Proceedings of The Institute of Acoustics

THE PREDICTION OF SPEECH INTELLIGIBILITY DURING DESIGN

HOWARD G. LATHAM

HERIOT-WATT UNIVERSITY

INTRODUCTION

Speech Intelligibility tests are well established as reliable subjective measures of auditorium speech acoustics. Reverberation Time was thought to be a design guide for Speech Intelligibility but it is now known that there is little correlation between Speech Intelligibility and Reverberation Time (1).

A significant advance in the prediction of Speech Intelligibility was made by the Signal-to-Noise Ratio of Lochner and Burger (2). This is a ratio of useful to detrimental speech energy, according to the measured integration characteristics of hearing. They obtained a high correlation with Speech Intelligibility when a constant-level speech-shaped noise was used to mask differences in ambient background noise level (3). Their procedure left unanswered the question of the significance of ambient background noise to Speech Intelligibility in reverberant conditions.

A modified Signal-to-Noise Ratio has now been developed which accounts for ambient background noise (4). This was achieved by a correlation analysis of objective and subjective measurements in a range of existing auditoria. The Speech Intelligibility Test used was Fairbanks' Rhyme Test (5), modified for auditorium testing by embedding the test words in continuous speech passages. The best-fitting correlation with Speech Intelligibility provided an empirical basis for the modification to the Signal-to-Noise Ratio (see Figure 1). The time and frequency varying characteristics of ambient background noise were interpreted by the L_{10} index and by the Preferred Noise Criterion (P.N.C.) curves. Both the ambient background noise level and the reflection pattern analysis are based on the 1 KHz whole octave frequency band. The modified Signal-to-Noise Ratio and its correlation with Speech Intelligibility provides a practical basis for an auditorium design guide (see 6).

SPEECH INTELLIGIBILITY PREDICTION : MEASUREMENT PROCEDURE

To predict Speech Intelligibility, four objective measurements must be quantified, and the Signal-to-Noise Ratio computed. The evaluation of an existing auditorium permits direct measurement but evaluation during design requires the application of mathematical or physical modelling techniques.

(i) Anechoic Speech Level at Source

Here we are concerned with the level of speech at source. For convenience a reference point at 1 m is used to measure the long-time average sound pressure level of the complete test passages S_R (dB re 2×10^{-5} Pa in 1 KHz whole octave band). (To standardise the source level for comparison purposes 62.5 dB has been used.)

(ii) Attenuation Factor

Here we are concerned with the attenuation of speech energy in the auditorium between the source reference point and the listener. To measure this factor A a continuous white-noise is emitted from the source position and measurements of sound pressure level (dB re 2 x 10^{-5} Pa 1 KHz band) are made at the reference L_R and listener L_L locations.

$$A = L_R - L_L \quad \text{dB re } 2 \times 10^{-5} \text{ Pa} \qquad \text{(Eq. 1)}$$

(iii) Reflections Ratio

This measurement concerns the distribution of reflections in the auditorium according to the method of Lochner and Burger. The impulse response at a position in the auditorium to a spark is processed by computer — the program squares, integrates, and applies Lochner and Burger's pressure-and-time dependent weighting factor α for the integration characteristics of hearing for speech, to give

$$R_f = 10 \log \frac{\int_0^{95} \alpha \; P^2_{(f,t)} dt}{\int_{95}^{1500} P^2_{(f,t)} dt} \text{ dB} \qquad \text{(Eq. 2)}$$

where R_f is the Reflections Ratio (dB) in the 1 KHz band, P is the absolute pressure (2 x 10^{-5} Pa), f is the 1 KHz whole octave frequency band, and t is time (msec re direct sound).

(iv) Ambient Background Noise Level

This measurement can obviously only be made accurately in a completed auditorium. However an estimate may be made during design to investigate the range of possible effects on Speech Intelligibility of expected noise sources. The ambient background noise level is expressed as an L_{10} P.N.C. level.

SPEECH INTELLIGIBILITY PREDICTION : COMPUTATIONAL PROCEDURE

The procedure to derive Signal-to-Noise Ratio and thus Speech Intelligibility from the measured values consists of five stages of computations.

(i) Total Speech Level at Listener

The total speech level at a listener S_L (dB re 2 x 10^{-5} Pa) is given by the difference between the anechoic speech level S_R and the attenuation factor A

$$S_L = S_R - A \text{ dB re } 2 \times 10^{-5} \text{ Pa} \qquad \text{(Eq. 3)}$$

(ii) Energy Density

As we are working in energy terms, the average speech and noise sound pressure levels L (dB re 2×10^{-5} Pa) are re-expressed in energy density terms E_x (joules/m^3) using the relation

$$E_x = 2.869 \times 10^{-15} \times 10^{(L/10)} \text{ joules/m}^3 \qquad \text{(Eq. 4)}$$

Thus we now have E_t the total speech energy, and E_n the equivalent energy of ambient background noise (both joules/m^3).

(iii) Speech Energy Fractions

The Reflections Ratio (Eq. 2) can be expressed more simply in terms of the energy fractions: useful energy E_u and detrimental energy E_d (joules/m^3)

$$R_f = 10 \log \frac{E_u}{E_d} \text{ dB} \qquad \text{(Eq. 5)}$$

The sum of the useful and detrimental energy fractions can be regarded as practically equal to the total energy level at the listener,

$$E_t = E_u + E_d \text{ joules/m}^3 \qquad \text{(Eq. 6)}$$

Absolute values for the energy fractions can therefore be derived from the known values for total energy and reflections ratio:

$$E_u = E_t / (10^{-(R_f/10)} + 1) \text{ joules/m}^3 \qquad \text{(Eq. 7)}$$

$$E_d = E_t / (10^{-(R_f/10)} + 1) \text{ joules/m}^3 \qquad \text{(Eq. 8)}$$

(iv) Signal-to-Noise Ratio

The derived energy values can now be combined to give the Signal-to-Noise Ratio S/N (dB) using the relation

$$S/N = 10 \log \frac{E_u}{E_d + E_n} \text{ dB} \qquad \text{(Eq. 9)}$$

(v) Speech Intelligibility

The predicted Speech Intelligibility SI (% Fairbanks Rhyme Test) is derived from the computed Signal-to-Noise Ratio S/N using the curve in Figure 1.

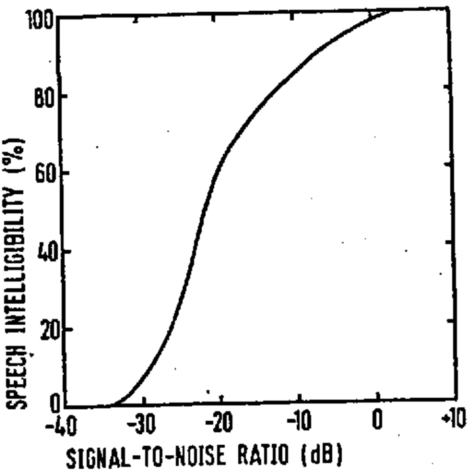

Figure 1 Speech Intelligibility v. Signal-to-Noise Ratio

CONCLUSION

The procedure outlined in this paper can be used to predict Speech Intelligibility in auditoria. It emphasises that auditorium design should be considered in terms of both reflection pattern and ambient background noise.

REFERENCES

(1) H.G. LATHAM 1976 Proceedings of the Institute of Acoustics, Edinburgh, paper 2 - 11. An investigation into the speech propagation properties of auditoria.

(2) J.P.A. LOCHNER and J.F. BURGER 1961 Acustica 11, 195 - 200. The intelligibility of speech under reverberant conditions.

(3) J.P.A. LOCHNER and J.F. BURGER 1964 Journal of Sound and Vibration 4, 426 - 454. The influence of reflections on auditorium acoustics.

(4) H.G. LATHAM 1977 Proceedings of the Ninth International Congress on Acoustics, Madrid, paper B5. The design of auditoria for the optimisation of speech intelligibility.

(5) G. FAIRBANKS 1958 Journal of the Acoustical Society of America 30, 596 - 601. Test of phonemic differentiation: the rhyme test.

(6) P.I. NEWMAN 1978 Proceedings of the Institute of Acoustics Meeting on Theatre Acoustics, Cambridge. The Grange Arts Centre, Oldham.