HAL open science

# How to optimise your speech intelligibility in a noisy restaurant

John Culling, Jacques Grange, Barry Bardsley, Eleanor Ainge

**To cite this version:**

John Culling, Jacques Grange, Barry Bardsley, Eleanor Ainge. How to optimise your speech intelligibility in a noisy restaurant. Forum Acusticum, Dec 2020, Lyon, France. pp.1277-1280, 10.48465/fa.2020.0193 . hal-03234189

**HAL Id: hal-03234189**

**https://hal.science/hal-03234189v1**

Submitted on 26 May 2021

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# How to optimise your speech intelligibility in a noisy restaurant

**John F. Culling  Jacques A. Grange  Barry Bardsley  Eleanor G. Ainge**
School of Psychology, Cardiff University, Park Place, Cardiff, U.K.

## ABSTRACT

A model of spatial release from masking in reverberant environments has proved successful in predicting a wide range of empirical data, particularly for noise interferers or where there are multiple interfering voices. Here, we show how the model can be used to successfully predict optimal behaviour in such environments. Head orientation of around 30° away from the target speaker was predicted, and then shown empirically to improve speech intelligibility for normally hearing listeners, hearing-impaired listeners and cochlear implant users. This degree of head orientation was also shown to be compatible with lip-reading and robust to reverberation. However, location within a room, even one uniformly distributed with interfering sound sources, was predicted to have a strong influence on the effective signal-to-noise ratio and on the benefits of head orientation and directional microphones. Certain locations were predicted to bring all these benefits together and others to largely confound them. These differences were confirmed empirically for the benefits of signal-to-noise ratio and head orientation using virtual acoustics.

## 1. INTRODUCTION

The model of spatial release from masking in reverberant environments examined here [1,2] is based on a combination of better-ear listening and binaural unmasking. Within a range of frequency channels, the signal-to-noise ratio (SNR) at the acoustically better ear is selected. An effective SNR value is then derived by adding the benefit of binaural unmasking predicted for that frequency from Equalization-Cancellation theory [3,4]. The effective SNRs for each channel are then weighted according to the frequency importance function of the speech intelligibility index [5], to give an overall effective SNR for speech.

Here we report a combination of published [6,7] and unpublished work showing accurate predictions of the effects of behavioural adaptations to a listening environment; the choice of seat and the orientation of the head.

## 2. HEAD ORIENTATION

In [6] & [7] we have previously shown that the model predicts the effect of head orientation using a simulation of a restaurant. The simulation was based upon measured binaural room impulse responses from a real restaurant (Mezzaluna in Cardiff) collected using a head and torso simulator with 8 simulated interfering sources in the room.

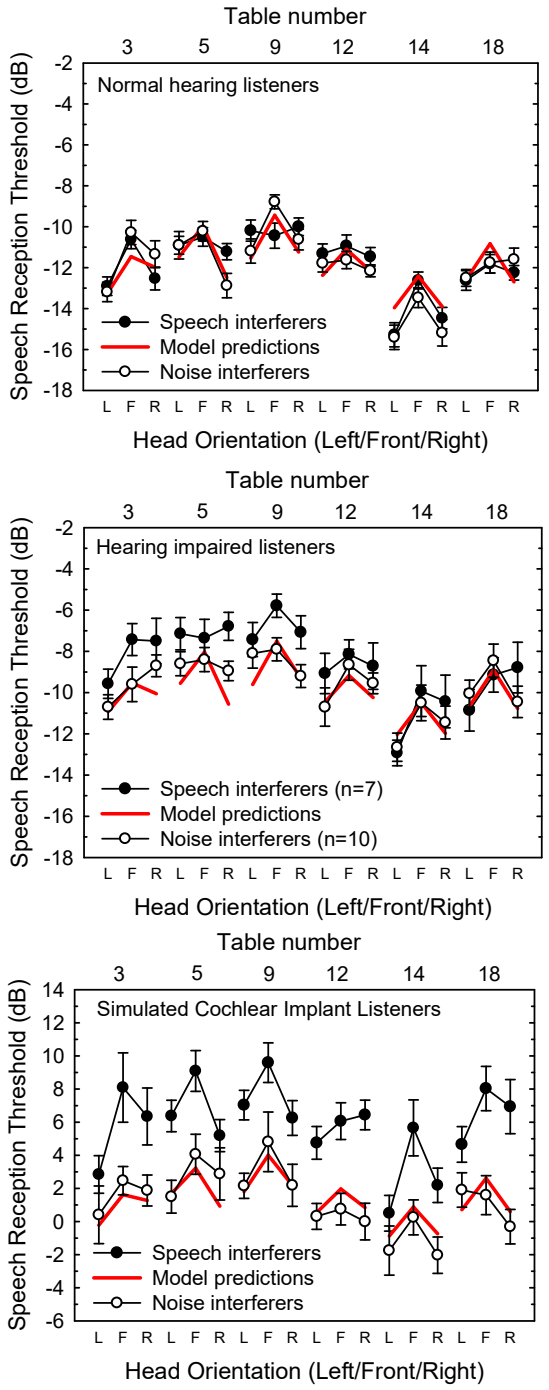


**Figure 1.** Speech reception threshold for different head orientations in 3 groups of listeners in a simulated restaurant. Error bars are one standard error of the mean. Model predictions from [2].

Speech reception thresholds (SRTs) showed a clear beneficial effect of head orientation ±30° away from the target voice, which was predicted by the model, but the model does not predict the absolute level of performance. Since SRTs differed between the three participant groups, the model predictions have been fitted to the overall level of performance against speech-shaped noise for each group. The results show that the benefit of head orientation is very robust; it occurs wherever one is seated and across diverse listener groups.

SRTs were always lowest for normal hearing listeners, who had similar thresholds for speech and for speech-shaped noise interferers. For hearing impaired listeners, SRTs were about 4 dB higher with speech-shaped noise interferers and 6 dB higher for continuous speech interferers. For cochlear implant simulations, implemented using the SPIRAL vocoder [8], SRTs are were 13 dB higher than normal hearing listeners for speech-shaped noise interferers and 19 dB higher for speech.

The fact that impediments to hearing caused accentuated difficulties with speech interferers suggests that these modes of hearing are more susceptible than normal hearing to modulation masking [9]. The speech-shaped noise is a steady-state masker, whereas, even with 8 talkers, the restaurant babble is relatively modulated. It is worth bearing in mind that the interferers varied substantially in distance from the listener for each listener position, so the degree of residual modulation after summing eight voices would have been higher than would occur if eight voices were summed at equal level and much higher than the intrinsic modulation of steady-state noise sources [10].

## 3. SEATING POSITION

### 3.1 Rationale

Although the above simulations of the Mezzaluna restaurant showed rather consistent benefits from head orientation, it also found that different tables in the restaurant produced different levels of intelligibility; there is a difference of 4-5 dB in average SRT between tables 9 and 14. Moreover, the measurements made in Mezzaluna, used one nominated speaking position and one nominated listening position in each table, making it impossible to determine whether the table itself was in a noisier or quieter spot, or whether the particular seat at a given table makes a difference.

In the present experiment, two seats on opposite sides of the same table are tested. In this way, we were able to examine whether it is better to face the wall or into the room and what influence a wall to the side of the table might have.

### 3.2 Method

In order to examine the choice of seat on a particular table, we made use of a simple room reverberation model [11] that was adapted to include appropriate head-related cues for each incident ray at the receiver position. A given binaural room impulse response was built up by adding copies of head-related impulse responses from [12] for each predicted acoustic ray arriving at the listening position. These impulse responses were selected according to their angle of incidence, delayed according to their path length and scaled according to the distance travelled and the acoustic absorbance of the room boundaries with which they had interacted. This model was previously used in [13] and the room design was the same as the simulated reverberant restaurant used in Expt. 1 of that paper, except that the absorption coefficients for walls, ceiling and floor were 0.2, 0.95 and 0.05, respectively. In addition, in [13] the listener was always seated on the central table, whereas in the present experiment, six different seats were tested (Fig. 2).

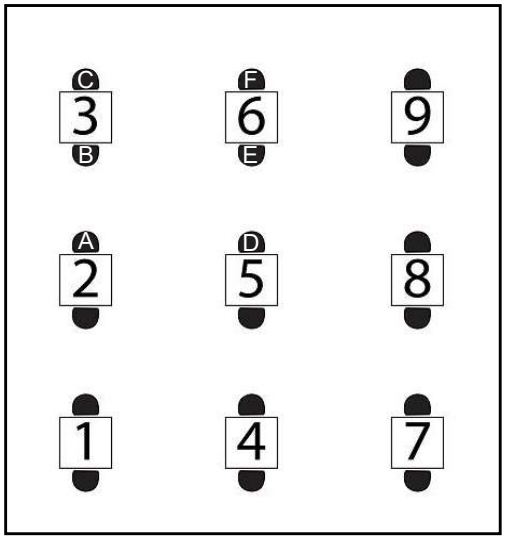


**Figure 2.** Layout of the simulated restaurant with the six tested listening positions labelled A-F.

Speech reception thresholds were measured using the same method as in [13], but with one continuous speech interferer (a book reading) placed at each of the other eight tables. Target speech (IEEE sentences) came, in each case, from the opposite side of the listener's table.

### 3.3 Results

Figure 3 show speech reception thresholds varied substantially across the 6 seating positions [$F(5,85)=8.25$, $p<0.001$], that there was an effect of head orientation [$F(2,34)=10.9$, $p<0.001$], and that the effect of head orientation was stronger at some tables than at others [$F(10,170)=5.0$, $p<0.001$]. Sometimes the effect of head orientation appears to have strong asymmetry, by which orientation to one side was much more effective than orientation to the other. The data fit well with the model predictions ($r=0.82$).

The results confirm the model's prediction that the effect of orienting the head is much larger in some seating

positions, such as B and E, than in others, and that the effect is asymmetric in some cases, notably B.

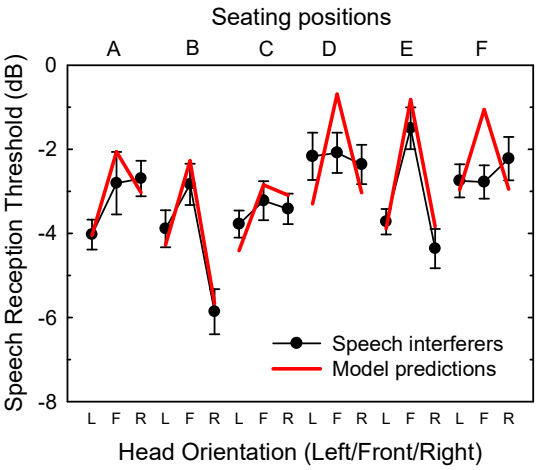


**Figure 3.** Speech reception thresholds in each of the six seatring positions illustrated in Fig. 2; model predictions from [2].

### 3.4 Discussion

Relating the results to the spatial configuration of the room, it is possible to interpret them intuitively in terms of the changes in target and interfering sound level at ear. The model predicts only small benefits in this crowded and reverberant listening situation from binaural unmasking (averaging 0.48 dB), so signal-to-noise ratio at the better ear is the dominant effect.

The highest thresholds are seen for seat D, which is at a table in the middle of the room. Here, the listener is surrounded by competing voices and it is impossible to derive substantial relief from the noise by any turn of the head. The model does predict a small effect, but the empirical data falls well short of even this prediction.

The lowest thresholds and greatest effects of head orientation are observed in seats B and E which are both facing towards the restaurant wall, rather than into the restaurant. It is thus predicted by the model, and confirmed empirically for this room simulation, that facing the edge of the restaurant, as in seats B and E, is advantageous compared to taking the opposite seats, C and F. Interestingly, the lowest thresholds is achieved in seat B with the head orientated to the right. In this position, the listener effectively has the target voice from seat C at the left ear and the rest of the restaurant at the right ear.

## 4. CONCLUSIONS

It is most acoustically advantageous to have a table towards the edge of a restaurant, to select a seat that faces the wall rather than into the room, and to orient one's head to one side of the target voice.

## 5. REFERENCES


[1] M. Lavandier and J. F. Culling, "Prediction of binaural speech intelligibility against noise in rooms.," *J. Acoust. Soc. Am.*, vol. 127, pp. 387–399, 2010.

[2] S. Jelfs, J. F. Culling, and M. Lavandier, "Revision and validation of a binaural model for speech intelligibility in noise.," *Hear. Res.,* vol. 275, pp. 96–104, 2011.

[3] N. I. Durlach, "Equalization and Cancellation Theory," *Foundations of Modern Auditory Theory Vol. II*, pp. 371–463, 1972.

[4] J. F. Culling, "Evidence specifically favoring the equalization-cancellation theory of binaural unmasking.," *J. Acoust. Soc. Am.*, vol. 122, pp. 2803–2813, 2007.

[5] ANSI "Methods for calculation of the speech intelligibility index," *ANSI* S3.5-1997, American National Standards Institute, New York, 1997.

[6] J. A. Grange and J. F. Culling, "Head orientation benefit to speech intelligibility in noise for cochlear implant users and in realistic listening conditions," *J. Acoust. Soc. Am.*, vol. 140, pp. 4061–4072, 2016.

[7] J. A. Grange, J. F. Culling, B. Bardsley, L. I. Mackinney, S. E. Hughes, and S. S. Backhouse, "Turn an Ear to Hear: How Hearing-Impaired Listeners Can Exploit Head Orientation to Enhance Their Speech Intelligibility in Noisy Social Settings," *Trends Hear.,* vol. 22, pp. 1–13, 2018.

[8] J. A. Grange, J. F. Culling, N. S. L. Harris, and S. Bergfeld, "Cochlear implant simulator with independent representation of the full spiral ganglion," *J. Acoust. Soc. Am.*, vol. 142, pp. EL484-489, 2017.

[9] S. Jørgensen and T. Dau, "Predicting speech intelligibility based on the signal-to-noise envelope power ratio after modulation-frequency selective processing.," J. *Acoust. Soc. Am.*, vol. 130, no. 3, pp. 1475–1487, 2011.

[10] M. A. Stone and B. C. J. Moore, "On the near non-existence of 'pure' energetic masking release for speech.," *J. Acoust. Soc. Am.*, vol. 135, pp. 1967–77, 2014.

[11] J. B. Allen and D. A. Berkley, "Image method for efficiently simulating small-room acoustics," *J. Acoust. Soc. Am.*, vol. 65, pp. 943–950, 1979.

[12] W. G. Gardner and K. D. Martin, "HRTF measurements of a KEMAR," *J. Acoust. Soc. Am*, vol. 97, pp. 3907–3908, 1995.

[13] J. F. Culling, "Speech intelligibility in virtual restaurants," *J. Acoust. Soc. Am*, vol. 140, pp. 2418–2426, 2016.