

INVESTIGATION OF THE OPTIMUM ACOUSTICAL CONDITIONS FOR SPEECH USING AURALIZATION

Wonyoung Yang and Murray Hodgson

School of Occupational and Environmental Hygiene, University of British Columbia, 2206 East Mall, BC, V6T 1Z3
wyang@interchange.ubc.ca

1. INTRODUCTION

Speech intelligibility is known to be mainly determined by the signal-to-noise level difference at a listener, and reverberation. It is directly related to signal-to-noise level difference and is inversely related to the reverberation time. However, their interaction results in a complicated situation in real rooms. Increased reverberation benefits speech intelligibility by increasing speech levels. Noise is also affected by the reverberation time, as are the speech levels. Thus, the spatial relationship between a listener and the sound sources – both speech and noise – affects the optimal reverberation for speech intelligibility in rooms [1].

In this project, an experimental approach to identify the optimal reverberation time in an idealized room, and validate theoretical prediction, considering babble noise sources inside the room, is presented using auralization. Realistic optimal reverberation times are found using speech-intelligibility tests with normal-hearing and hard-of-hearing subjects. The best metric predicting speech intelligibility is presented for both normal-hearing and hard-of-hearing groups.

2. METHOD

2.1 Subjects

Subject groups for the study were normal-hearing and hard-of-hearing people with a mild to moderate sensorineural hearing loss. The hearing-loss criterion for the hard-of-hearing subjects was less than 20 dB at 250 Hz and 500 Hz and greater than 30 dB at 1 kHz to 8 kHz. Twenty-four normal-hearing subjects (mean age=27) and ten hard-of-hearing subjects (mean age=58) completed the tests.

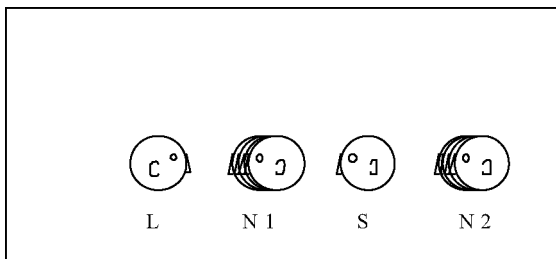


Fig. 1. Room elevation showing the speaker (S), listener (L) and noise-source (N) positions.

2.2 Test materials and simulated sound fields

The Modified Rhyme Test (MRT) was used as the speech-intelligibility test. The MRT lists were processed through the CATT-Acoustics room-acoustical prediction and auralization program with 4talker babble noise. The volume of the virtual room used for the simulation was 385 m³ (11 m x 7 m x 5 m). Fig. 1 shows the elevation of the virtual room, and the relative positions of the listener, the speaker, and the noise sources. In this study, the object was to model an idealized room with exponential sound decay. The effects of the different distributions of various surface materials on the walls, floor and ceiling were excluded by using the same absorption coefficients and diffusion coefficients for all octave bands and for all of the surfaces.

A total of 16 different sound fields were created, consisting of the combinations of 2 different speech- and noise-source output levels (SNS = 0 and +5 dB), 4 different reverberation times (RT = 0, 0.2, 0.4, and 0.8 s), and 2 different positions of the noise source (see Fig. 1). The speech-to-noise level difference at the listener (SNR) varied from -6 dB to 8.5 dB.

The completed auralization test materials were transferred to a compact disc for presentation using a CD player. Hearing screening tests were done prior to the main speech-intelligibility test to identify the hearing categories of subjects. The tests were processed through a Sony MDR V600 headphone in a soundproof room.

3. RESULTS

Fig. 2 shows the mean speech-intelligibility scores with 95 % confidence intervals. The speech-intelligibility scores were analyzed statistically by analysis of variance (ANOVA). Between normal-hearing and hard-of-hearing groups, there was a significant difference ($\alpha = 0.05$) in the speech-intelligibility scores. However, there was no relationship between the mean difference and SNR.

When the distance of the listener from the noise is farther than the distance from the speaker (N2), there was an overall trend for speech intelligibility to decrease with increasing RT for both normal-hearing and hard-of-hearing groups (Figs. 2b and d); i.e. the optimum reverberation time was zero. At the listener position, the noise level decreased more than the speech level in this case. Therefore, reverberation had a detrimental effect on speech intelligibility. When the noise was positioned between

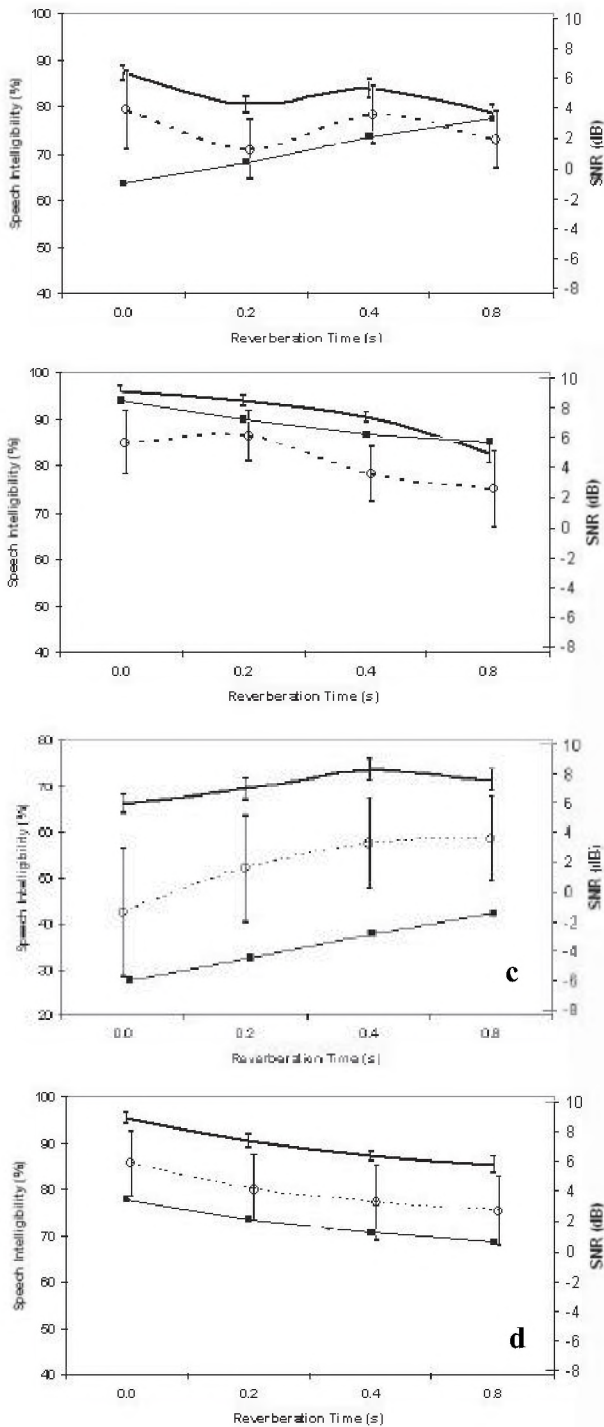


Fig. 2 Mean speech-intelligibility scores with 95% confidence intervals and signal-to-noise ratios (a: SNS=5 dB + N1; b: SNS=5 dB + N2; c: SNS=0 dB + N; d: SNS=0dB + N2. The lines have been offset to avoid overlapping. —: Normal-hearing; - - - : Hard-of-hearing; ___ : SNR).

Table 1. Strength of third-order polynomial regression

R ²	Normal hearing		Hard-of-hearing	
	Calculated	Measured	Calculated	Measured
U ₈₀	84.8	83.3	91.3	91.3
U ₇₀	86.5	84.3	92.9	91.5
U ₆₀	87.6	85.1	93.0	91.5
U ₅₀	88.4	85.9	92.2	91.2
U ₄₀	88.2	86.2	89.6	89.9
U ₃₀	86.2	86.7	83.4	87.4
SNR	-	72.7	-	84.9

the listener and the speaker (N1), there were significant differences among the RTs ($\alpha = 0.05$). For normal-hearing with SNS = 0 dB (Fig. 2c), the speech intelligibility increased with increasing RT until 0.4 s then decreased ($\alpha = 0.05$). For hard-of-hearing subjects with SNS = 0 dB, there were similar trends as for normal-hearing subjects. In this case, the speech level decreased more than the noise level at the listener's position. Therefore, reverberation increased the speech-intelligibility scores. When the SNS was 5 dB (Fig. 2a), unexpected results occurred. The peak occurred at RT = 0.4 s with SNS = 0 dB; however, at RT = 0 s, there was another peak.

For each sound-field configuration, useful-to-detrimental ratios were calculated from predicted impulse responses combined with speech and noise levels. As early-time limits of the useful and detrimental energies, 30, 40, 50, 60, 70 and 80 ms were used in the unweighted frequency spectrum. Third-order polynomial regression analyses were done with the mean MRT scores of each sound field configuration. Table 1 shows the strength of the relationship, R², between each measure and speech intelligibility for normal and hard-of-hearing people. Since the form of the fit and the number of data points was the same in every case, the success of each measure can be compared in terms of the corresponding R² values. In both calculation methods, there were similar trends. U₅₀ values were best suited for predicting speech intelligibility for the normal-hearing group. However, for the hard-of-hearing group, the early-time limits were slightly higher than those for the normal-hearing group.

4. CONCLUSION

Optimal reverberation times depended on the signal-to-noise level difference and the spatial relationship between the listener and the sound sources for both normal and hard-of-hearing subjects, as predicted by theory. Hard-of-hearing people needed high signal-to-noise level difference rather than shorter reverberation time to achieve better speech intelligibility. They required longer early energy than normal hearing did.

REFERENCE

1 Hodgson and Nosal, "Effect of noise and occupancy on optimal reverberation times for speech intelligibility in classrooms", *J. Acoust. Soc. Am.*, 111(2) 931-939.