

PAPER

The effects of ceiling height and absorber placement on speech intelligibility in simulated restaurants.

John F. Culling*, Rada Gocheva, Yanyu Li, Nur Kamaludin

School of Psychology, Cardiff University, Tower Building, Park Place, Cardiff, CF10 3AT, U.K.

(Received July 12, 2018)

Abstract: The intelligibility of speech was measured in simulated rooms with parametrically manipulated acoustic features. The rectangular rooms were designed to simulate restaurant environments with either three or nine occupied tables, using either speech or noise as interfering sounds. The existence of more detailed acoustic features, such as furniture was also modelled. The measurements revealed that reverberation time was poorly correlated with speech intelligibility. In contrast, a psychoacoustic model of spatial release from masking produced accurate predictions for noise interferers and ordinally correct predictions for speech interferers. It was found that rooms with high ceilings facilitated higher speech intelligibility than rooms with lower ceilings and that acoustic treatment of walls facilitated higher speech intelligibility than equivalent treatment of ceilings. Ground-level acoustic clutter, formed by furniture and the presence of other diners had a substantial beneficial effect. Where acoustic treatment was limited to the ceiling, it was found that continuous acoustic ceilings were more effective than suspended panels, and that the panels were more effective if acoustically absorbent on both sides. The results suggest that the most effective control of reverberation for the purpose of speech intelligibility is provided by absorbers placed vertically and close to the diners.

Keywords: Binaural unmasking, spatial release from masking, reverberation, acoustic treatment.

PACS number: 00.00Aa

1. INTRODUCTION

1.1 Problem

Speech is often heard against background noise and in reverberant spaces; spaces that are intended for conversational interactions are often too noisy and reverberant for effective communication. This paper asks what exactly is wrong with these spaces, and how their deficiencies would best be addressed. The particular example of restaurants is explored.

1.2. Reverberation time (T_{60})

The acoustic quality of a room is frequently assessed by measurement of its T_{60} , the duration required for the sound level in the room to drop by 60 dB after the offset of a sound source. T_{60} increases with the volume of the room, V , and decreases with the total absorption of the room's surfaces, A , a relationship captured by the Sabine equation.

$$T_{60} = 0.161 \frac{V}{A}$$

However, the use of this very simple equation to predict T_{60} has been criticized as neglecting the influence of ab-

sorber placement [1]. Indeed, for rooms dominated by specular reflection of sound, any deviation from equal dimensions will also lead to large inaccuracies [2]. A more reliable way to derive T_{60} is to calculate it from a room impulse response using reverse integration [3]. The required impulse response can be measured acoustically from a real room or predicted from plan using ray-tracing software.

T_{60} was originally developed for the assessment of concert halls [4]. It is a convenient measure, because a single value can be derived for a particular room, and the value is largely independent of measurement position. Perhaps as a consequence of its convenience as a single-value, off-the-shelf measure, it has since been applied much more widely, such as in the regulation of classroom acoustics, for which maximum permissible values are specified [5], but it is questionable whether such use is appropriate.

1.3. The speech transmission index (STI)

The STI was developed as a specific predictor of speech intelligibility in reverberant rooms [6]. The understanding of speech is highly reliant on the transmission to the listener of modulations in the intensity of the speech within each

frequency band. The STI is based on the modulation transfer function from one point in a room to another. It can predict the intelligibility of speech delivered between these two positions with high accuracy [6]. In contrast to T_{60} , therefore, the STI is not a single-value measurement for a given room, but is location dependent. The ideal application is a situation in which the locations of speaker and listeners are highly predictable, such as a lecture theatre. For a lecture theatre, one can derive the STI for each seat in the audience for a speaker located at the lectern.

The STI can also make predictions for the effect of noise in the room, but it is limited, in this respect, to noise that is totally diffuse. In the real world, interfering noise sources, such as concurrent conversations, can be nearby, and consequently much more intense at one ear than the other. Moreover, in the restaurant application considered here, it is likely that the target speech will also be very close to the listener. In this situation, the target speech is dominated by direct sound and so direct effects of reverberation on its intelligibility are minimal. Instead, the effects of reverberation on the interfering sounds affect the ease of conversation [7]. The STI cannot model the effect of reverberation on interfering sounds.

1.4. Binaural models

Binaural models of spatial release from masking take into account the differences in timing and level of target and interfering sounds at the two ears [8,9,10]. They can be used to predict intelligibility in combinations of noise and reverberation for specific spatial configurations of listener, speaker and interferers. They are thus more appropriate for predicting intelligibility in social spaces, such as restaurants or classrooms used for group work.

However, predictions from a binaural model will change with each change in spatial configuration, so the assessment of the *room* will depend upon the exploration of a range of listening scenarios that might be encountered in that room. The calculation of multiple scenarios introduces a potential computational explosion in which every permutation is tested. This computational load must be contained by adopting a representative sample of spatial configurations and by using efficient computation. The binaural model used here [10] is very efficient, because it operates directly upon binaural room impulse responses rather than first generating binaural simulations based on those impulse responses.

2. EXPERIMENTS 1 & 2

2.1 Rationale

The Sabine equation shows that T_{60} increases with greater room volume and that it decreases with greater total

absorption. If a long reverberation time is bad, one would therefore expect that an increase in the height of a restaurant would impair intelligibility. Applying the Sabine equation uncritically, one might also expect that the distribution of acoustic absorbance in the room would be irrelevant, provided that the total absorbance is constant. Experiments 1 and 2 test these predictions.

2.2. Methods

Room impulse responses were generated using a source-image model [11], which calculates specular ray paths within a rectangular box. The model implementation used appropriate head-related impulse responses for each ray reaching the listener's head. The resulting binaural room impulse responses were convolved with speech-shaped noise sources, continuous speech interferers and target sentences in order to create virtual simulations of different listening scenarios for a listener and speaker at the central table of a notional 3×3 array of tables within a $6.4 \text{ m} \times 6.4 \text{ m}$ room.

Experiment 1 compared speech intelligibility for absorbance placed mainly on the ceiling (2.5 m high), or mainly on the walls, but keeping the total absorbance constant. Ceiling absorbance was 0.95 or 0.05; wall absorbance was 0.05 or 0.62. Floor absorbance was 0.1, in each case. Experiment 2 compared speech intelligibility for high (5 m) and low (2.5 m) ceilings, for a room with absorbance of 0.1 on walls and ceiling and 0.2 on the floor. Both experiments compared cases with 2 and 8 interferers and both experiments compared speech and noise interferers.

Speech reception thresholds (SRTs) were measured [12]. These were the ratios of target power to interferer power *in the room* (not at the listener's ears) that gave 50% intelligibility of words at the listening position. This measure is used because a pertinent effect of room acoustics is to alter the signal-to-noise ratio at the listener's ears. The interferers were played continuously during the measurement of the SRT for 50% intelligibility of IEEE sentences. Target speech materials were counterbalanced across conditions. See [13] for further details of the experimental procedure.

2.3 Results

Fig. 1 shows that acoustic absorbance located on the ceiling produced higher SRTs, and was therefore less effective than equivalent absorbance located on the walls [$F(1,7)=53.1, p<0.001$]. The effect is very similar for both speech and noise interferers (interaction non-significant). Speech interferers produced less masking than noise [$F(1,7)=22.0, p<0.02$] and 8 interferers produced more masking than 2 [$F(1,7)=31.6, p<0.001$], but this effect was largely limited to speech interferers as reflected by a significant interaction [$F(1,7)=21.1, p<0.001$]. It should be noted

that, as in [13], sound levels for different numbers of interferers were equalized, such that any effect of interferer number reflects an influence of the spatial distribution of the interferers on auditory perception, rather than their combined sound level.

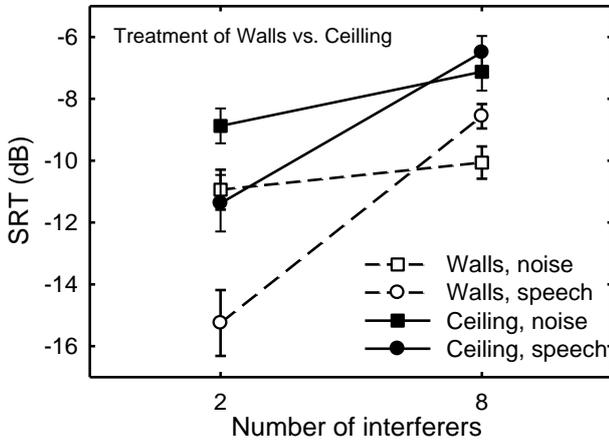


Fig. 1. SRTs for wall vs. ceiling treatment for rooms with the same total absorbance. Error bars are one standard error of the mean

Fig. 2 shows that high ceilings produced consistently lower (better) SRTs than low ceilings [$F(1,7)=29.4$, $p<0.001$]. There was, again, a detrimental effect of 8 compared to 2 interferers [$F(1,7)=82.9$, $p<0.001$], but, again, the effect was largely limited to speech interferers [$F(1,7)=79.8$, $p<0.001$]. There was no main effect of interferer type or any other interaction.

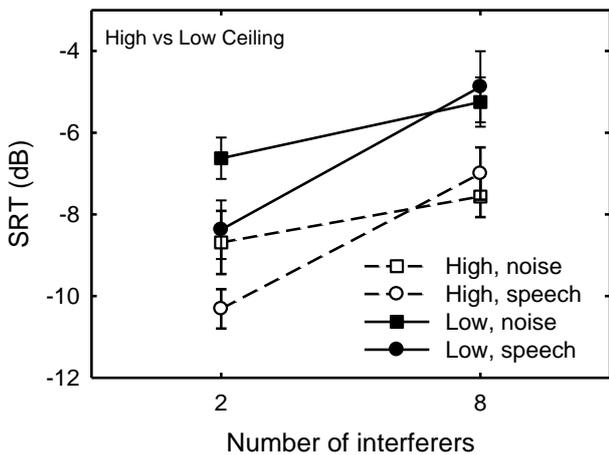


Fig. 2. Mean SRTs for high vs. low ceilings. Error bars are one standard error of the mean.

2.4 Discussion

In experiment 1, rooms with equivalent absorption produced markedly different speech intelligibility under the same conditions. This result indicates that the total absorption, and thus the Sabine equation, cannot be relied upon as a means of predicting the suitability of rooms for social gatherings. The actual T_{60} values, as revealed by reverse integration, showed that the room with absorption distributed

across the walls was substantially less reverberant than the one in which absorption was concentrated on the ceiling. The measured intelligibility in these rooms showed that wall treatment also led to substantially better intelligibility, so the experimental results are nonetheless consistent with the idea that longer T_{60} leads to poorer intelligibility.

In experiment 2, however, increasing the ceiling height increased the measured T_{60} , but led to *improved* intelligibility. The results of the two experiments thus show that T_{60} is an inaccurate predictor of intelligibility within the room. The correlation between T_{60} for each of the four rooms tested and the average SRTs for those rooms is only 0.62. This outcome can be contrasted with the effectiveness of the binaural model of spatial release from masking [10]. Importantly, the model correctly predicted that intelligibility would improve with increased ceiling height as well as with distribution of absorbance to the walls. Fig. 3 shows that the model predictions gave a correlation of 0.91 with the eight noise-interferer SRTs from the two experiments (regression slope = 1.11). The binaural model thus provided a more reliable index of intelligibility than T_{60} .

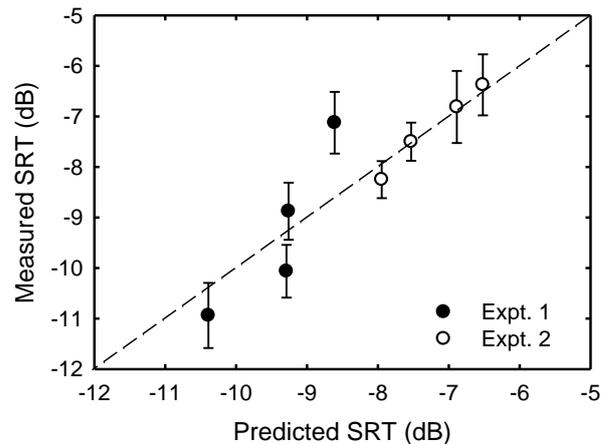


Fig 3. Comparison of SRTs predicted by the Jelfs et al. model [10] with the measured SRTs for the noise conditions of Experiments 1 and 2.

3. EXPERIMENTS 3 & 4.

3.1 Rationale

The source-image model used in Experiments 1 and 2 is only capable of simulating specular reflections in an empty rectangular space. Experiments 3 and 4 employed CATT Acoustic™ (8.0), a commercial software package that can simulate complex geometries, surface scattering of sound and sound source directivity. Fig. 4 illustrates an example room geometry. Experiment 3 looked again at ceiling height, but in the context of the effects of acoustic clutter formed by furniture and the bodies of room occupants. The presence of such clutter substantially increases the area of

absorbent material in the room and thus the total absorbance. It also scatters sound. Experiment 4 examined the benefits of different forms of ceiling treatment, a full acoustic ceiling, and separate panels. The separate panels were either absorbent on both sides or reflective on the lower surface so that they might provide beneficial early reflections to the target voice on the other side of the table.

3.2. Methods

The complex room geometries were first designed using Google Sketchup™, defining a rectangular room with or without planar representations of tables and diners. The geometry was then imported into CATT Acoustic™, where surfaces were allocated absorption spectra across 6 octave bands (0.125–4 kHz) for plastered walls (0.05, 0.05, 0.05, 0.05, 0.05), wooden tables (0.19, 0.23, 0.20, 0.20, 0.15, 0.20) or clothed humans (0.16, 0.24, 0.56, 0.69, 0.81, 0.78). The room was 6 m × 10 m and of variable height. An array of 3×5 tables was included, in which the central group of 3×3 tables was used in the same way as in Experiments 1 and 2. Default sound directivity for a human head was used for each source, facing directly across their respective tables.

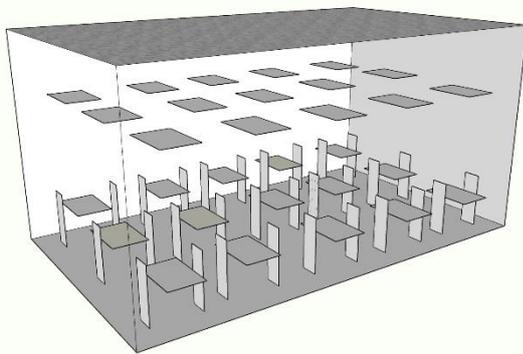


Fig. 4. Example room geometry used in Experiment 4, showing table surfaces, diners and acoustically absorbent ceiling panels.

Both experiments used the same interferer distributions as in experiments 1 and 2, but only speech interferers were used. In experiment 3, ceiling heights of 2.5 or 5 m were tested, with and without the absorbing and scattering effects of furniture and people. In experiment 4, a 5-m ceiling was used with 4 different acoustical treatments: untreated, a complete acoustic ceiling, fully absorbent suspended panels, 3.5 m above the floor (see Fig. 4), and similar suspended panels with a reflective lower surface. Furniture and people were included in all conditions of experiment 4.

3.3 Results

Fig. 5 shows the results of experiment 3. High ceilings,

again, provided superior speech intelligibility to low ceilings [$F(1,7)=7.1$, $p<0.05$], but only in the absence of furnishing and people [$F(1,7)=9.6$, $p<0.02$]. The greater absorbance produced by this clutter appears to overwhelm the benefit of a higher ceiling. The largest effect was of the presence of furnishing/people itself which improved SRTs by about 6 dB [$F(1,7)=696$, $p<0.001$], followed by the familiar influence of the number of interferers [$F(1,7)=23.7$, $p<0.005$].

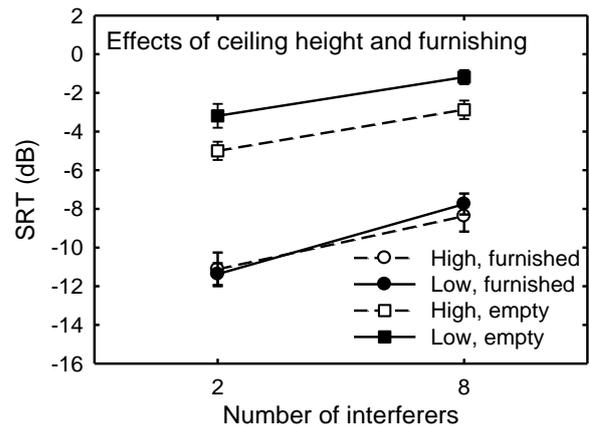


Fig. 5. Mean SRTs for high vs. low ceilings, with and without absorbance and scattering from furniture and people. Error bars are one standard error of the mean.

Fig. 6. shows the results of experiment 4. There were significant effects of different ceiling treatments [$F(3,21)=39.5$, $p<0.001$] and of the number of interferers [$F(1,7)=65.3$, $p<0.001$], but no interaction. Bonferroni-corrected t -tests showed that the full ceiling treatment was superior to each of the other options, and that the two-sided suspended panels were superior to an untreated ceiling ($p<0.002$, in each case).

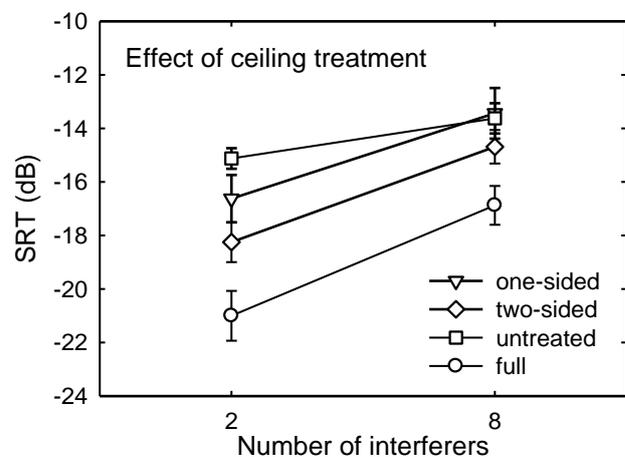


Fig. 6. Mean SRTs for different forms of ceiling treatment: one- and two-sided suspended panels, a bare ceiling and a full acoustic ceiling. Error bars are one standard error of the mean.

3.4 Discussion

Experiment 3 repeated the ceiling-height manipulations of experiment 2, but using a room-modelling program that is not limited to modelling specular reflection, and which can model the scattering effects of the sort of acoustic clutter that would normally be present in the floor area of a restaurant. These changes did not reverse the fundamental effect of ceiling height, although the acoustic clutter attenuated the effect to negligible levels. This reduction is probably brought about by substantially lower reverberation in the presence of the clutter, which increases the absorbance of the room and so reduces the masking energy from the interferers.

Experiment 4 tested the effects of different types of ceiling treatment. A conventional continuous acoustic ceiling proved to be most effective, presumably because it provides the greatest total absorbance and so a reduction in the acoustic energy density produced by the interferers. The suspended panels with reflective lower surfaces did not prove to be significantly more effective than an untreated ceiling. They might be more effective if suspended lower in order to produce a stronger reflection across each table.

It is noteworthy that the SRTs from experiment 4 were lower than in the other experiments. In particular, one would expect results from the untreated 5-m-high ceiling with furnishing in experiments 3 and 4 to give similar thresholds, but SRTs from experiment 4 are about 3 dB lower in this case. The reason was traced to an error in the room geometry in experiment 4, for which the representation of seated people were 0.88 m high rather than 0.72 m. Since the sources were located just in front of the body of the room occupant, the higher absorber acted rather as though the person was seated in a high-backed, acoustically reflective chair. This “chair back” removed direct sound paths between the listener and some of the interfering sound sources, particularly those behind the listener and with their back to the listener.

The binaural model of spatial release from masking was successful in predicting the ordinal relationship among the SRTs from experiment 3 and 4, giving an overall correlation of 0.96. However, it did not accurately predict the size of the effects, with a regression slope of 0.62.

4. GENERAL DISCUSSION

4.1 Ceiling height

Experiments 2 and 3 demonstrated that high ceilings provide a better environment for conversation than low ceilings. The effect is not large, but contradicts the notion that a long T_{60} is intrinsically bad for intelligibility. As reflected in the Sabine equation, any increase in room volume

will increase the reverberation time, as the volume will always increase faster than the surface area, even if only one dimension is changed. The effect also stands in contrast to intuitive expectations. High-ceilinged rooms sound reverberant, and people also intuitively, and correctly, associate reverberation with conversational difficulty, so they understandably expect a high ceiling to be detrimental.

The combination of theory and intuition leads to confusion that can be seen even in the academic literature on restaurant acoustics. For instance, the relationship of greater ceiling height to longer reverberation time has been inferred to produce “a negative effect on the intelligibility” [14], and decreasing ceiling height has been listed as one of several potential methods for improving intelligibility [15]. The reason for this counter-intuitive effect is that the increase in room volume serves not only to increase the reverberation time, but also to reduce the noise level. The sound energy spreads into the additional space, reducing the sound energy density.

4.2 Absorber placement

Experiment 1 showed that equivalent absorbance on the walls is more effective than on the ceiling. There are potentially two factors at work, here.

First, when the absorbance is distributed more evenly, specularly reflected rays meet more absorbance per unit time than if they are only absorbed when encountering one of the six surfaces. This can be seen from a marked difference in reverberation time for the two rooms used in experiment 1. Although the Sabine equation would predict the same T_{60} of 360 ms for each room, using the reverse integration method [3] on the room impulse responses showed a T_{60} of 1770 ms for the room with an absorbent ceiling and 520 ms for the room with absorbent walls. This difference will have been exaggerated by the use of a totally specular model of room reverberation, but since real rooms will have some specular reflection, the principle stands: an even distribution of absorbance across the room boundaries is more effective than treating one surface.

Second, it is possible that the lateral positioning of ears on a human head may mean that binaural processing is susceptible to reverberation travelling in the same plane as the ears. Humans use their binaural system to exploit differences in the timing and sound level of sound arriving at their two ears in order to improve speech understanding in noise [16,17]. The timing and level differences are caused by the spacing of the ears and the presence of an acoustically reflective head between them. This arrangement means that sounds coming from the side have a greater influence on these cues than sounds coming from above or below, and, consequently, lateral reflections from the walls

can be disruptive to binaural processing. Effects of this sort are likely to have been small in the present experiment, because interferers were always located on both sides of the listener, but may be larger in other circumstances.

The binaural model [10] was successful in predicting the intelligibility of speech in these four experiments, primarily because it evaluates the importance-weighted signal-to-noise ratio at the listener's ears. It also accounts for the two main binaural mechanisms of speech intelligibility in noise, better-ear listening and binaural unmasking. Better-ear listening is based on the signal-to-noise ratio at the better ear, and will reflect the acoustic energy density of the interferers at the listening position as well as the proximity of the speaker. Binaural unmasking is a process driven by timing differences between target and interfering sound, and is less effective the less interaurally coherent the interferer is. A multiplicity of interferers and lateral reflections from walls tend to reduce the interaural coherence.

4.3 Implications for room design

Overall, the results of these experiments indicate that high ceilings are a favourable feature in rooms designed for social interaction. An acoustically absorbent ceiling can improve intelligibility further, its success being based primarily on an increase in the total absorbance, but absorbent materials in other locations tend to be much more effective.

Experiment 1 showed that the same absorbance on the walls was markedly more effective than when it was placed on the ceiling, and experiment 3 showed that acoustic clutter in and around the room occupants, including the occupants themselves, has a large beneficial effect. It would seem, therefore, to be more productive to introduce wall absorbers and acoustically absorbent furniture rather than to treat the ceiling. The serendipitous finding that high seat backs can have a big effect indicates that any way to interrupt the space at head level has a disproportionate benefit. Restaurants with booth-style tables, pillars or vegetation at head level will benefit from this effect.

REFERENCES

- [1] M. R. Schroeder, "Computer Models for Concert Hall Acoustics" *Am. J. Phys.*, 41, 461–471, (1973).
- [2] M. C. Gomperts, "Do the classical reverberation formulae still have a right for existence?" *Act. Acust. u. Acust.*, 16, 255–268, (1965).
- [3] M. R. Schroeder, "New method of measuring reverberation time" *J. Acoust. Soc. Am.*, 37, 1187–1188, (1965).
- [4] W. C. Sabine, "Architectural Acoustics" *Proc. Am. Acad. Art. Sci.*, 42, 51–84, (1906).
- [5] "Building Bulletin 93: Acoustic design of schools: performance standards" U.K. Dept for Education. (2015).
- [6] T. Houtgast, and H. J. M. Steeneken, "A review of the MTF concept in room acoustics and its use for estimating speech intelligibility in auditoria" *J. Acoust. Soc. Am.*, 77, 1069–1077, (1985).
- [7] M. Lavandier and J. F. Culling, "Speech segregation in rooms: effects of reverberation on both target and interferer" *J. Acoust. Soc. Am.*, 122, 1713–1723, (2007).
- [8] R. Beutelmann and T. Brand, "Prediction of speech intelligibility in spatial noise and reverberation for normal-hearing and hearing-impaired listeners" *J. Acoust. Soc. Am.*, 120(1), 331–342, (2006).
- [9] M. Lavandier and J. F. Culling, "Prediction of binaural speech intelligibility against noise in rooms" *J. Acoust. Soc. Am.*, 127, 387–399, (2010).
- [10] S. Jelfs, J. F. Culling and M. Lavandier, "Revision and validation of a binaural model for speech intelligibility in noise." *Hear. Res.*, 275, 96–104, (2011).
- [11] J. B. Allen and D. A. Berkley, "Image method for efficiently simulating small-room acoustics" *J. Acoust. Soc. Am.*, 65, 943–950, (1979).
- [12] R. Plomp, R. and A. Mimpen. "Improving the reliability of testing the speech reception threshold for sentences." *Audiology*, 18, 43–52. (1979).
- [13] J. F. Culling, "Speech intelligibility in virtual restaurants" *J. Acoust. Soc. Am.*, 140, 2418–2426, (2016).
- [14] J. Kang "Numerical modelling of the speech intelligibility in dining spaces" *Appl. Acoust*, 63, 1315–1333, (2002).
- [15] M. Nahid and M. Hodgson "Prediction of optimal conditions for verbal-communication quality in eating establishments" *J. Acoust. Soc. Am.*, 129, 2005–2014, (2011).
- [16] A. W. Bronkhorst, R. Plomp "The effect of head-induced interaural time and level differences on speech intelligibility in noise" *J. Acoust. Soc. Am.*, 83, 1508–1516, (1988).
- [17] M. L. Hawley, R. Y. Litovsky, J. F. Culling "The benefit of binaural hearing in a cocktail party: Effect of location and type of interferer" *J. Acoust. Soc. Am.*, 115, 833–843, (2004).