On the importance of early reflections for speech in rooms

Bradley, J. S.; Sato, H.; Picard, M.

Publisher's version / Version de l'éditeur:
https://doi.org/10.1121/1.1570439
Journal of the Acoustical Society of America, 113, June 6, pp. 3233-3244, 2003-06-01

On the importance of early reflections for speech in rooms

Bradley, J.S.; Sato, H.; Picard, M.

NRCC-45431

A version of this document is published in / Une version de ce document se trouve dans :
Journal of Acoustical Society of America, v. 113, no. 6, June 2003, pp. 3233-3244

On the Importance of Early Reflections for Speech in Rooms

J.S. Bradley and H. Sato

National Research Council, Ottawa, Canada, K1A 0R6

M. Picard

Ecole d'orthophonie et d'audiologie, Université de Montréal, Montréal, Canada, H3C 3J7

Shortened title: "Early reflections for speech"

PACS numbers: 43.55.Hy, 43.71.Gv

Received: ____

ABSTRACT

This paper presents the results of new studies based on speech intelligibility tests in simulated sound fields and analyses of impulse response measurements in rooms used for speech communication. The speech intelligibility test results confirm the importance of early reflections for achieving good conditions for speech in rooms. The addition of early reflections increased the effective signal-to-noise ratio and related speech intelligibility scores for both impaired and non-impaired listeners. The new results also show that for common conditions where the direct sound is reduced, it is only possible to understand speech because of the presence of early reflections. Analyses of measured impulse responses in rooms intended for speech show that early reflections can increase the effective signal-to-noise ratio by up to 9 dB. A room acoustics computer model is used to demonstrate that the relative importance of early reflections can be influenced by the room acoustics design.

I INTRODUCTION

The beginnings of our understanding of how we perceive sound reflections that arrive within a short time after the direct sound can be traced to the work of Joseph Henry in the 1850s [1]. Work by Haas [2] and Wallach et al. [3] showed explicitly how early reflections are integrated with the direct sound to make the direct sound seem to be effectively louder. Lochner and Burger [4] carried out extensive experiments to determine exactly how delayed reflections affected articulation test results as a function of the amplitude and delay time of an early-arriving reflection. Although these early results demonstrated an extensive understanding of the benefits of early reflections to speech intelligibility in rooms, they sometimes focussed on the negative effects of early reflections. For example, Haas considered the point at which early reflections became disturbing. Although the early work of Lochner and Burger and others provides a solid basis for the importance of early reflections, they provide little information on the expected improvements to speech intelligibility scores in actual rooms due to early reflection energy.

Lochner and Burger's results suggest that speech energy in early arriving reflections is more or less equivalent to similar amounts of direct speech energy in terms of increasing speech intelligibility scores. They developed the concept of useful-to-detrimental sound ratios in which 'useful' is defined as the sum of the direct and early-reflected speech energy and 'detrimental' is the sum of the later-arriving speech reflections and the ambient noise. This concept has been shown to be well correlated with speech intelligibility test scores in a wide range of rooms [5-7] and to be strongly correlated [8] with the newer speech transmission index (STI) concept.

However, Bess [9] has recently claimed that the benefits of early reflections are not experienced by impaired listeners. This is in contradiction with the results of Nábělek and Robinette [10] that indicated normal hearing and impaired listeners benefited equally from a single added early reflection. In addition, the importance of early reflections in rooms is apparently not well appreciated in conventional room acoustics design which is usually based primarily on obtaining a particular preferred reverberation time [11]. A lack of appreciation of the importance of early reflections is no doubt responsible for some recommendations for very short reverberation times for rooms for speech. There is often a misconception that reverberation time must be minimized rather than optimised. Of course, very short reverberation times require increased absorption on room surfaces which is likely to lead to reduced early reflection energy and hence to reduced speech intelligibility.

There are many situations in speech communication in rooms where early reflections would appear to be particularly important such as a teacher talking to students from somewhere in the middle of a classroom or an actor on a thrust stage. Students behind the teacher would receive much reduced direct sound and presumably benefit significantly from early reflections of the speech sounds. In many other situations where the talker directs their attention to some part of a group of listeners, those in other directions benefit from early reflections of the speech sounds. Although we typically include early reflections within the first 50 ms after the direct sound as being useful early reflections, we do not have a quantitative indication of the relative magnitude (and hence importance) of this early reflection energy in typical rooms.

The purpose of the new work reported in this paper was to first confirm the importance of early reflections to speech intelligibility test scores in conditions representative of real rooms. This was done using speech intelligibility tests for both impaired and non-impaired subjects in simulated sound fields. These simulations allowed comparison of the benefits of added early reflection energy with that of increased direct sound energy for cases with and without reverberant sound. Tests also considered the critical value of early reflections for situations where the direct sound is reduced in amplitude. The second part of this paper examines the magnitude of the benefit of early reflections in a range of rooms intended for speech communication and illustrates how modern room acoustics computer models can be used to explore various room designs in terms of improved early reflection levels.

II SPEECH INTELLIGIBILITY TESTS IN SIMULATED SOUND FIELDS

The goal of the speech intelligibility tests in simulated sound fields was first to confirm directly that increased speech energy in early reflections has a similar effect to increased direct speech energy. The second goal was to demonstrate that both hearing-impaired and non-impaired listeners benefit from early arriving reflections of speech sounds. The benefits of increased early reflections are then shown to be present in more realistic sound fields that also included reverberant sound typical of many rooms for speech communication. Finally, in situations where the direct sound is reduced in level such as when the talker's head is turned, early reflections are shown to be essential to achieving adequate speech intelligibility.

A. Method

1. Sound field simulation procedures

All simulated sound fields were produced using an 8-channel electroacoustic system with loudspeakers arranged around the listener in an anechoic room. The 8 loudspeakers were located at a distance of 1.7 m from the listener and their angular locations relative to the listener are described in Table I. Each of the 8 channels of electronics included programmable digital equalizers that included time delays and reverberators that could all be changed under computer control via a MIDI interface. The loudspeaker responses were corrected to be flat ±3 dB from 80 Hz to 12 kHz.

The loudspeaker located directly in front of the listener produced the simulated direct sound (first arriving sound) and in some experiments also produced reverberant sound. The other seven loudspeakers each produced one early reflection and in some experiments reverberant sound. The early reflections arrived at the listener within the first 50 ms after the direct sound. Figure 1 illustrates a measured impulse response for a condition that included a direct sound, early reflections and reverberation. Some sound fields included only a direct sound component, others included a direct sound and early reflections and others included a direct sound combined with both early reflections and reverberant sound. Where reverberant sound was included it was delayed to start just after the 50 ms early time interval as seen in the example in Figure 1. The overall amplitudes of each of these three component groups (direct sound, early reflections and reverberant reflections) were varied but the arrival times and relative amplitudes of early reflections were not changed.

Each loudspeaker also reproduced simulated ambient noise with a spectrum shape corresponding to that of an NC 40 contour and with a measured overall level at the listener of 47.6 dBA. The noise signals to each loudspeaker were passed through different digital reverberators to ensure that they were not exactly coherent.

Each experimental condition was measured with our RAMSoft room acoustics measurement software that uses a maximum length sequence signal to obtain impulse responses. A range of standard room acoustics parameters were measured in octave bands and are included here where relevant.

2. Subjects and speech intelligibility tests

Speech intelligibility scores were obtained using a Fairbanks rhyme test as modified by Latham [5] and as used in previous tests [6-8]. The test words were embedded in the sentence “Word number ___ is ___ write that down” and were spoken by a male talker.

Each 50 word test took a total of about two and a half minutes and the speech material was presented at a rate of approximately 3 syllables per second.

Subjects varied from 20 to 74 years of age. The wide range of ages was intended to ensure subjects included a variety of hearing abilities representative of those typically found for listeners in many rooms for speech. However, no subjects were included who used hearing aids or who had known serious hearing impairment. Many of the younger subjects had excellent hearing with minimal hearing loss. A number of the older listeners had mild to moderate hearing loss typical of many middle-aged listeners.

In some analyses the results of the subjects were divided into two groups according to their measured hearing loss (HL). The group with the least HL will be referred to as 'non-impaired' and the other group will be referred to as the 'impaired' listeners. Figure 2 plots the average HL (± 1 standard deviation) for each group. The non-impaired (and generally younger) group, on average, showed only very small reductions below threshold at 6 and 8 kHz. They had a mean age of 28 years and their mean hearing loss results shown in Figure 2 are very similar to the median for this age [12]. The impaired (older) group, on average, showed increasing HL above 1 kHz. Most of the impaired group were apparently unaware that they had any hearing loss. This group had a mean age of 60 years and again their mean hearing loss was similar to the expected median for this age [12]. The high frequency pure tone average (from 3, 4 and 6 kHz results and from both ears) for the non-impaired listeners was 5.2 dB and 30.5 dB for the impaired group.

B. Results of speech intelligibility tests

1. Comparison of effects of varied direct sound and varied early reflection levels on speech intelligibility scores.

The first comparisons were based on the results of tests in which subjects performed speech intelligibility tests for sound fields with varied speech signal-to-noise ratio (S/N) and for two types of reflection conditions. In one series of tests the sound fields consisted of only a direct (speech) sound and varied S/N was obtained by varying the amplitude of the direct speech sound with constant noise level. In the other series of tests the direct speech sound was fixed and S/N was varied by adding increased levels of early reflections in combination with the same constant noise level. Table II showing a summary of the measured conditions for each series of tests, includes A-weighted signal-to-noise ratios, S/N(A), varying from –5.8 to +4.4 dB for the direct sound only cases and -5.8 to +0.5 for the cases with added early reflections. The three added-reflections cases corresponded to increasing the total speech level by 0, 3 and 6 dBA. These were thought to be representative of conditions in typical rooms but later analyses (included in section III) showed that increases of up to 9 dBA can occur in typical rooms for speech.

Results were first considered for all 21 subjects and were not divided according to their level of hearing threshold shift. The complete group of all subjects can be thought of as representing a typical mixed group of listeners in a theatre audience or at a public meeting. Figure 3 plots the mean speech intelligibility scores (± 1 standard deviation) for each sound field condition for this combined group of listeners. An analysis of variance test (ANOVA) of the speech intelligibility scores was performed for the comparable S/N(A) cases with and without added early reflections. There was a statistically significant effect of S/N(A) ($p < 0.001$) but no significant effect of the two types of reflection conditions (direct sound only versus direct plus early reflections).

For both series of reflection conditions, speech intelligibility scores increased in a statistically significant manner with increasing S/N(A) as was expected. However, there was no identifiable difference between cases where the same increased S/N(A) was due only to varied direct sound and those where it was due to varied added early reflections. In these tests increased early reflection energy had the same effect on speech intelligibility scores as increased direct sound level.

2. Comparison of the effects of varied early reflection levels for impaired and non-impaired listeners.

In a second analysis of the same data, subjects were divided into two almost equal-sized groups according to their measured hearing loss. Eleven subjects were classified as ‘non-impaired’ and 10 as ‘impaired’. Their average HL characteristics were given in Figure 2. The mean speech intelligibility scores are plotted versus measured S/N(A) in Figure 4 separately for the impaired and non-impaired listeners. As for the combined results in Figure 3, the mean speech intelligibility scores increase with increasing S/N(A) but for corresponding S/N(A) values the impaired listener group always had lower intelligibility scores. That is, the trends of the results are in agreement with expectation from previous studies that have considered hearing-impaired listeners [13].

Again the statistical significance of the results was tested using ANOVA on the conditions with similar S/N(A). For the non-impaired subjects, there was a statistically significant effect of S/N(A) but no effect of the difference between sound fields with only direct sound and those with added early reflections. Exactly the same result was obtained for the impaired listeners' results. However, when impaired and non-impaired listeners were compared for either condition of direct sound only cases or added early-reflections cases, there was a statistically significant effect of the differences in HL.

Therefore these results show that both the impaired and non-impaired listeners benefit in a similar manner from added reflections. Increasing the S/N(A) by adding early reflections has the same effect as increasing the level of the direct sound on the resulting speech intelligibility scores for both no-impaired and impaired listeners.

These results also reconfirm that hearing impaired listeners need better conditions with higher S/N(A) values to enjoy the same level of speech recognition. The results in Figure 4 suggest that these particular impaired listeners would require approximately 5 dB higher S/N(A) values to have the same intelligibility scores as the non-impaired listeners. (For these listeners there was an approximate linear relationship between their measured HL and their speech intelligibility scores. Thus more impaired listeners would appear to require even better S/N(A) for them to be equally able to understand speech as less impaired listeners).

3. **Evaluation of the effect of varied early reflection levels in the presence of later arriving speech sounds.**

It was thought that the benefits of early reflections might be masked or diminished by the presence of reverberant sound which would normally be present in typical rooms intended for speech communication. Therefore, a third series of conditions was created that included varied levels of early reflections in the presence of a fixed level of the direct sound and reverberant sound. These conditions were identical to the previous series of sound fields in which the level of early reflections were varied but with a constant amount of added reverberation as illustrated in Figure 1. The acoustical measurements of the conditions are summarized in Table III and include reverberation times of 1.1 s representing a realistically reverberant condition for a room intended for speech communication. The speech levels and the S/N(A) values in Table III are termed 'effective' because the speech levels are based on the direct and early reflection energy and do not include the later arriving speech energy. Thus these 'effective' speech levels and S/N(A) values are exactly the same as in the test series where early reflection level was varied without the inclusion of reverberation (given in Table II).

Some subjects in this series of tests were different than in the original tests. There were 10 'non-impaired' subjects and 6 'impaired' subjects. However, the average HL characteristics of the 'impaired' and 'non-impaired' groups were very similar to the previous group averages shown in Figure 2. The group average HL values for the subjects in these tests differed by less than 2 dB at most frequencies with a few differences in higher frequency hearing loss of up to 4 dB.

The mean intelligibility scores are plotted versus effective S/N(A) in Figure 5 along with the results for the previous tests repeated from Figure 4. The speech intelligibility scores for the sound fields including reverberation follow similar trends to those of the previous results and there seems to be no significant effect of adding reverberation. Again this was confirmed by ANOVA tests of the significance of the results. There was a highly significant effect of S/N(A) ($p < 0.001$) but no significant effect of the differences in the types of reflections present.

These results confirm that the benefits of increased early reflection energy occur for both the impaired and non-impaired listeners in the presence of a realistic amount of reverberation.

One might expect that adding reverberation would at least reduce intelligibility scores for all listeners. This does not occur because the detrimental effects of the added reverberation were relatively small compared to the negative effects of the simulated ambient noise. This is confirmed by the almost invariant U_{50} values between corresponding cases shown in Table III for sound fields including reverberation and those without reverberation shown in Table II. U_{50} is a useful to detrimental sound ratio [6] where 'useful' is the sum of the direct and early-reflection energy arriving in the first 50 ms and 'detrimental' is the sum of later-arriving speech sounds and the ambient noise. Thus the reverberant energy is relatively small compared to the ambient noise energy and therefore its detrimental effect is negligible for these cases. Of course, this is similarly reflected in the almost identical speech transmission index (STI) values in Table III to those in Table II for the corresponding early reflections conditions.

4. Demonstration of the benefits of early reflections when the talker's head is turned

In most situations in rooms the presence of early reflections increases intelligibility by effectively enhancing the direct sound component and hence increasing the signal-to-noise ratio. However, in a number of situations where the direct sound is blocked or reduced in amplitude, the intelligibility of speech is more critically dependent on the presence of early reflection energy. One particular example is when the talker is not pointing directly toward the listener but is directing their speech in some other direction.

For example, when a teacher is talking to a class from the middle of the classroom, there are listeners both in front of and behind the talker. Due to the directionality of the human voice, those listeners not directly in front of the talker will experience reduced direct speech sound and this is especially so at the higher frequencies, which are critical for recognizing consonant speech sounds.

The final speech intelligibility tests in simulated sound fields were intended to demonstrate the importance of early reflections when the direct speech sounds are reduced in magnitude. The effect of the talker's head turning were simulated by modifying the spectrum of the direct speech sound to be equivalent to measured speech spectra at angles of 0, 90 and 180 degrees relative to straight ahead of the talker [14]. Subjects listened to speech representing these three talker angles in sound fields with only a direct sound and also in sound fields that also included early reflections. The sequence of early reflections was the same as used in the previous experiments and illustrated in Figure 1. No attempt was made to estimate changes in early reflections with talker head turning because these would be as likely to increase as to decrease in amplitude and hence to not systematically affect intelligibility.

Figure 6 plots the mean speech intelligibility scores versus talker angle for cases with a direct sound only and also for cases with added constant early reflections. Without early reflections, mean intelligibility decreases dramatically with talker head turning similar to Plomp and Mimpen's [15] results. However, when early reflections were included, there was only a small reduction in intelligibility even when the taker's head was turned 180 degrees (that is, facing away from the listener). These changes in intelligibility scores can be related to corresponding changes in S/N(A) values. Without early reflections

S/N(A) values changed from –2.2, to –8.7 and to –17.5 dB for angles of 0, 90 and 180 degrees. However when early reflections were included the corresponding S/N(A) values were –0.6, -3.9 and –5.4 dB.

The results in Figure 6 indicate that if there were no early reflections and the talker's head turned 90 degrees, listeners would find it very difficult to understand speech. If the talker's head was turned 180 degrees it would be completely impossible to understand speech without the benefit of early reflections. Clearly in many situations, it is only possible to understand speech in rooms because of the presence of early reflections.

III THE RELATIVE IMPORTANCE OF EARLY REFLECTION ENERGY IN TYPICAL ROOMS

The results presented in the previous section have confirmed the benefit of early reflections to obtaining adequate speech intelligibility in rooms for both impaired and non-impaired listeners. Early reflections were seen to effectively enhance the direct sound and also to compensate for reduced or weaker direct sound components. This section is intended to provide an initial examination of the relative magnitude of the early reflection energy that is found in actual rooms.

The direct sound energy will decrease with distance so that direct speech levels would frequently be unacceptably low at more distant listening positions in many rooms. In many cases this lack of direct sound energy is compensated for by added early arriving speech reflection energy. The experiments in the previous section confirm that the speech energy in early reflections is equally beneficial to intelligibility as similar speech energy in the direct sound. Thus we can expect speech energy arriving within the first 50 ms after the direct sound (E_{50}) to be useful to increasing intelligibility. If the direct sound is represented approximately by the energy arriving in the first 10 ms (E_{10}), then the ratio of E_{50}/E_{10} is a measure of the benefit provided by early arriving speech reflection energy. Thus the early reflection benefit (ERB) is proposed as a simple measure of the effectiveness of a room's acoustical design obtained by measuring the relative benefit of the early reflection energy.

$$\text{ERB} = 10\log\{E_{50}/E_{10}\}, \text{ dB}$$

Values of ERB were determined for several rooms used for unamplified speech communication. The ERB values were calculated from measures of the relative sound level or strength (G) for the first 10 ms (G_{10}) and the first 50 ms (G_{50}) of the impulse responses in each octave band. The relative sound level G is given by,

$$G = 10\log\left\{ \int_0^\infty p^2(t)dt \,/ \int_0^\infty p_A^2(t)dt \right\}, dB$$

where: p(t) is the pressure response in the measured impulse response and

$p_A(t)$ is the response to the same source in a free field at a distance of 10 m.

G_{10} and G_{50} are similarly calculated with the upper limit of the first integration set to 10 and 50 ms respectively.

The early reflection benefit (ERB) was then calculated as,

$$\text{ERB} = G_{50} - G_{10}, \text{ dB}$$

The current initial analyses were based on averages of the 1 to 4 kHz octave band G_{50} and G_{10} values because these frequencies are most important for the intelligibility of the weaker consonant sounds.

1. Boardroom

Acoustical data from a 390 m^3 boardroom in which impulse response measurements had been made were first considered. The room contained a large elongated table with seating for 22 people. Measurements were made in the unoccupied condition. Figure 7 plots ERB values versus source-receiver distance in this boardroom. These ERB values indicate that with added reflection energy the effective speech level is 2 to 8 dB greater than the direct sound level and that there is a systematic increase in the benefit from early reflections with increasing source-receiver distance. There is some scatter about the mean trend that may have been influenced by the presence of a large glass dome in the ceiling over the boardroom table.

These results suggest that it is quite possible for early reflection energy to increase the effective or useful speech level by as much as 8 dB. This is even greater than the maximum increase of +6 dB included in the speech intelligibility tests in simulated sound fields.

2. Four rooms for speech

The generality of the results from the boardroom were further explored by considering measurements in 4 other rooms that represent a range of conditions intended for unamplified speech communication. Descriptions of the boardroom mentioned above and the 4 other rooms are included in Table IV. They varied from a round table meeting room with a considerable amount of added sound absorption (used for teleconferencing) to a small auditorium and a small theatre.

ERB values were again calculated from averages of the 1 to 4 kHz octave-band G_{50} and G_{10} values obtained from impulse response measurements. All measurements were for unoccupied conditions. The resulting ERB values for positions in all 4 rooms are plotted versus source-receiver distance in Figure 8. There is again an approximate linear trend for ERB values to increase with increasing source-receiver distance. The maximum ERB values at the largest source-receiver distances are close to +9 dB.

These results indicate that the effect of early reflections in rooms is equivalent to an increase in the direct speech level of up to 9 dB. Increases tend to be greatest for listeners farthest from the talker. That is, rooms naturally tend to help intelligibility most where the help is most needed farthest from the source. Figure 9 compares measured impulse responses at two locations 'A' (closer to the source) and 'B' (farther from the source) in the Playhouse theatre. In Figure 8 the ERB at position 'B' is seen to be approximately 4 dB greater than at position 'A'. The impulse responses in Figure 9 clearly illustrate the much greater number of significant reflections at the more distant 'B' position. Presumably the amount of early reflection benefit can be influenced by the room acoustics design. At the rear of many rooms it may only be possible to understand speech because of the benefits of early reflections. Adding too much absorbing material or inappropriate shaping of reflecting surfaces could reduce the level of beneficial reflections and the related level of speech intelligibility.

The slope of the mean trend for the data from the boardroom in Figure 7 is steeper than that for the data from the 4 rooms in Figure 8. A brief examination of data from larger auditoria indicated even lower slopes. Thus the variation of benefits of early reflections may depend on the overall room size as well as the source-receiver distance. The ERB measure indicates the relative importance of early reflections relative to the direct sound and is not simply related to overall sound levels or strength values (G).

IV DESIGN EXAMPLE CALCULATIONS

This section explores the use of a modern room acoustics computer model to evaluate the effects of room design on the relative strengths of early reflections in rooms for speech. The ODEON (version 4.0) software was used to calculate impulse responses in a small 200 seat lecture theatre with a volume of just over 1000 m^3. The room had a steeply sloping seating area and is illustrated in Figure 10. It was similar to the room referred to as M50 in Figure 8 and Table IV.

In the first calculation example, the magnitude of early reflection energy was varied by varying the height of this room. The audience seating area was absorptive and the other surfaces were highly reflective. When the ceiling height (at the front of the room) was increased from 7 m to 10 m the volume increased from 1092 m^3 to 1777 m^3. Figure 11 shows the calculated ERB values versus source-receiver distance for both ceiling heights. The details of the design cases are summarized in Table V. The calculated results show that the lower ceiling increases the benefit from early reflections by up to 3 dB. Of course, the lower ceiling case also has a lower reverberation time (see Table V), which would also be preferable for speech.

Figure 12 compares calculated results for the same lower ceiling room shape with varied absorptive treatments. The 'live' case is the same as the lower ceiling case in the previous comparison. Only the audience area was absorptive for this case. The 'dead' case corresponds to adding material that is 60% absorptive to all of the walls and ceiling in addition to the absorptive audience area. The 'optimized' case corresponds to 70% absorptive material added to the shaded areas in Figure 10. It was hoped that this 'optimized' design would better control reverberant sound without unduly reducing early reflection energy. Figure 12 shows that for a given source-receiver distance the ‘dead’ case shows much smaller ERB values than the other two cases. That is, when there is too much absorptive treatment, the early reflection energy is reduced by up to 5 dB relative to the ‘live’ case. However, the ERB values for the ‘optimized’ case are almost the same as for the ‘live’ case. Thus, it seems possible to adjust the details of the design to maximize the relative magnitude of early reflection energy.

The measurement results in Figures 7 and 8 as well as the calculation results in Figures 11 and 12 show how the relative importance of early reflections tends to increase with increasing source-receiver distance. These types of results can be more completely understood when shown in terms of the total early-arriving level. Figure 13 plots results for the same calculations as in Figure 12 in terms of G_{50} values versus ERB values. For the ‘dead’ case G_{50} values decrease by 6 dB when moving from nearer to the source to the farthest receiver position. For the ‘live’ and ‘optimized’ cases G_{50} values decrease by only 3 dB with increasing source-receiver distance because there are larger amounts of beneficial early reflection energy which lead to larger ERB values. By comparing G_{50} values for the ‘dead’ case with corresponding values for the ‘optimized’ case, it is seen that optimizing the absorptive treatment can increase G_{50} values by as much as 4 dB. That is, the effective signal-to-noise ratio was increased by up to 4 dB. In many situations a 4 dB increase in effective signal-to-noise ratio would lead to significant increases in speech intelligibility ratings. (It is interesting to note that if the treated and untreated surface areas are reversed the same increased ERB does not occur).

Of course, there are many possible approaches to optimizing a particular room design and larger improvements in ERB values than indicated in these results may be possible.

However, the complete room acoustics design process must also include consideration of late-arriving sound levels. Increasing early reflection levels by decreasing the total amount of sound absorption will also tend to increase the later arriving sound levels. The optimum design will also have to consider the level of ambient noise because a truly optimum design must maximize the effective signal-to-noise ratio, which is expected to correspond to maximizing the useful-to-detrimental sound ratios. However, to explore all of the various combinations of room designs and ambient noise conditions is much beyond the scope of this paper.

V DISCUSSION

These new results have confirmed that early reflections of speech sounds are important for achieving adequate speech intelligibility in rooms and have approximately the same effect as increased direct sound energy for both non-impaired and impaired listeners. Where the direct sound is particularly weak, such as when the talker's head is turned away from the listener, or at positions towards the rear of many rooms, early reflections are essential to achieving adequate intelligibility.

The impaired subjects in this study were assumed to have peripheral hearing loss due to some combination of presbycusis (as the dominant factor) and noise induced hearing loss. However, the precedence effect is thought to function at a more central cognitive level [16]. There is therefore no reason to expect that subjects with mild to moderate peripheral hearing loss will not benefit fully from the precedence effect. Thus the current results, showing that both impaired and non-impaired listeners benefited equally from early reflections, are readily explained and lend credence to the belief that the precedence effect is due to higher level processing in the brainstem and auditory cortex.

The early reflection energy can be up to 9 dB greater than that of the direct sound. It is therefore more important than the direct sound and it is of primary importance that the room acoustics design attempts to maximize early reflection energy.

The common practice of focussing on reverberation time as the primary acoustical design parameter can distract us from the more important details of the acoustical design of rooms for speech. Obtaining an optimum reverberation time should not be thought of as a primary design goal but as something that is a consequence of the need to maximize early reflections without including excessive later arriving reflection energy. Ignoring the critical benefits of early reflections would lead to the conclusion that a reverberation time of 0 s. would be preferred. This misconception has been encouraged by experiments in which the positive effects of increased sound levels were either ignored or deliberately controlled [e.g. 17]. A minimum reverberation time is of course not a desirable goal and the added absorption required to achieve very short reverberation times could severely attenuate early reflections that make it possible for us to hear speech in many situations. The result of the real need to maximize early reflection energy without excessive later arriving reflections leads us to a related need for a non-zero optimum reverberation time. The actual optimum reverberation time will vary with room size and ambient noise level. However, there is not a precise optimum value for some particular condition but a relatively broad range of acceptable values [18].

Although later arriving reflections are undesirable, controlling them should not be the first priority. The first priority for the acoustical design of rooms for speech should be to maximize the total energy in the direct sound and early-arriving reflections of the speech sounds. (The 'live' case in Section IV is an example of only maximizing the early reflections). A second priority would be to ensure, that there is not excessive later-arriving reflection energy, usually by determining that conventional goals for optimum reverberation times are approximately met. (The 'optimum' case in the Section IV is a simple example of combining these first two design steps). However, it is much more important for designers to focus efforts on maximizing early reflections and hence increased ERB values than on small differences in reverberation times. A room that is slightly too reverberant is probably better for speech than one that is too dead and hence likely to be lacking in critical early reflection energy. Early reflections can increase the effective speech level by up to 9 dB, but even doubling the reverberation time would only increase late arriving sound by only about 3 dB and even this increase in unwanted late arriving sound can be insignificant relative to excessive ambient noise. Of course, achieving adequately low ambient noise levels is usually even more important than any aspect of room acoustics design [19].

VI CONCLUSIONS

The results of the new studies presented in this paper show that increased early reflection energy has the same effect on speech intelligibility scores as an equal increase in the direct sound energy. This was true for both non-impaired listeners and for listeners with mild to moderate hearing threshold shifts. These impaired listeners are thought to be representative of a significant portion of the population.

In typical rooms for speech, early reflection energy increases the effective S/N(A) by up to 9 dB. This would lead to very important increases in speech intelligibility scores in typical rooms for speech. That is, early reflections are important for good speech communication and in many situations where the direct sound is reduced they are essential to satisfactory speech communication. For example, when the talker is not facing the listener or for listeners near the rear of many rooms, it is only due to the benefits of early reflections that we are able to satisfactorily understand speech.

Room acoustics design for speech should focus first on maximizing early reflection energy. Although it is also important to avoid excessive reverberant sound, adding large amounts of absorption to achieve very short reverberations times may degrade intelligibility due to reduced early reflection levels.

The ratio of the early arriving energy in the first 50 ms of impulse responses, to the energy associated with the direct sound is termed the early reflection benefit, ERB, and is proposed as a useful measure of the effectiveness of a room's acoustical design.

While the hearing impaired listeners in this study benefited from added early reflections, there is a need to verify that this is also true for some other special groups. Further studies are required to determine whether younger and older listeners as well as more severely impaired listeners similarly benefit from early reflections.

Much of this is not totally new. In the introduction to a 1964 review paper Lochner and Burger stated [4], "…we know that reverberation time in itself gives very little indication of the suitability of a room for speech; rather, given the integration and masking characteristics of the hearing system, the intelligibility of speech will be determined by the reflection pattern of the room". Perhaps it is time to make a more serious effort to apply our understanding of the importance of early reflections to the design of rooms for speech.

ACKNOWLEDGEMENTS

The authors would like to thank Wai Lyn Wong who carried out many of the speech intelligibility experiments, and also the subjects who volunteered their time.

REFERENCES

[1] R.S. Shankland, "Architectural acoustics in America to 1930", JASA 61 (2) 250-254 (1977).

[2] H. Haas, "Uber den Einfluss des Einfachechos auf die Horsamkeit von Sprache", Acustica 1, 49-58 (1951).

[3] H. Wallach, E.B. Newman and M.R. Rosenzweig, "The precedence effect in sound localization", A, J. Psychol. 52, 315-336 (1949).

[4] J.P.A. Lochner and J.F Burger, "The influence of reflections on auditorium acoustics". J. Sound Vibr. 1 (4) 426-454 (1964).

[5] H.G. Latham, "The Signal-to-Noise Ratio for Speech Intelligibility –An Auditorium Acoustics Design Index", App. Acoust. Vol. 12, pp. 253-320, (1979).

[6] J.S. Bradley, "Predictors of Speech Intelligibility in Rooms", J. Acoust. Soc. Am., Vol. 80, pp. 837-845, (1986).

[7] J.S. Bradley, "Speech Intelligibility Studies in Classrooms", J. Acoust. Soc. Am., Vol. 80, No. 3, 846-854, (1986).

[8] J.S. Bradley, "Relationships Among Measures of Speech Intelligibility in Rooms", J. Aud. Eng. Soc., Vol. 46, No. 5, pp. 396-405 (1998).

[9] F.H. Bess, "Classroom Acoustics: An Overview", Volta Review, vol. 101, no. 5, pp. 1-14 (2001).

[10] A.K. Náběl ek, and L. Robinette, "Influence of the precedence effect on word identification by normally hearing and hearing-impaired subjects", J. Acoust. Soc. Am. Vol. 63, no. 1, pp. 187-194 (1978).

[11] Anon, “American National Standard Acoustical Performance Criteria, Design Requirements And Guidelines For Schools”, ANSI Draft Standard S12.60-200X (2001).

[12] Anon, “Acoustics – Threshold of hearing by air conduction as a function of age and sex for otologically normal persons”, ISO 7029 – 1984 (E), International Organisation for Standardisation .

[13] M. Picard, and J.S. Bradley, “Revisiting Speech Interference and Remedial Solutions in Classrooms”, Audiology, Journal of Auditory Communication, vol. 40, no. 5, pp. 221-244, (2001).

[14] W.T. Chu, and A.C.C. Warnock, “Directivity of Human Talkers”, J. Acoust. Soc. Am., vol. 110, no. 5, part 2, p. 2664 (2001).

[15] R. Plomp and A.M. Mimpen, “Effect of the Orientation of the Speaker’s Head and Azimuth of a Noise Source on the Speech-Reception Threshold for Sentences”, Acustica, vol. 48, pp. 325-328 (1981).

[16] R.Y. Litovsky, and H.S. Colburn, “The precedence effect”, J. Acoust. Soc. Am., vol. 106, no. 4, pp. 1633-1654 (1999).

[17] A.K. Nábělek, and J.M. Pickett, “Reception of consonants in a classroom as affected by monaural and binaural listening, noise, reverberation, and hearing aids”, J. Acoust. Soc. Am., vol. 56, no. 2, pp. 628-639 (1974).

[18] R.D. Reich, and J.S. Bradley, “Optimizing Classroom Acoustics Using Computer Model Studies”, Canadian Acoustics, Vol. 26, No. 3, pp. 15-21,1998.

[19] J.S. Bradley, R.D. Reich, and S.G. Norcross, “On the Combined Effects of Signal-to-Noise Ratio and Room Acoustics on Speech Intelligibility”, J. Acoust. Soc. Am. Vol. 106, No. 4, pp. 1820-1828, (1999).

Table I. Orientation of loudspeakers relative to the listener. Horizontally straight ahead of the listener at ear level is 0 degrees in both planes.

Loudspeaker	**Horizontal angle, degrees**	**Vertical angle, degrees**
1. Centre low (direct sound)	0	0
2. Centre high	0	25
3. Left low	-32	0
4. Right low	+32	0
5. Left high	-37	28
6. Right high	+37	28
7. Far left	-115	0
8. Far right	+115	0

Table II. Summary of acoustical measures for the Direct only and Direct+early reflections sound field cases. (Noise level 47.6 dBA).

	Speech level dBA	**S/N(A) dB**	**C_{50} (1 kHz) dB**	**U_{50} (1 kHz) dB**	**STI**
Direct only					
1	41.8	-5.8	21.9	-5.7	0.32
2	45.4	-2.2	19.3	-2.0	0.44
3	48.8	1.2	19.9	1.5	0.55
4	52.0	4.4	20.1	4.6	0.66
Direct+early					
1	41.8	-5.8	21.9	-5.7	0.32
2	44.9	-2.7	21.1	-2.6	0.41
3	48.1	0.5	21.6	0.5	0.51

Table III. Summary of acoustical measures for the Direct+early+reverberant reflections sound field cases. 'Speech Level' is the total speech level including direct sound, early reflections and late arriving speech sounds. 'Effective Speech Level' includes only the direct and early reflection speech energy. (Noise level 47.6 dBA).

	Speech Level dBA	**S/N(A) dB**	**Effective Speech Level dBA**	**Effective S/N(A) DB**	**C_{50} (1 kHz) dB**	**U_{50} (1 kHz) dB**	**STI**	**EDT (1 kHz) s**	**RT (1 kHz) s**
Direct+early+reverb									
1	45.5	-2.1	41.8	-5.8	1.4	-5.7	0.31	1.2	1.1
2	47.2	-0.4	44.9	-2.7	4.1	-3.0	0.38	0.9	1.1
3	49.6	2.0	48.1	0.5	7.2	0.0	0.46	0.5	1.1

Table IV. Descriptions of the rooms from which acoustical measurement data were obtained.

Name	**Description**	**Number seats**	**Volume, m^3**	**RT(1 kHz), s**
Boardroom	Boardroom.	22	390 m^3	1.4
Ridgemont	School auditorium	750	2800 m^3	1.2
M50	Lecture theatre	200	1020 m^3	0.6
Playhouse	Small theatre	690	7000 m^3	1.0
RCMP	Round table teleconferencing/ meeting room	15	867 m^3	0.26

Table V. Details of the calculation examples for the room illustrated in Figure 10. ‘Treatment’ corresponds to the shaded area in Figure 10. Audience areas had a 1 kHz absorption coefficient of 0.8.

Case	Height, m	Volume, m^3	Absorption coefficients (1 kHz)		RT (1 kHz), s
			Walls & ceiling	Treatment	
High-live	10	1777	0.03	None	3.5
Live	7	1092	0.03	None	1.9
Dead	7	1092	0.6	None	0.34
Optimized	7	1092	0.03	0.6	1.1

List of Figure Titles

Figure 1. Illustration of the initial 0.2 s of the measured impulse response of a simulated sound field including a direct sound and 7 early reflections followed by a reverberant decay.

Figure 2. Average measured hearing loss (HL), (as per ANSI S3.6 1996), ± 1 standard deviation for the non-impaired and impaired groups of listeners.

Figure 3. Mean intelligibility scores ± 1 standard deviation for each sound field condition for all listeners for 'direct only' and 'direct+early' reflections cases.

Figure 4. Mean intelligibility scores for each sound field condition for 'direct only' and 'direct+early' reflections cases plotted for each hearing loss group.

Figure 5. Mean intelligibility scores for each sound field condition for 'direct only', 'direct+early' and 'direct+early+reverberant' reflections cases plotted separately for each hearing loss group. In calculating Effective S/N(A) values, the reverberant speech sound was excluded from the signal level.

Figure 6. Mean speech intelligibility scores versus horizontal angle of the talker's head relative to directly pointing towards the listener for sound fields with and without added early reflections.

Figure 7. Measured values of the early reflection benefit in a boardroom.

Figure 8. Measured values of the early reflection benefit in four rooms intended for unamplified speech communication.

Figure 9. Measured broadband impulse responses in terms of pressure squared versus time at locations ‘A’ and ‘B’ from Figure 8 in the Playhouse theatre. They are normalised to the same direct sound amplitude and the same start time to facilitate the comparison.

Figure 10. Sketch of the room used in the calculation examples showing the ‘optimized’ configuration with a room volume of 1092 m^3. The eight receiver positions (solid circles) and the position of the source (open circle) having the directivity of a human talker are also shown.

Figure 11. Calculated 1 kHz ERB values versus source-receiver distance for the ‘high-live’ (higher ceiling) and ‘live’ (lower ceiling) room design cases. (Table V includes descriptions of the designs).

Figure 12. Calculated 1 kHz ERB values versus source-receiver distance for the ‘live’, ‘dead’ and ‘optimized’ room design cases. (Table V includes descriptions of the designs).

Figure 13. Plot of the relative early sound level, G_{50} versus ERB for the 1 kHz results shown in Figure 12.

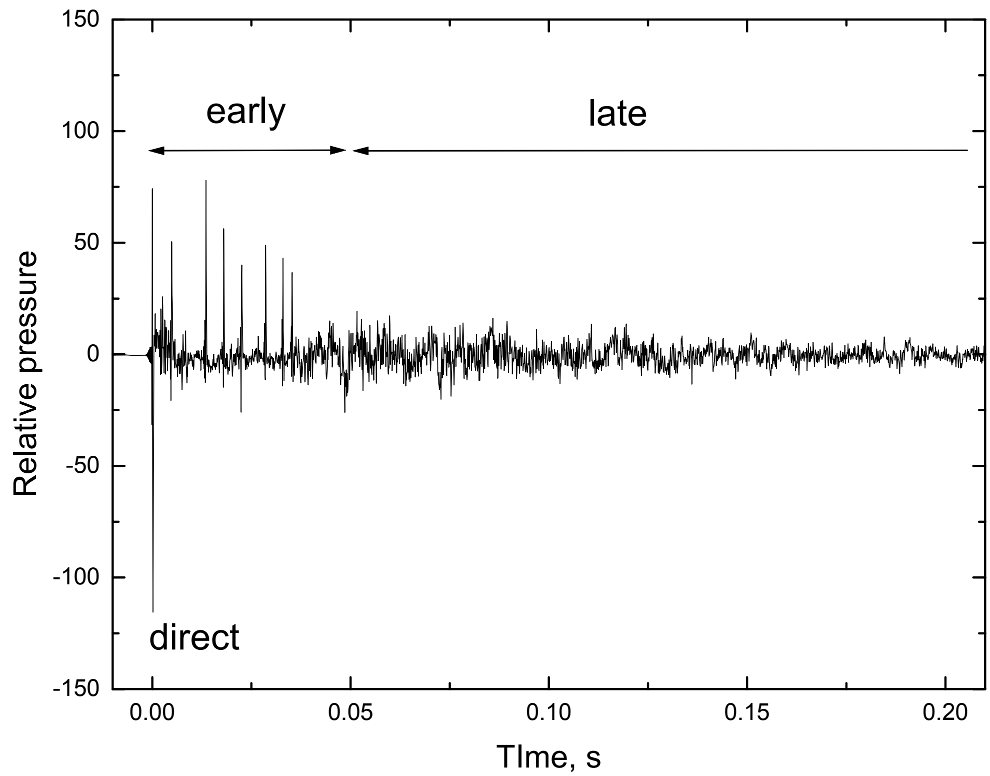

Figure 1. Illustration of the initial 0.2 s of the measured impulse response of a simulated sound field including a direct sound and 7 early reflections followed by a reverberant decay.

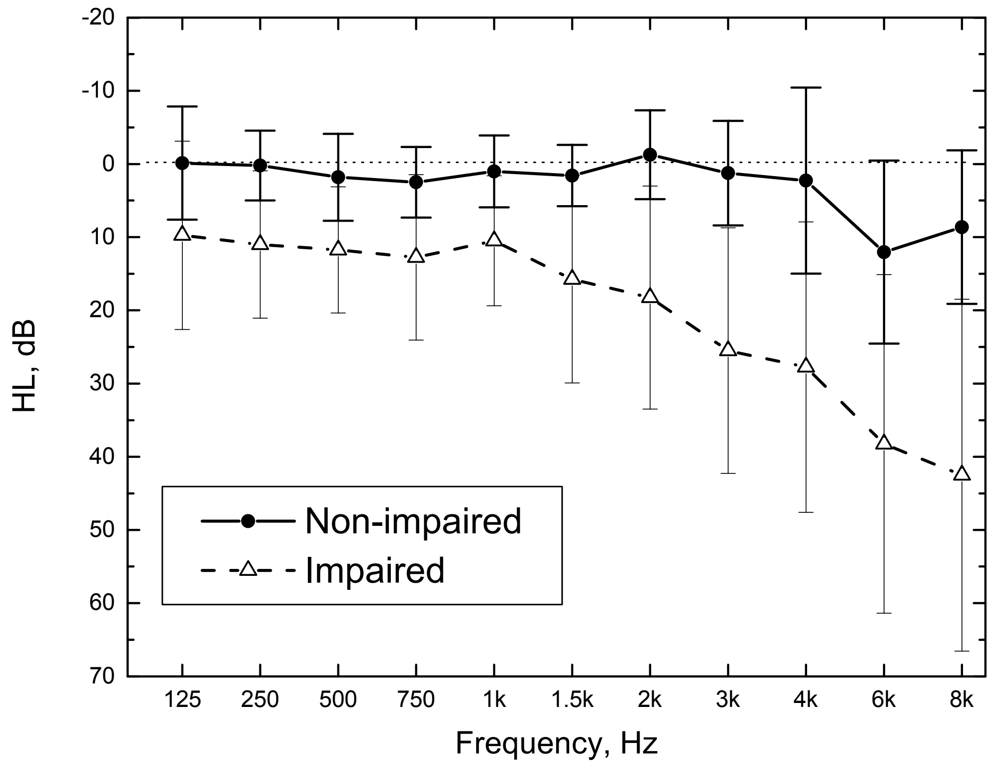

Figure 2. Average measured hearing loss (HL), (as per ANSI S3.6 1996), ± 1 standard deviation for the non-impaired and impaired groups of listeners.

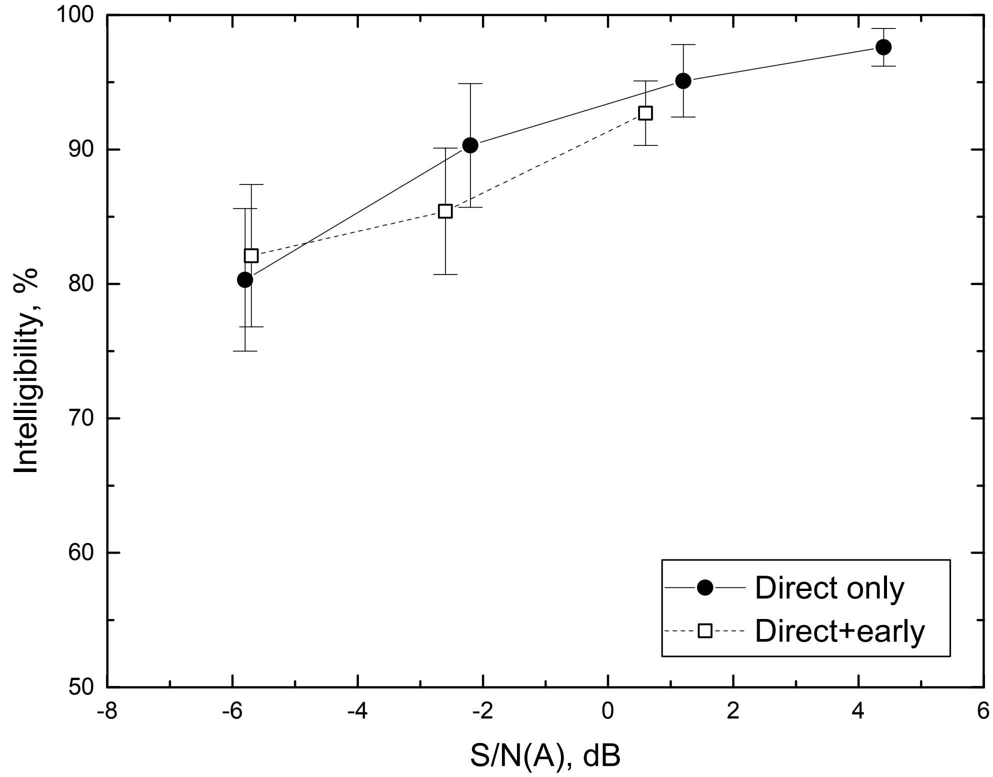

Figure 3. Mean intelligibility scores ± 1 standard deviation for each sound field condition for all listeners for 'direct only' and 'direct+early' reflections cases.

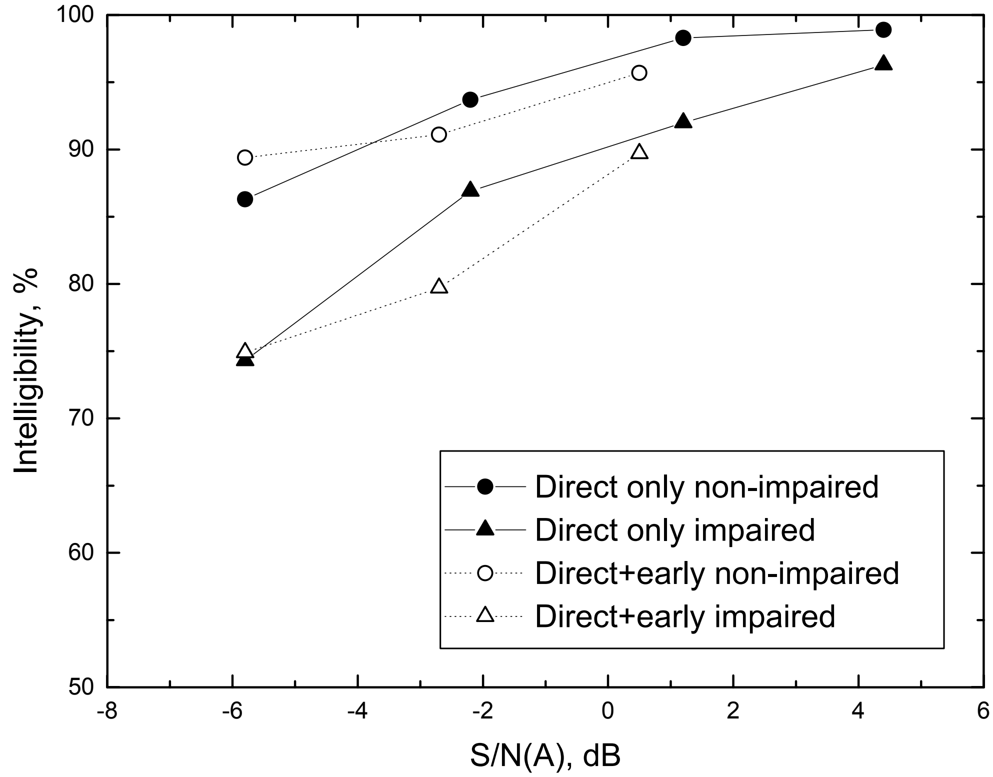

Figure 4. Mean intelligibility scores for each sound field condition for ‘direct only’ and ‘direct+early’ reflections cases plotted for each hearing loss group.

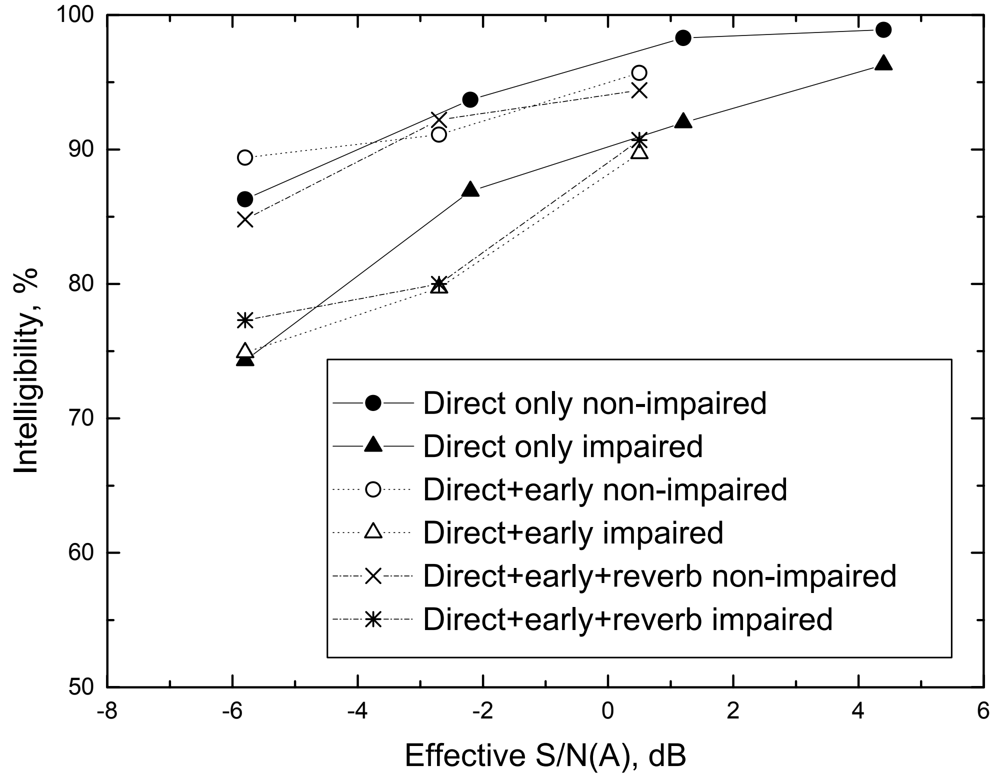

Figure 5. Mean intelligibility scores for each sound field condition for 'direct only', 'direct+early' and 'direct+early+reverberant' reflections cases plotted separately for each hearing loss group. In calculating Effective S/N(A) values, the reverberant speech sound was excluded from the signal level.

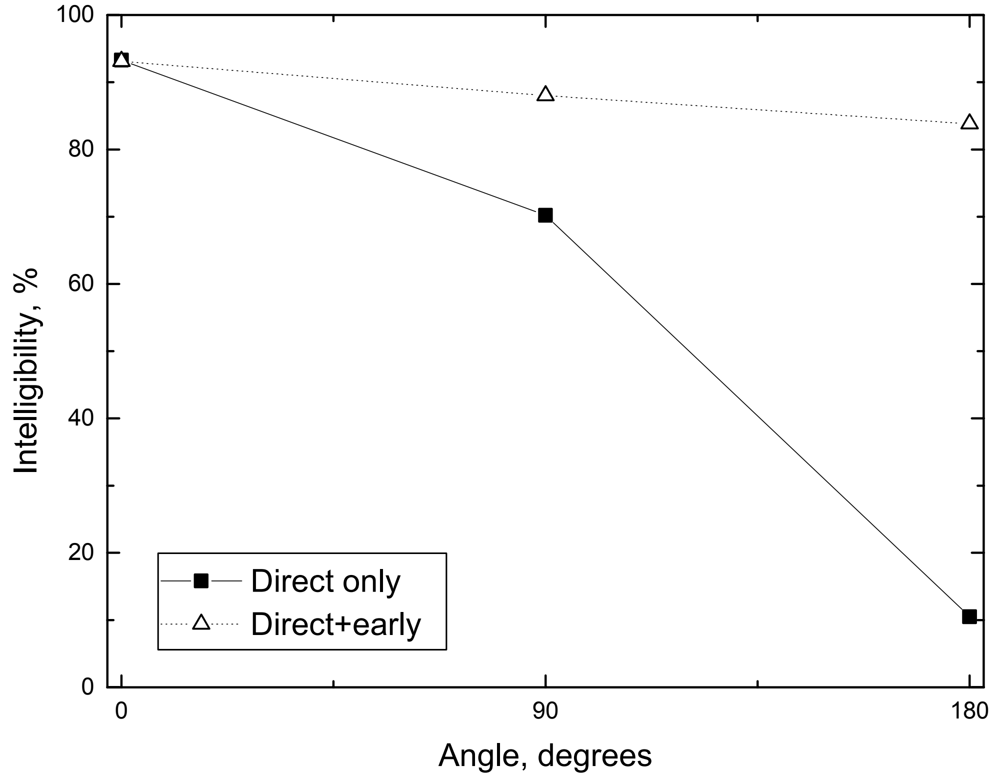

Figure 6. Mean speech intelligibility scores versus horizontal angle of the talker's head relative to directly pointing towards the listener for sound fields with and without added early reflections.

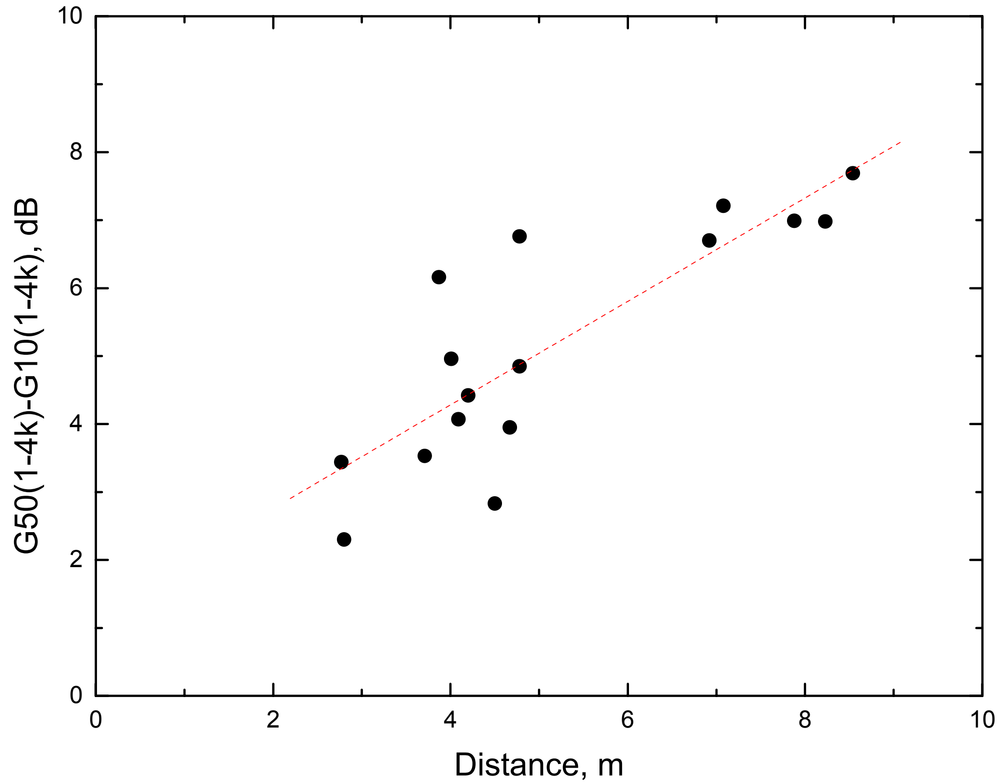

Figure 7. Measured values of the early reflection benefit in a boardroom.

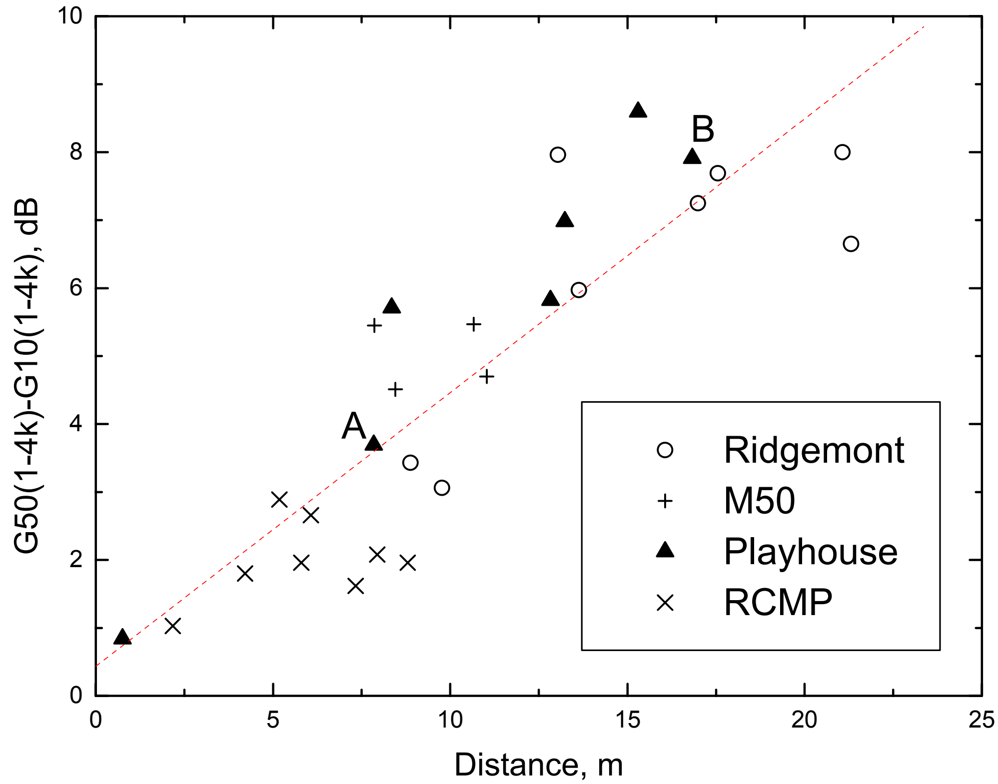

Figure 8. Measured values of the early reflection benefit in four rooms intended for unamplified speech communication.

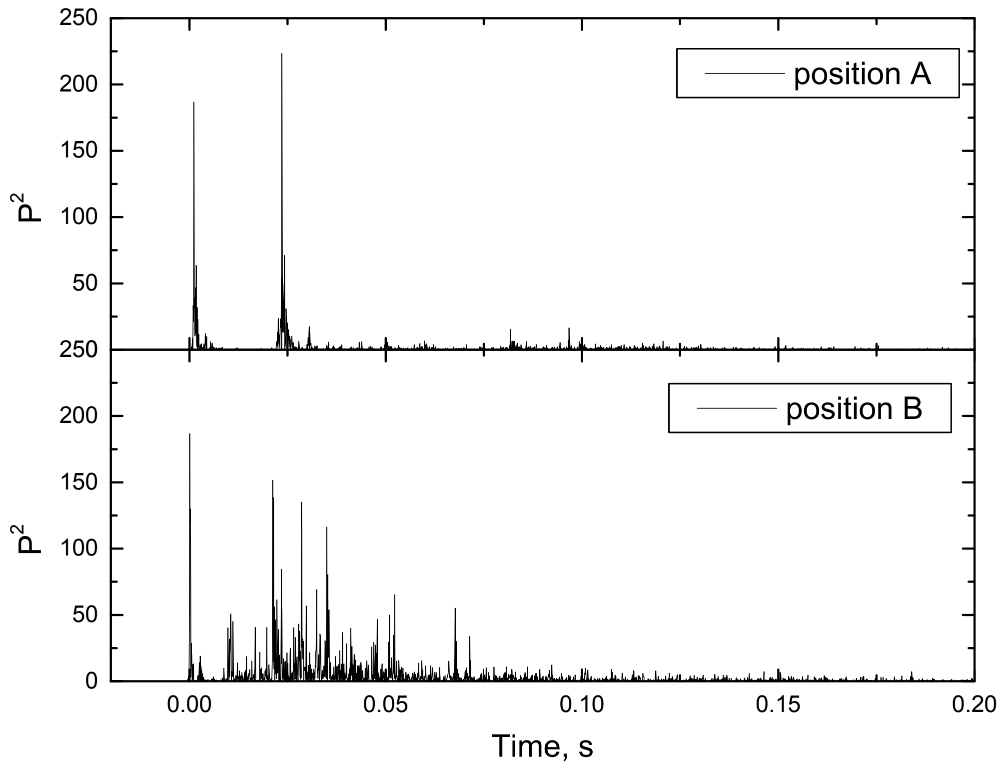

Figure 9. Measured broadband impulse responses in terms of pressure squared versus time at locations ‘A’ and ‘B’ from Figure 8 in the Playhouse theatre. They are normalised to the same direct sound amplitude and the same start time to facilitate the comparison.

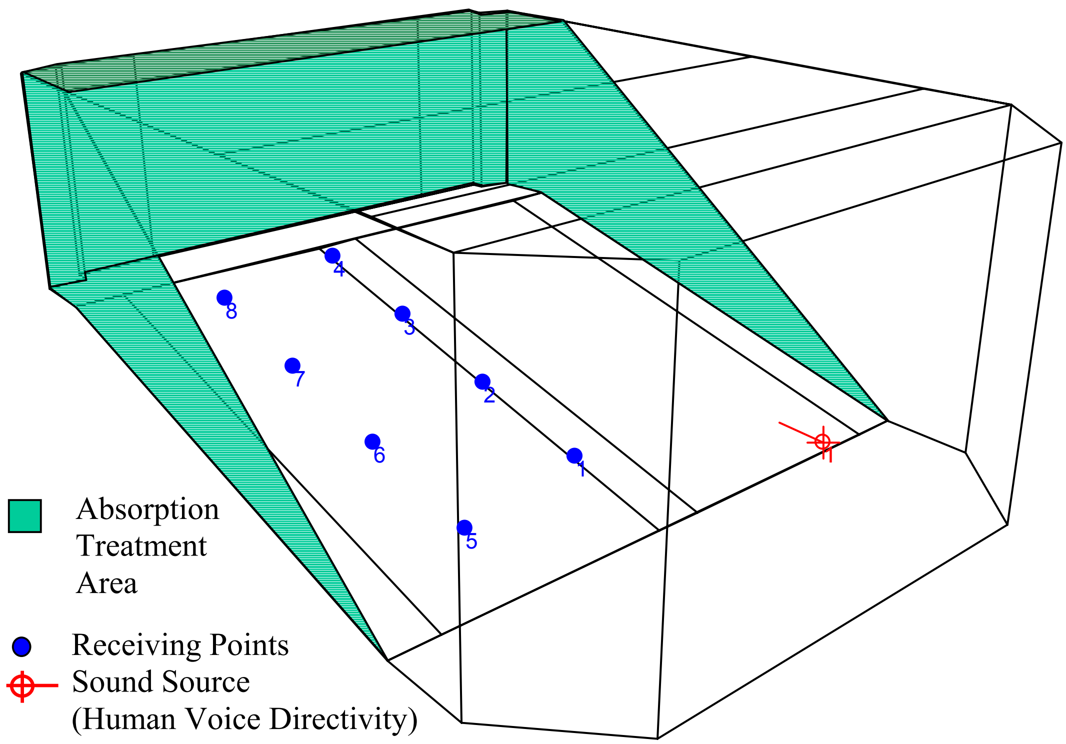

Figure 10. Sketch of the room used in the calculation examples showing the 'optimized' configuration with a room volume of 1092 m^3. The eight receiver positions (solid circles) and the position of the source (open circle) having the directivity of a human talker are also shown.

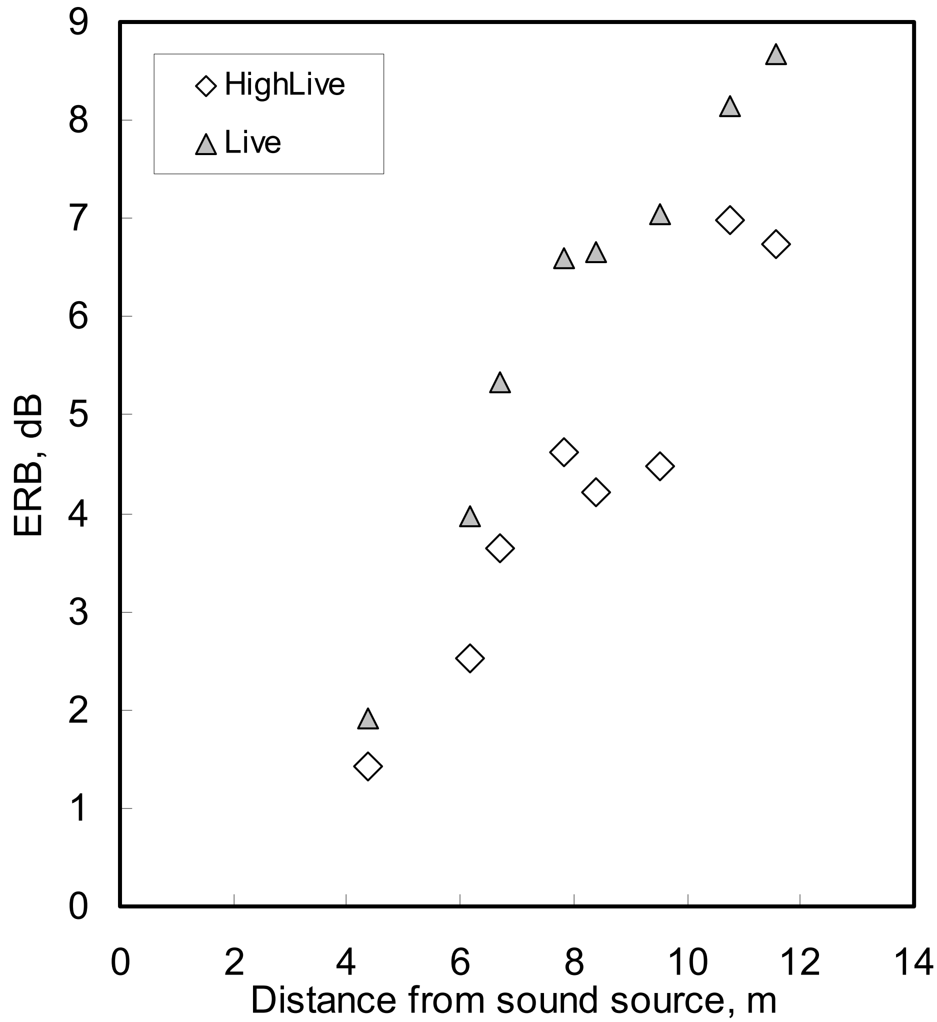

Figure 11. Calculated 1 kHz ERB values versus source-receiver distance for the 'high-live' (higher ceiling) and 'live' (lower ceiling) room design cases. (Table V includes descriptions of the designs).

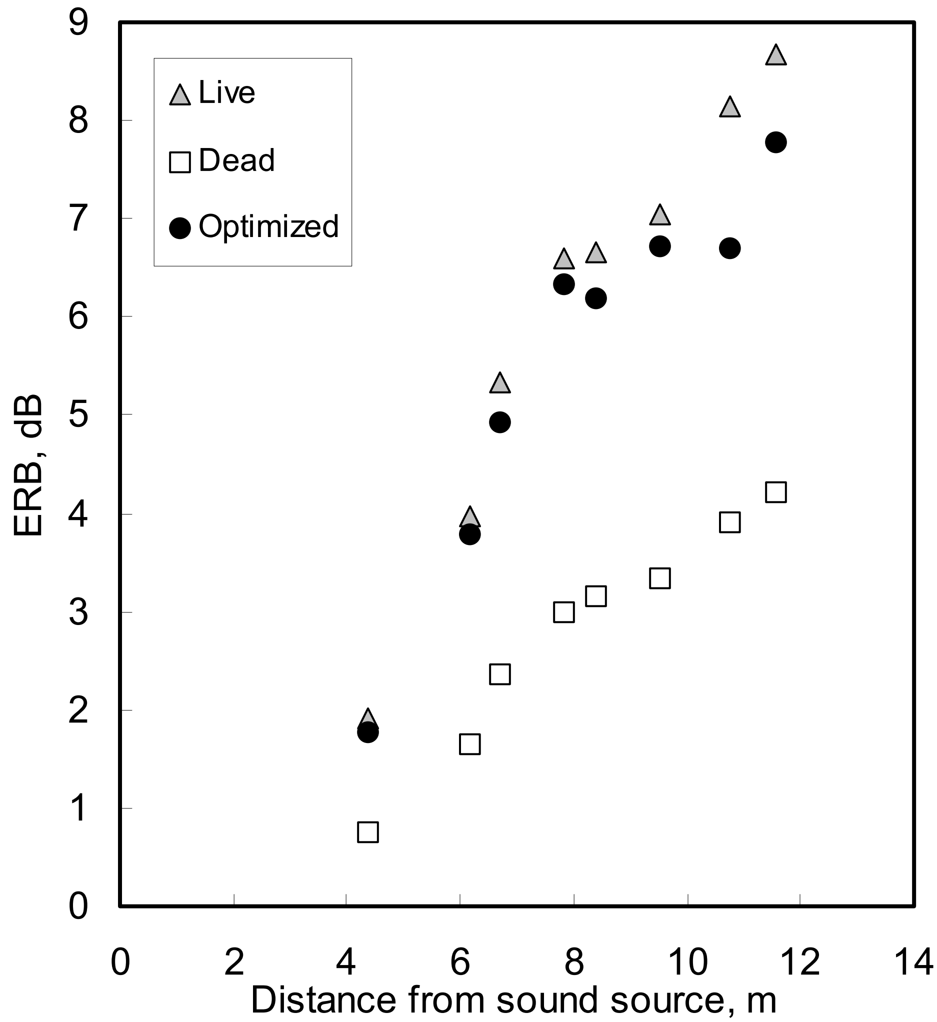

Figure 12. Calculated 1 kHz ERB values versus source-receiver distance for the ‘live’, ‘dead’ and ‘optimized’ room design cases. (Table V includes descriptions of the designs).

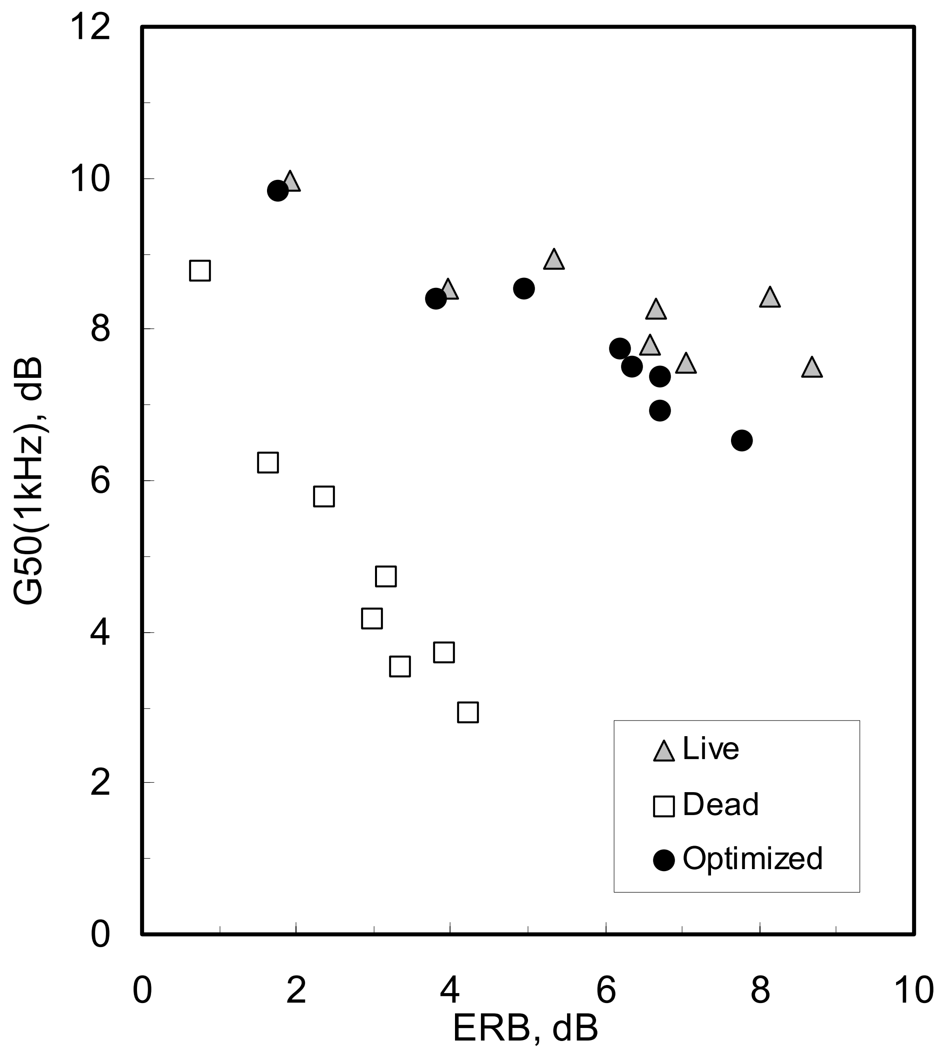

Figure 13. Plot of the relative early sound level, G_{50} versus ERB for the 1 kHz results shown in Figure 12.