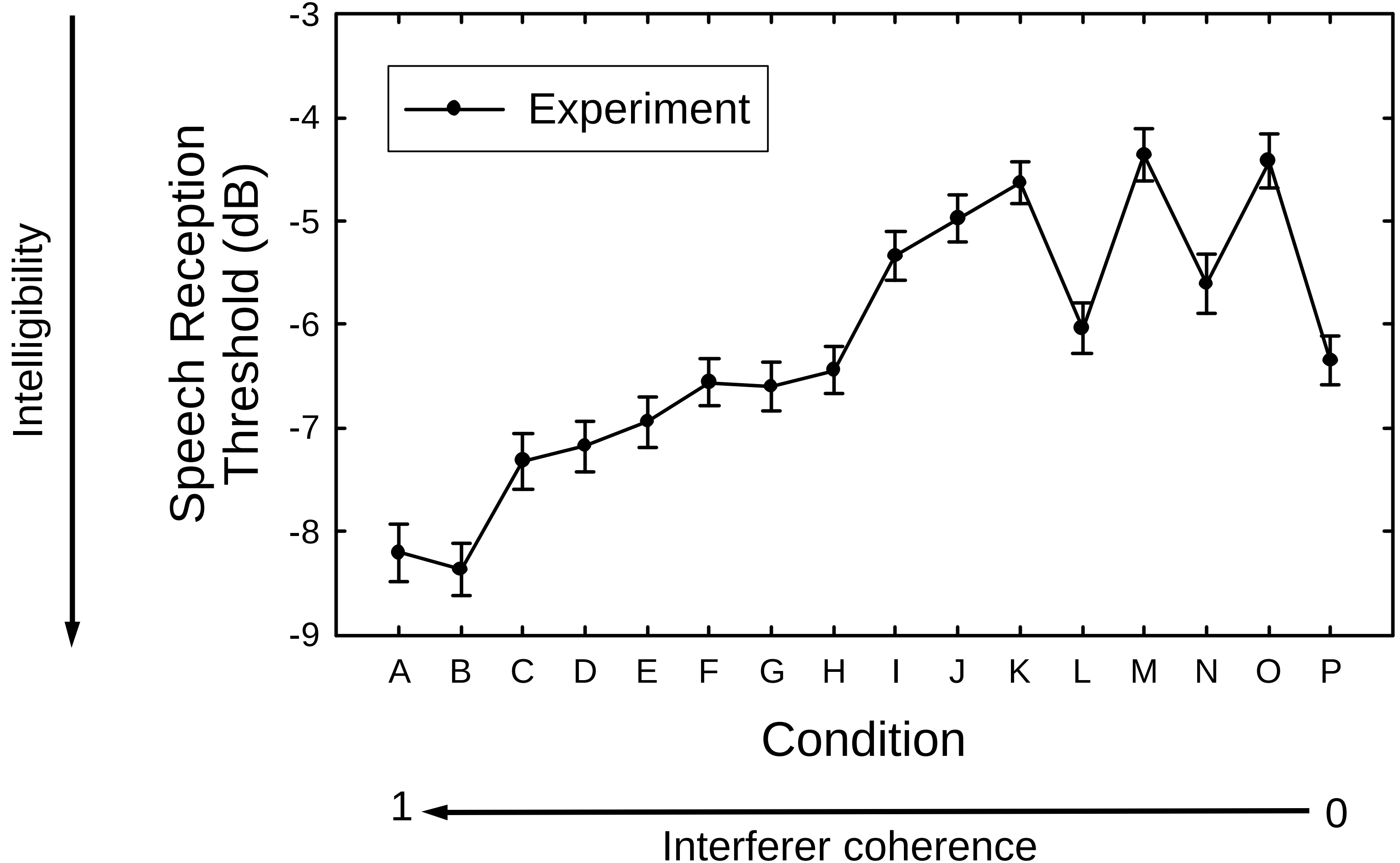
- Size of the room
- Masker source distance
- Masker source absorption coefficients
- Symmetry of room configuration

Other factors were kept constant:-

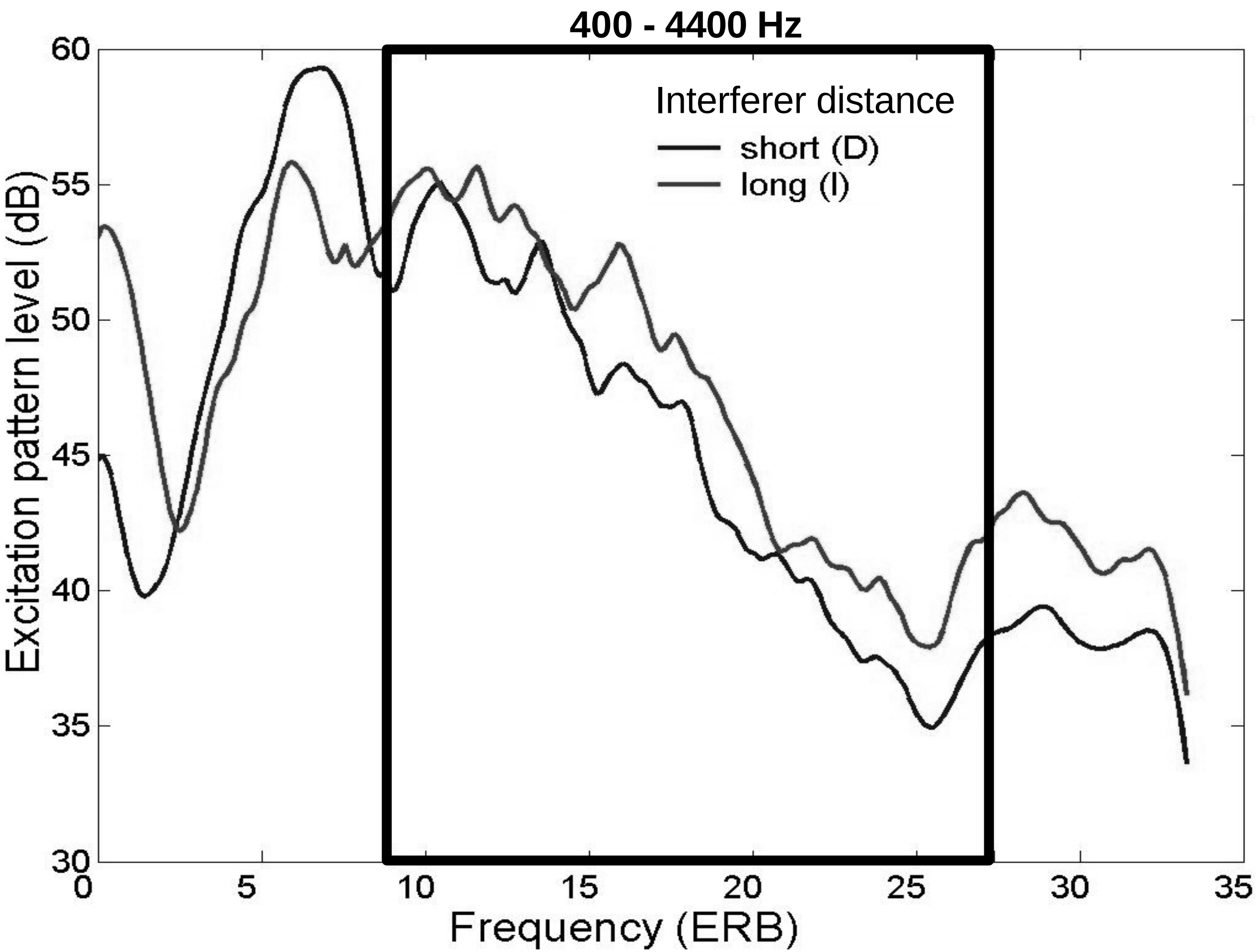
- Target speech always anechoic
- RMS level at each ear normalised
- Spatial separation between target and masker

Lavandier and Culling (2010) JASA ***127****, 387-399*

Experimental Results



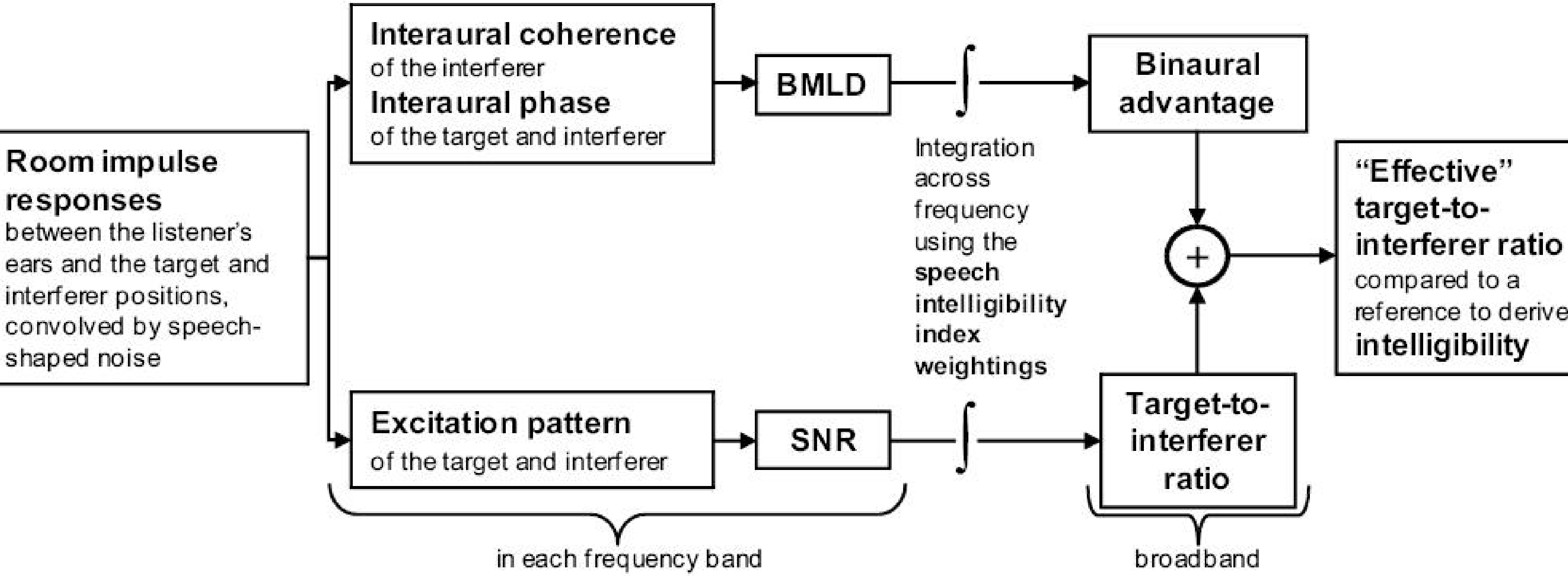
Room colouration



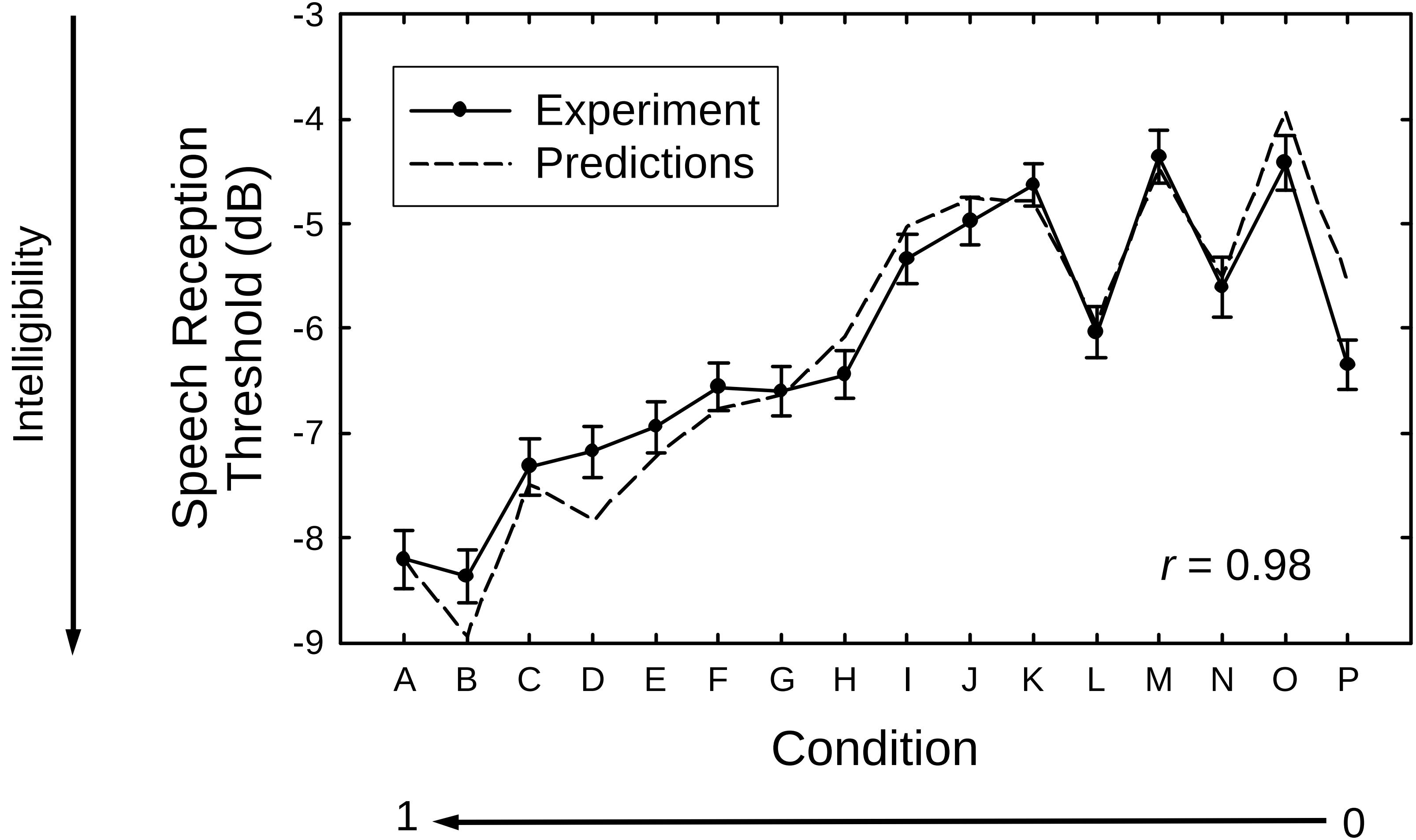
Room coloration influences masking/intelligibility

Model structure

Binaural unmasking (differences in ITD)



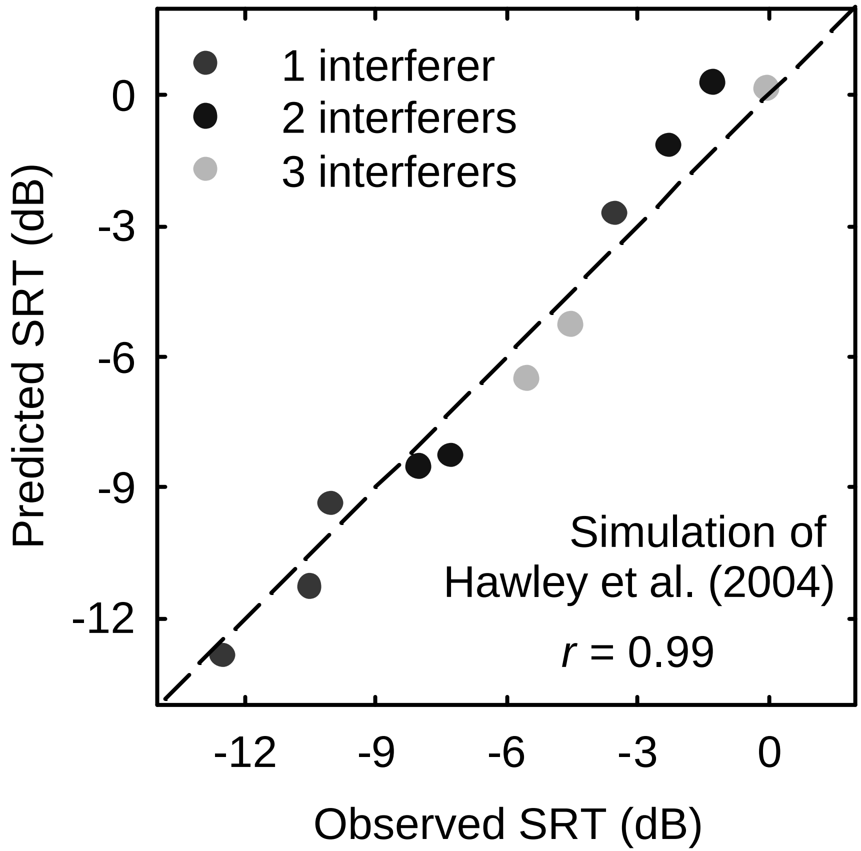
Better-ear listening (head shadow, room colouration, source distances).



*Lavandier and Culling (2010), JASA, **127**, 387-399*

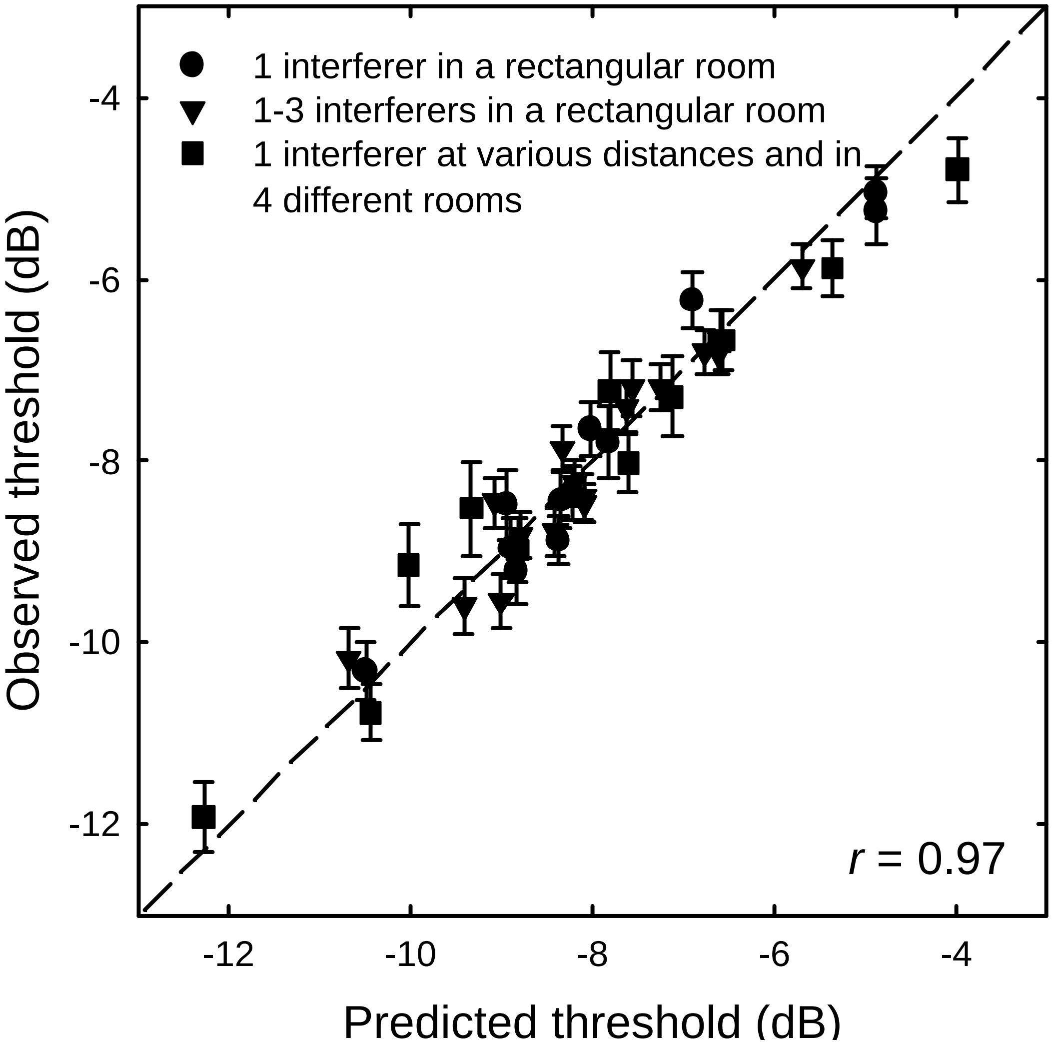
Modelling of data from the literature with head shadow effects and multiple interferers.

Hawley et al. (2004) presented 1, 2 or 3 speech-shaped noise interferferers in a variety of spatial configurations in a simulated anechoic space.



Modelling multiple interferers with natural reverberation and a realistic directional source.

Speech shaped noise emitted by a B&K head and torso simulator and recorded from a KEMAR acoustic mannequin (unpublished data).



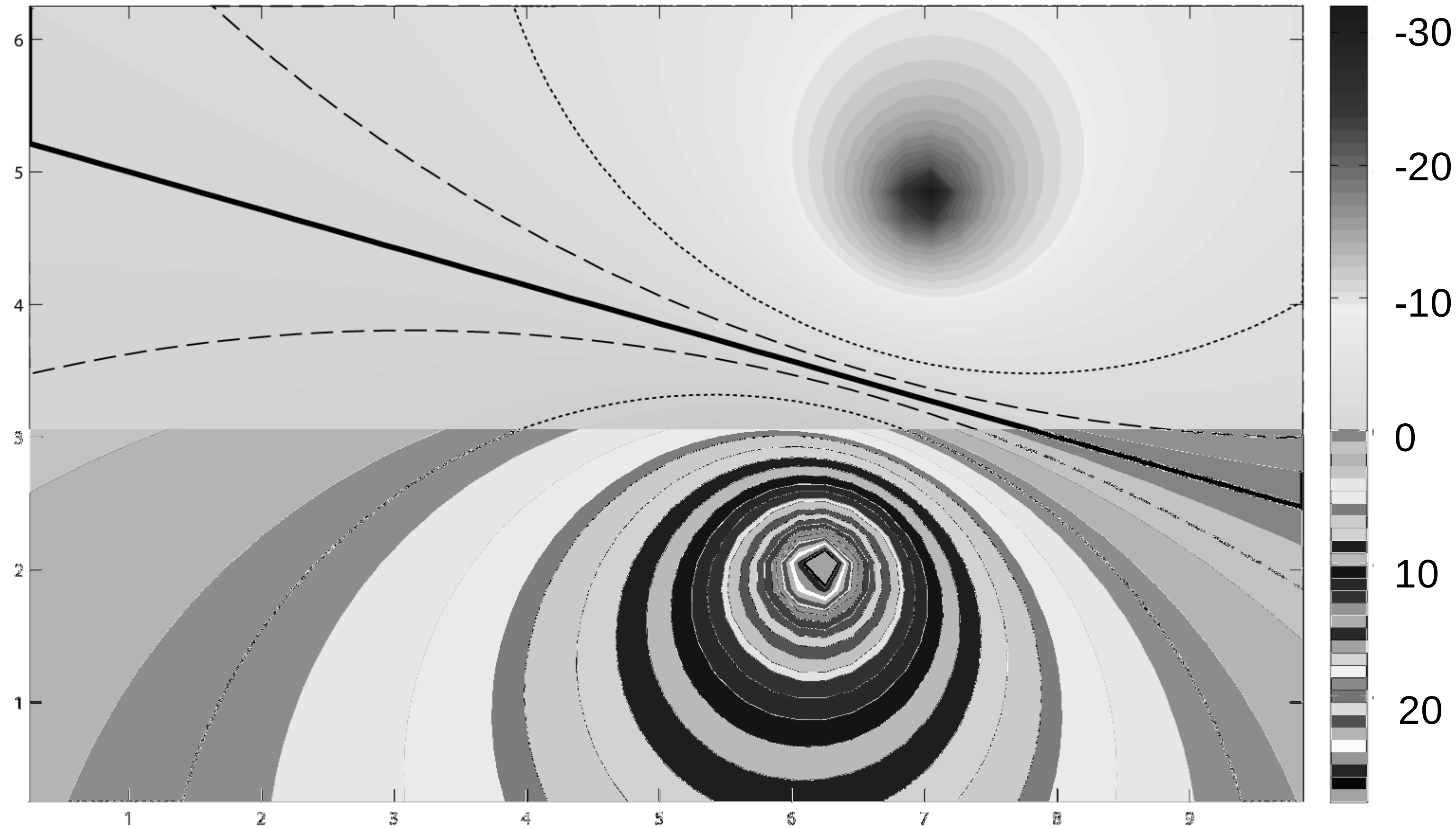
BRIRs supplied by Tony Watkins' group at Reading.

A computational tweak…

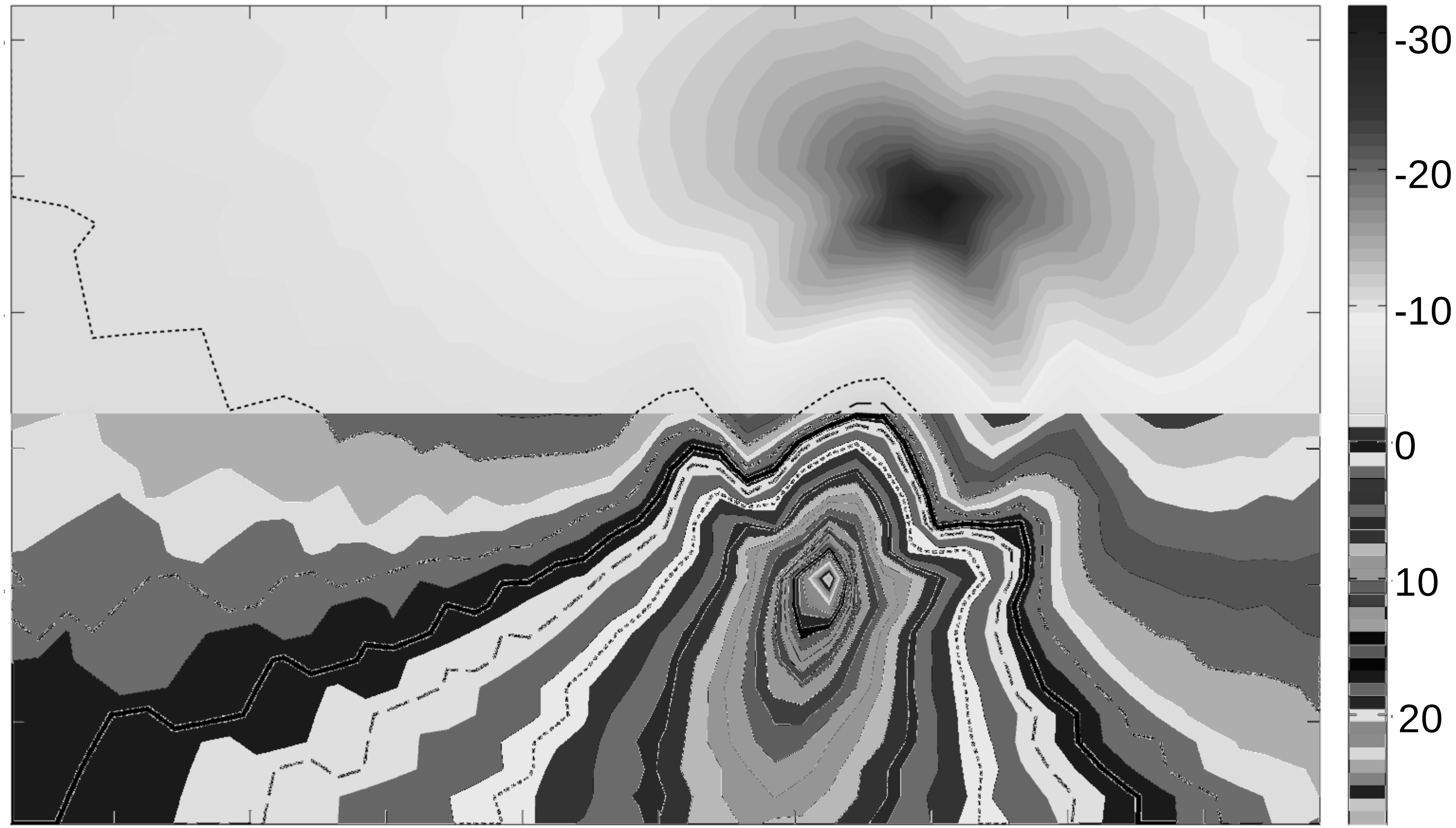
For the case of BRIRs considerable computational efficiencies can be made by directly cross-correlating the BRIRs, rather than convolving them with noise first.

Armed with this greater efficiency, we can make intelligibility maps…

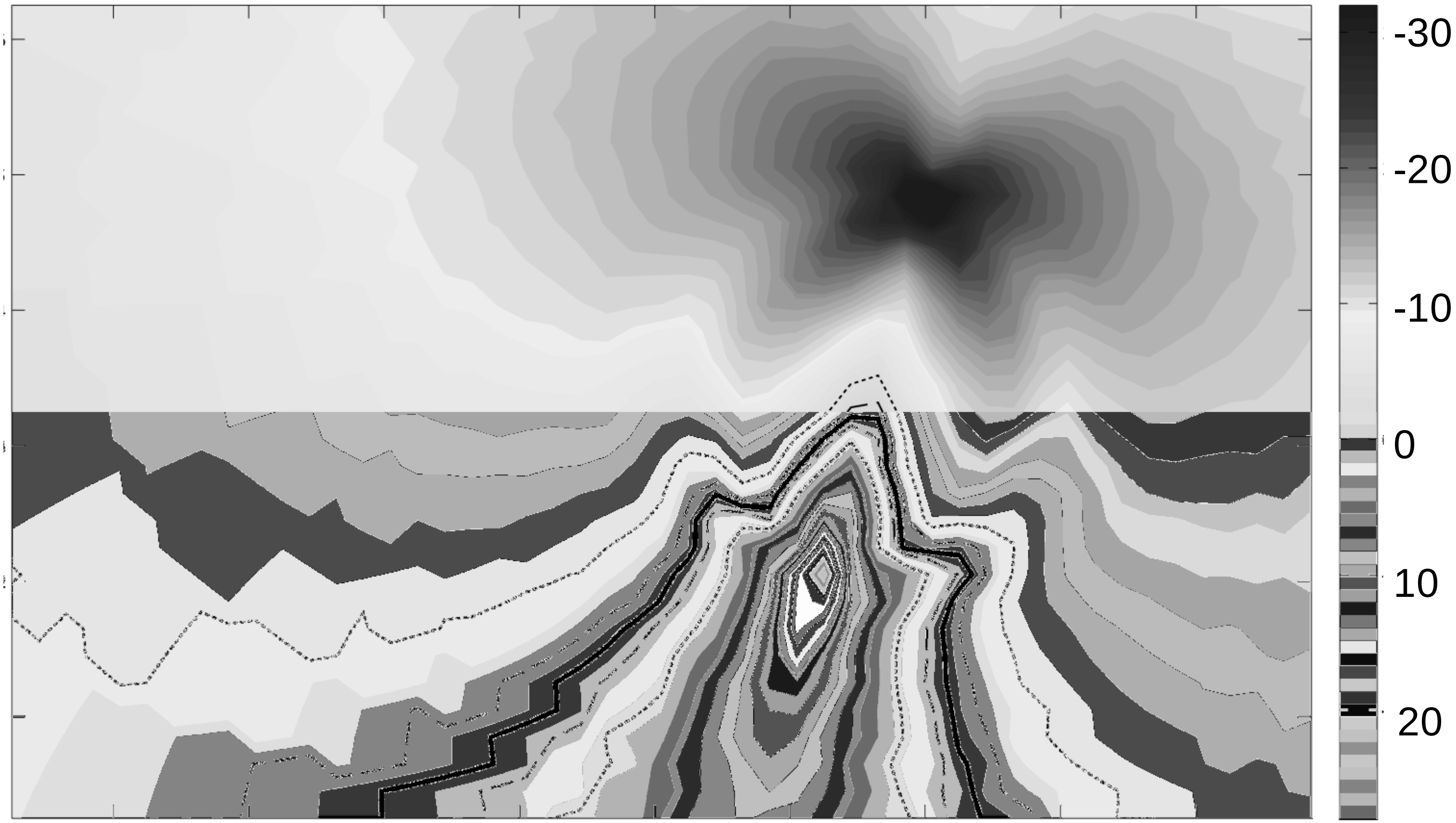
Intelligibility map of an anechoic room: an omnidirectional microphone.



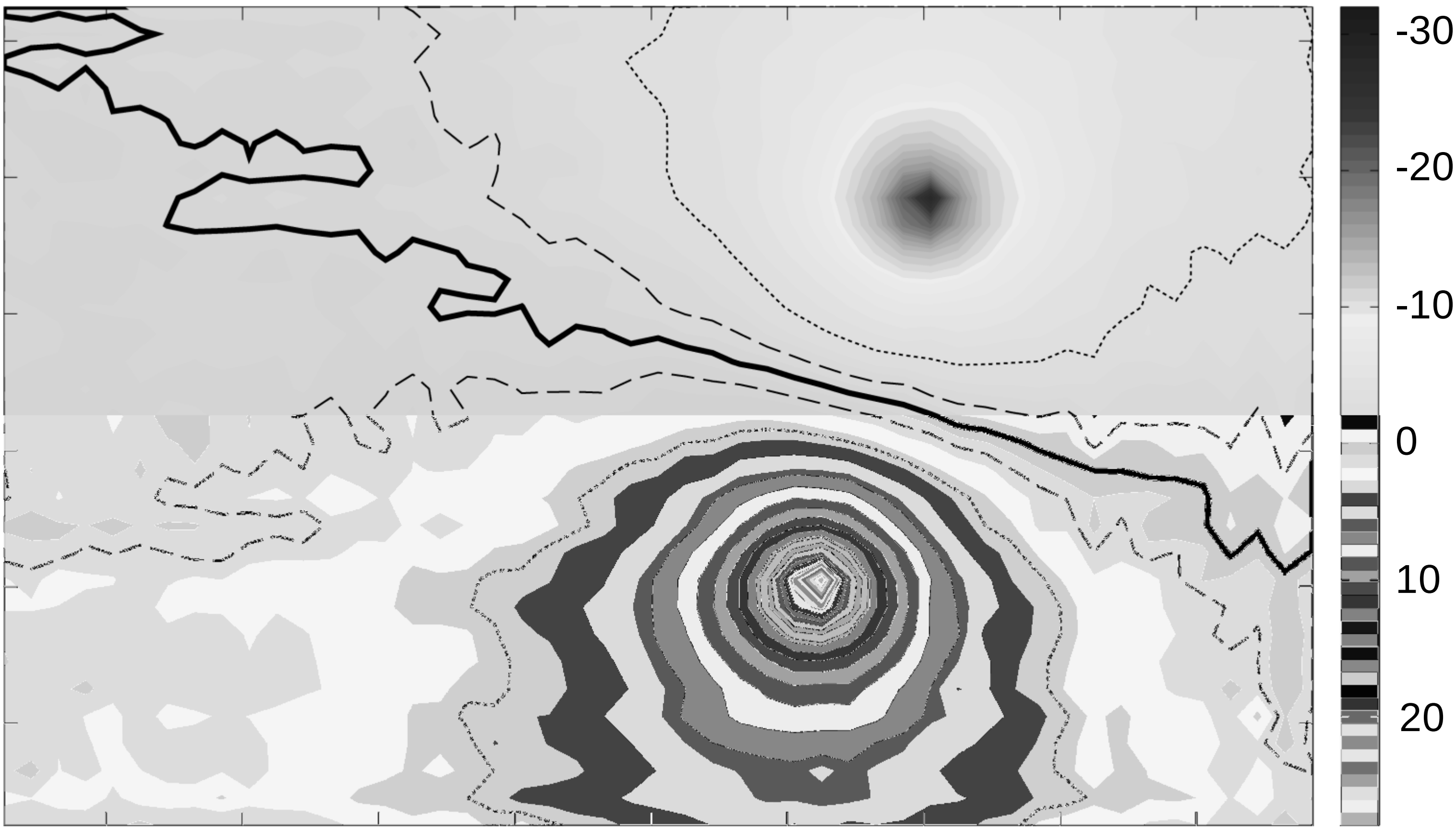
Intelligibility map of an anechoic room: better-ear listening only.



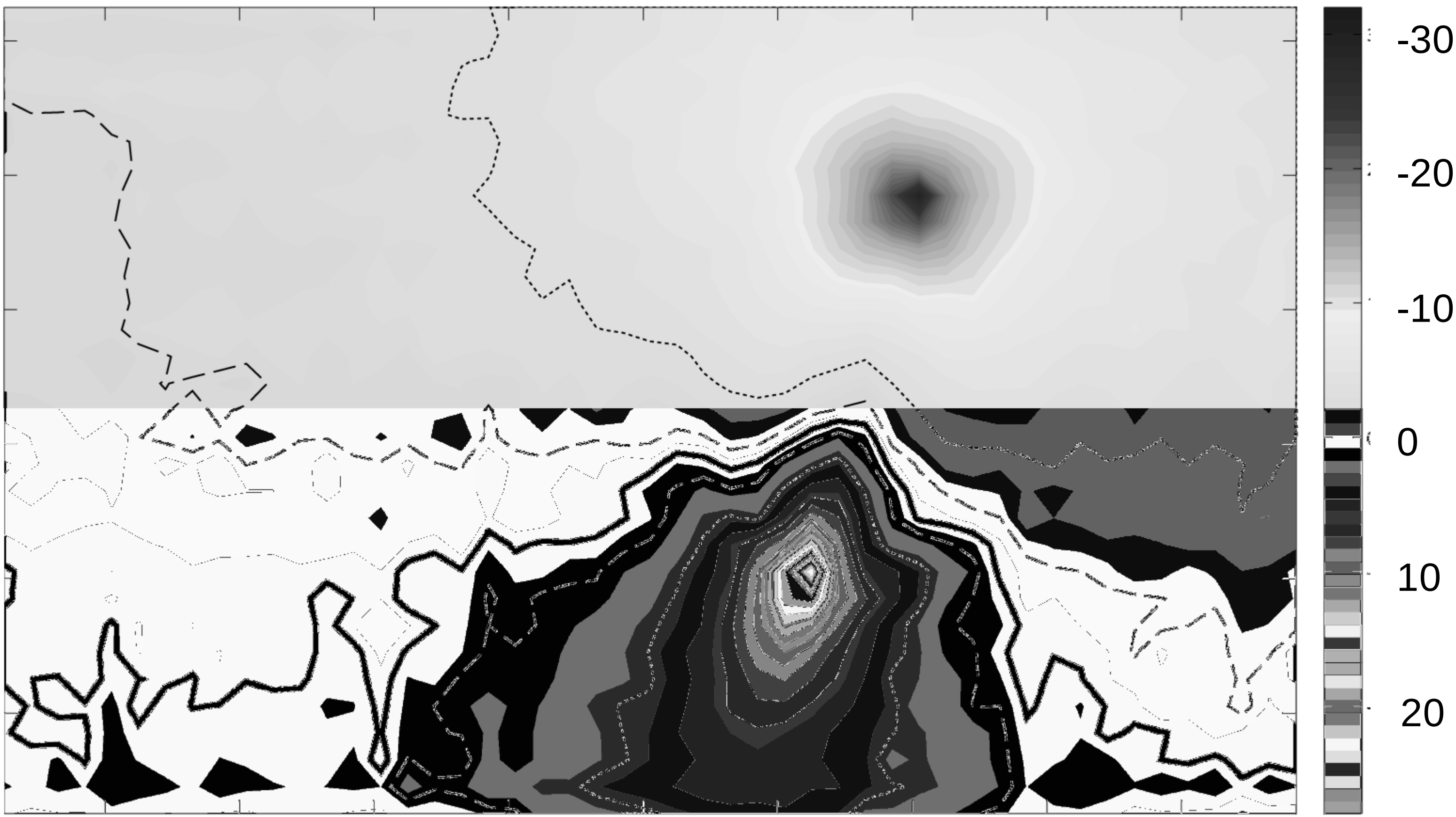
Intelligibility map of an anechoic room: better-ear listening + unmaksing



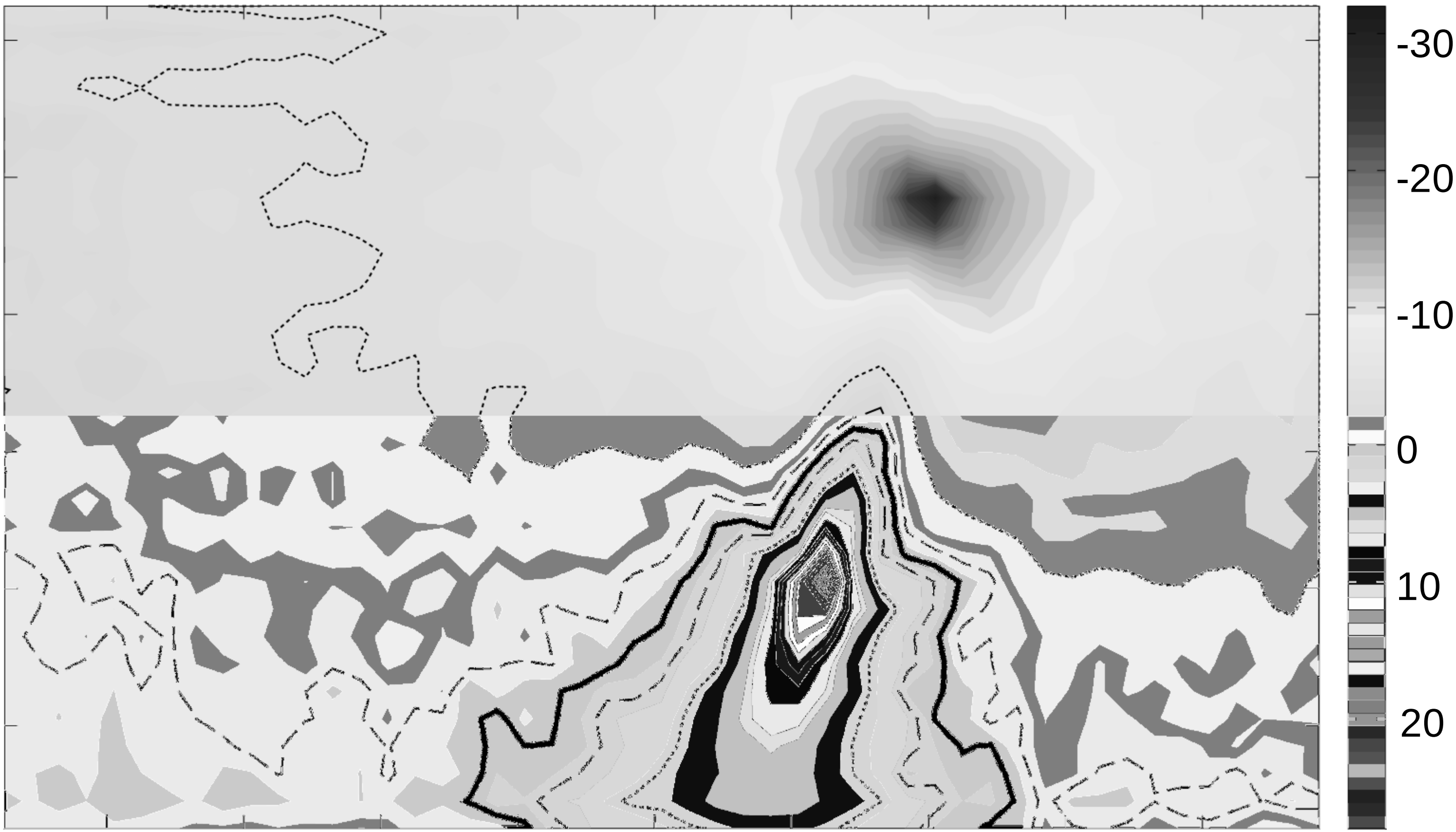
Intelligibility map of a reverberant room: an omnidirectional microphone



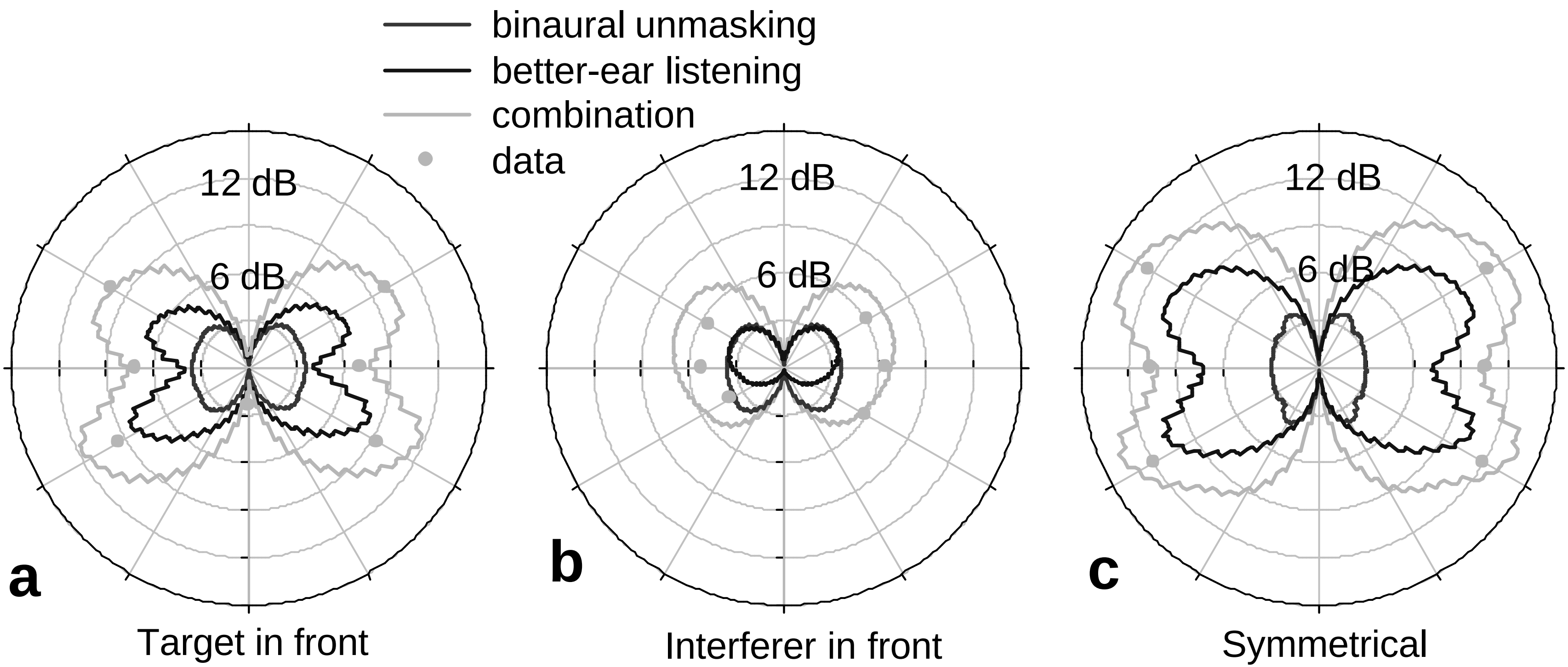
Intelligibility map of a reverberant room: better-ear listening only



Intelligibility map of a reverberant room: better-ear listening + unmasking

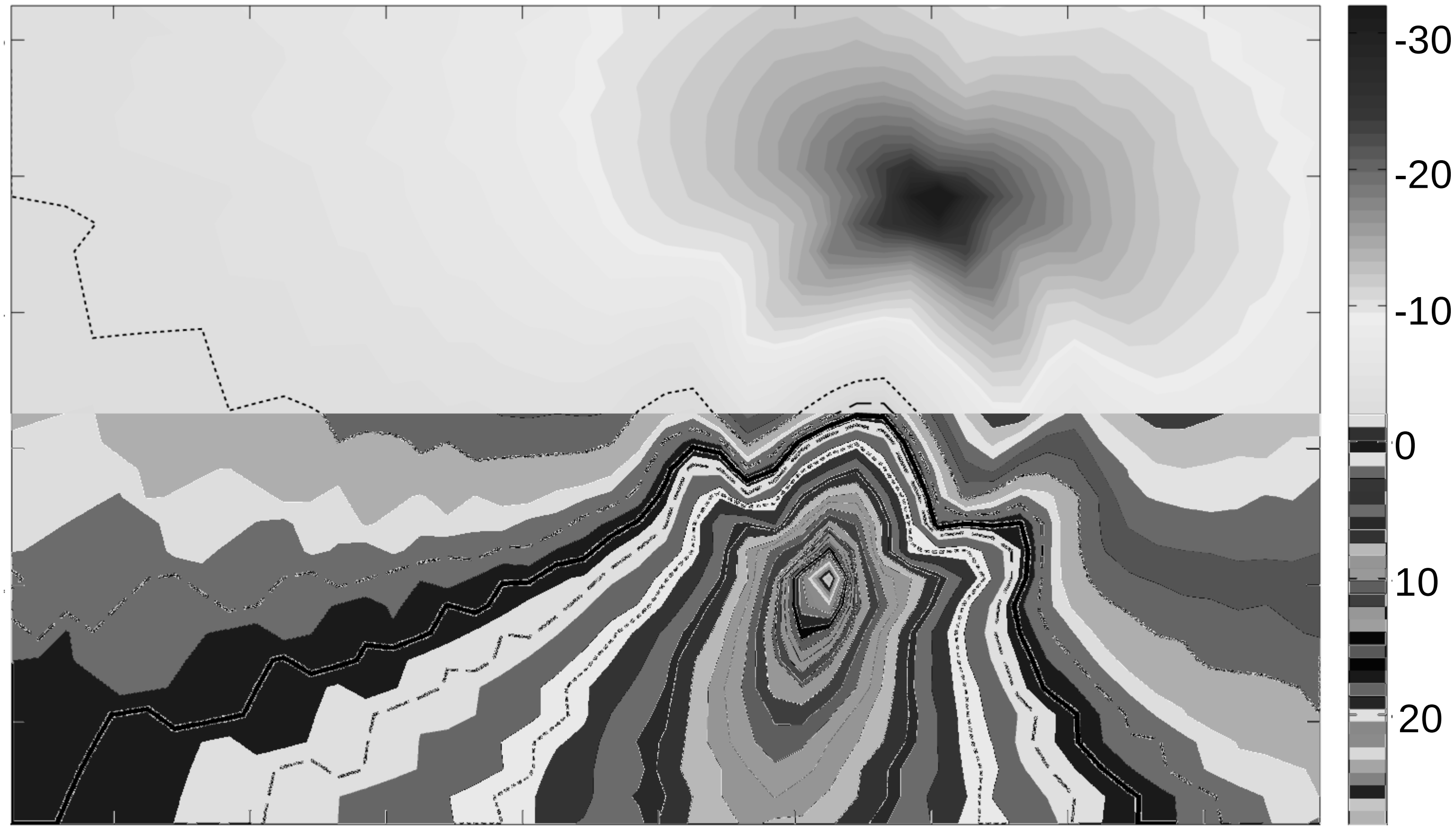


Effects of head orientation.

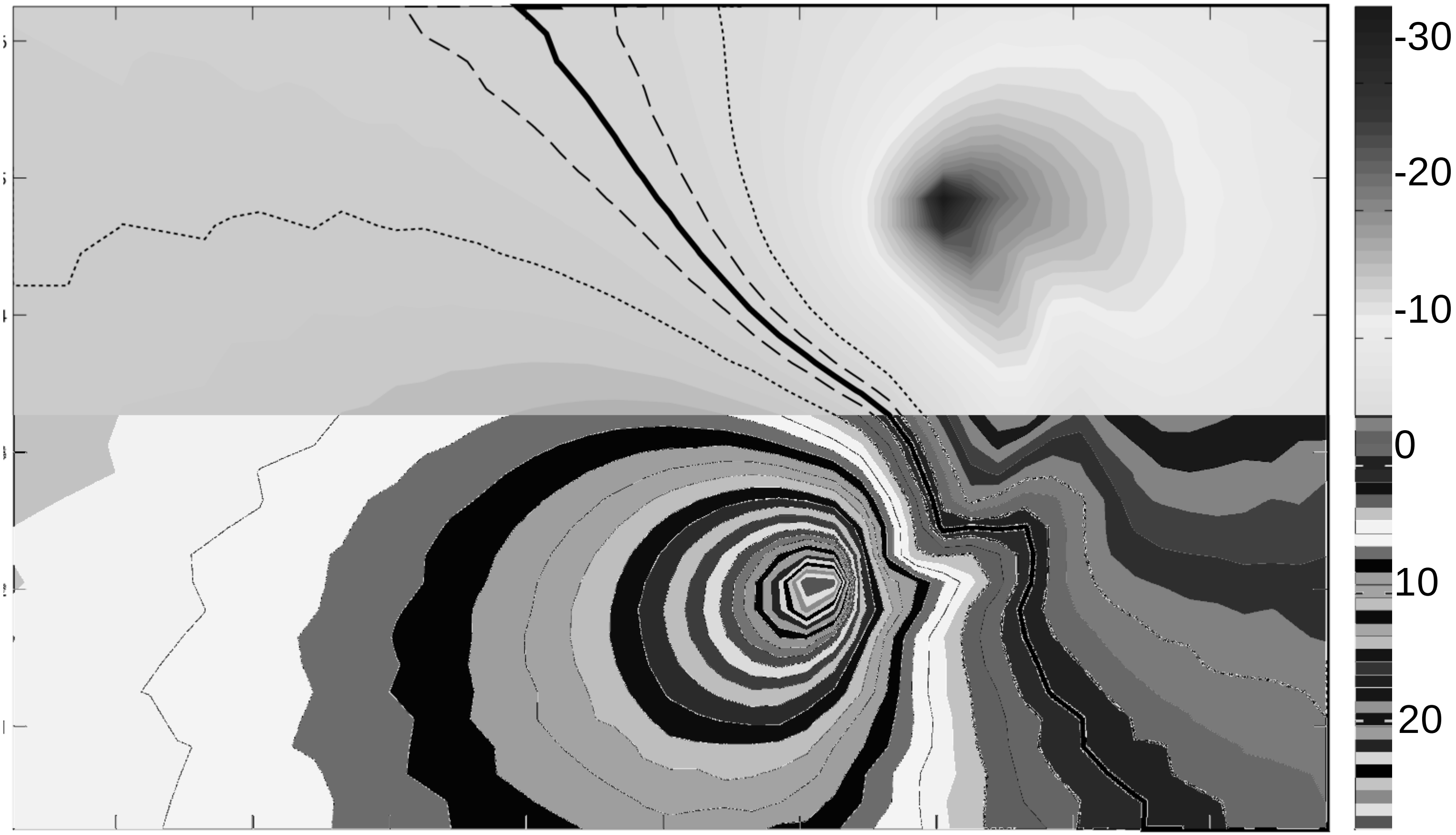


Empirical data supplied by Juan-Pablo Ramirez of Deutsche Telekom.

Intelligibility map of an anechoic room: better-ear listening only.



Intelligibility map of an anechoic room: right ear only.



Conclusions

The new E-C equation is more general than the set of equations from Durlach (1972).

The equation is well-suited to conditions whose noise has convolutive distortion such as room reverberation.

The equation can be used to accurately and efficiently predict speech intelligibility in noisy rooms.

The effect of headshadow is strongly dependent on head orientation.