Proceedings of The Institute of Acoustics

SPEECH IN THEATRES : WHAT ARE THE IMPORTANT CONSIDERATIONS ?

M. Barron

Department of Architecture, Scroope Terrace, Trumpington St. Cambridge CB2 1PX

INTRODUCTION

To be able to predict speech intelligibility in theatres, at least six influences need to be considered:

Theatre impulse response
Speaker orientation
Frequency effects
Speech level at listener position (due to theatre)
Speaker sound power
Background noise level

Arguments have also been put forward that speech intelligibility is not the only important consideration in rooms [1]. This paper will endeavour to show which of the influences listed deserve the major attention in theatre design. Most of the data discussed here was derived in an extensive objective survey of 12 British theatres and a complementary subjective survey of three of these theatres. It is appropriate however to start by discussing the sound source.

ACTORS' VOICE LEVELS

A measurement was made of a trained actor's and an actress's voice level. (The sample is regrettably very small but union regulations complicate what would otherwise be a simple operation.) The procedure was for the speaker to stand at a position on a theatre stage and to speak lines of his/her choice which were recorded via a microphone in the auditorium. The transmission between source and receiver position was measured afterwards with a calibrated loudspeaker source with directionality close to a human speaker (referred to below as the speech source). This method gave results of sound power level (SWL) directly. (Most measurements on the human voice are of sound pressure level (SPL) in front of the speaker; to convert from SWL to SPL or v.v. it is necessary to include the directivity factor for the human voice.) Sound level measurements were made by L_{eq} analysis in octaves. The following sound power levels were recorded:

1. Actor at loud conversational level	82 dB
2. Actor normal level to a full house	78 dB
3. Actress medium level to a full house	78 dB
Normal male speaking level [2]	75 dB
Normal female speaking level [2]	72 dB
Male shouting level [2]	102 dB

The final items of data are included for comparison. People normally increase voice level in noisy environments by on average 17 dB for 90 dB SPL of noise [3], which implies 92 dB voice power level in that situation. Reference [4] contains further interesting data for comparison after Knudsen for six lecturers at the University of California, Los Angeles (UCLA) who varied in voice power from 66 to 78 dB SWL ! Our figures therefore suggest that actors operate a few decibels above normal speakers, but way below levels achievable by untrained persons in stress situations.

A comment should also be made about the spectra recorded. Recording 2 (actor, normal level) has a spectrum virtually identical to that quoted for four (ordinary) Dutch speakers [5]. The difference between recording 1 and 2 is only at 500 Hz and above with an increase of 6 dB, but with no change in level at 125 and 250 Hz. Without wishing to suggest that our actor was shouting (he certainly was not), it is interesting to note a very similar characteristic which occurs between speaking and shouting [2]: the level at 250 Hz and below is almost constant and increases of 30 dB occur at 500 Hz and above. In the case of shouting the laryngeal frequency more than doubles its normal value.

NOISE LEVELS IN THEATRES

This consists of three major components: intruding environmental noise, ventilation noise and audience breathing noise. Nothing needs to be said about the former other than that it is not uncommon in theatres ! Background noise levels, principally due to ventilation, were measured in ten of the theatres, the results range from NC23 to NC35. Beranek [6] suggests as a criterion for 'large drama theaters' a noise level of not over NC20, which implies that several of these theatres rate as noisy. Breathing noise has been quantified by Kleiner [7] and is typically responsible for levels of 27 dB at mid-frequencies in large theatres, with a reasonably flat spectrum. While this may mask ventilation noise at mid-frequencies, a relaxation of the ventilation noise criterion of more than a few dB at these frequencies is undesirable if one wants to retain the possibility of a dramatic hush. Ventilation noise dominates at low frequencies over breathing noise.

ACCEPTABLE SIGNAL-TO-NOISE RATIOS FOR SPEECH INTELLIGIBILITY

In general speech intelligibility is determined by two factors: the source-receiver impulse response and the signal-to-noise ratio. Clearly in large open air theatres the latter is going to be the dominant concern, but in enclosed spaces both can be important. Most objective measures of speech intelligibility are just concerned with one factor or other, such as Articulation Index [8] which is a signal-to-noise measure. An inherent advantage of the Modulation Transfer Function (MTF) approach is that the two factors are automatically included [9]. Given the prediction method available, it is possible to investigate the effect of poor signal-to-noise on predicted intelligibility. To perform the calculation it is necessary to assume particular room impulse responses, or rather modulation functions. It is clear if one pursues this calculation that with signal-to-noise levels greater than 10 dB only minor degradation of speech intelligibility will occur. Is this something which is achieved in theatres ?

SIGNAL-TO-NOISE RATIOS IN THEATRES

To calculate this, as well as speaker sound power and the background noise level, we need the speech level at individual seats in theatres. This has been measured with the calibrated speech source already mentioned for two speaker orientations on stage. For background noise due to ventilation noise alone, only 5% of positions measured have mid-frequency signal-to-noise (S/N) ratios of less than 10 dB. This increases to 13% if breathing noise is included, but inevitably the same seats and theatres are involved. Introducing breathing noise into the calculation just exacerbates an already poor S/N situation. However all measured S/N ratios were better than 5 dB.

It is worth mentioning in passing the signal-to-noise situation at 125 Hz. The speech spectrum decreases in the bass, whereas most ventilation noise follows NC curves which rise sharply at low frequencies. This will produce a poor signal-to-noise ratio at 125 Hz, to the extent that masking in that particular octave is probably common in theatres.

Examination of locations with poor signal-to-noise shows that of the ten theatres measured only two theatres have S/N problems at more than one seating location. One of these problem theatres is the Roundhouse, Chalk Farm, London, which with theatre-in-the-round and a background noise level of NC34 was predictably marginal (the theatre is now closed). The other theatre, X, has a poor ventilation noise level of NC27 and low speech levels, which combine to provide poor S/N ratio.

We can conclude this section by saying that signal-to-noise ratio in theatres looks rarely likely to be a problem affecting speech intelligibility. This will particularly be true when ventilation noise levels are close to or better than NC20. It is interesting that theatre X has not been criticised for its acoustics since its opening, whereas other theatres with deficient impulse responses have been criticised. Perhaps the particular care which actors take at good articulation contributes here. Nevertheless this analysis suggests that a good voice level for actors is necessary; some of Knudsen's UCLA lecturers would not have made the grade !

SUBJECTIVE QUESTIONNAIRE STUDIES

The discussion now turns to the more general question of what theatre-goers are looking for in the acoustics of a theatre. Expert listeners familiar with questionniare studies attended performances at three theatres : the Olivier Theatre, National Theatre, the Shakespeare Theatre, Barbican (these are both in London) and the Festival Theatre, Chichester. A total of eleven seat positions were used with on average 5.8 responses per seat. The scales on the questionnaire were of the bipolar semantic differential type as follows:

Intelligibility, Overall
Intelligibility, Actors facing forward
Intelligibility, Actors facing away across stage
Intelligibility, Actors at rear of stage
Ease of listening
Reverberance
Intimacy
Loudness
Echo disturbance
Background noise
Overall impression

The choice of scales, and the obvious bias towards scales for intelligibility, reflects the fact that the results were intended for two forms of analysis. The aim has been to determine which characteristics of the sound are important in theatres, as well as to provide results suitable for correlations with objective measures of intelligibility. The questionnaire design may therefore compromise the results, particularly for the determination of important characteristics for theatres. These results should therefore be interpreted as preliminary. Correlation analysis between responses on the different scales shows that responses on all intelligibility scales are highly intercorrelated. Intelligibility is also highly correlated with 'Ease of listening' and 'Overall impression'. The standard technique in situations like this is factor analysis, which separates results into orthogonal dimensions. Three clearly independent dimensions emerge from factor analysis:

Intelligibility / Ease of listening

Intimacy

Reverberance

Of these only the intelligibility dimension correlates with 'Overall impression', and that with a high coefficient (r=0.85). Subjects judge acoustic intimacy and reverberance consistently but do not consider that they contribute to the overall judgement.

The remainder of the discussion will concentrate on the results of the intelligibility judgement. It should be stressed here that listeners were assessing intelligibility on a scale between 'Very poor' and 'Very good' and not measuring syllable intelligibility as such. Latham and Newman [1] have pointed out that even with better than 97% syllable intelligibility subjects respond consistently to a need for good 'speech quality'. However 'speech quality' was found to be well correlated with a modified objective measure normally associated with intelligibility [10], implying that 'speech quality' is an extension of syllable intelligibility. In our case these two concerns are likely to be accommodated in the single bipolar intelligibility scale.

OBJECTIVE CORRELATES OF SPEECH INTELLIGIBILITY

The following objective measures have been calculated from the measurements in theatres:

1. Early energy fraction (EEF)
2. Lochner and Burger's measure (LB)
3. Centre Time (CT)
4. Modulation transfer function (MTF)
5. Modulation depth (MD)

The first four measures were derived from measured impulse responses with the directional speech source; results have been calculated in the octaves 125 Hz to 4 kHz. The early energy fraction is the fraction of energy arriving within 50 ms of the direct sound. Lochner and Burger's measure [11] is similar in concept but avoids the sharp temporal cut-off of the EEF. Early sound is taken as before 95 ms, with a reduced weighting for later early reflections; their measure involves converting to decibels the ratio between this weighted early and the late energy after 95 ms. The Centre Time [4] is the first moment of area of the squared impulse response, with units of time; shorter values of Centre Time correspond to higher intelligibility. The Modulation Transfer Function has already been mentioned. For each octave filtered impulse response it is calculated by Fourier transformation at 18 modulation frequencies and results are averaged according to the specified procedure [9]. The Modulation Depth [12] is a crude version of the MTF, which is particularly easy to measure. A gated noise signal (frequency range 500-2000 Hz) is radiated via the speech source into the theatre, the modulation depth is measured by comparing the envelope and signal level. These five measures are all well intercorrelated with correlation coefficients of greater than r=0.80. None of these measures incorporated signal-to-noise terms for this part of the exercise.

Hawkes [13] has made perceptive observations about the suitability of individual objective measures. He suggests that for a certain measure to become established it has to fulfil certain criteria, principal among these are Predictability and Measurability. In Table 1 the ease of measurement and prediction for these measures are listed. The measurability refers to the ease of measurement on site, given normal equipment development facilities. For our study the measures 1-4 have been calculated on a specially programmed mini-computer; with this non-portable facility all measures are easily measurable ! The firm of Brüel and Kjaer [14] has recently introduced an instrument which measures the Modulation Transfer Function at nine points on the signal frequency/modulation frequency matrix compared with the 108 points on the whole matrix at six octave signal frequencies; it is referred to as the RASTI technique.

Table 1. Acceptability of different speech intelligibility measures

Measure	Measurability	Predictability
Early energy fraction	Easy	Difficult
Lochner and Burger measure	Fairly easy	Very difficult
Centre Time	Difficult	Impossible?
Modulation Transfer Function	Very difficult	Impossible?
" (RASTI)	Easy	Impossible?
Modulation Depth	Easy	Impossible?

Prediction of any of these measures can readily be made given the impulse response, such as might be calculated by a ray-tracing computer program. Calculation from drawings and acoustical absorption data is considerably more difficult, it is this which is listed in Table 1. Reverberation time and source-receiver distance are definitely inadequate parameters for theatre conditions. Prediction methods for the early energy fraction are discussed in reference [15]. Prediction of other measures in spaces like theatres has not been attempted and appears impossible for several of them.

For correlations between subjective and objective results, the subjective results for eleven seat positions were used. At eight of these positions, listeners had scored on intelligibility scales for actors both facing forwards and away across stage. Directly comparable objective data existed for these two source orientations, allowing 19 comparisons between subjective intelligibility results and objectively measured values. For correlation purposes subjective results were averaged across listeners at each seat position. The correlation coefficients between subjective and objective results are plotted in Figure 1. For the objective measures three means across signal frequency have been used. In the first column the mean for each measure is taken across the octaves 500 Hz to 2kHz. The modulation depth is derived from a single measurement with a two octave wide source signal. The RASTI technique involves

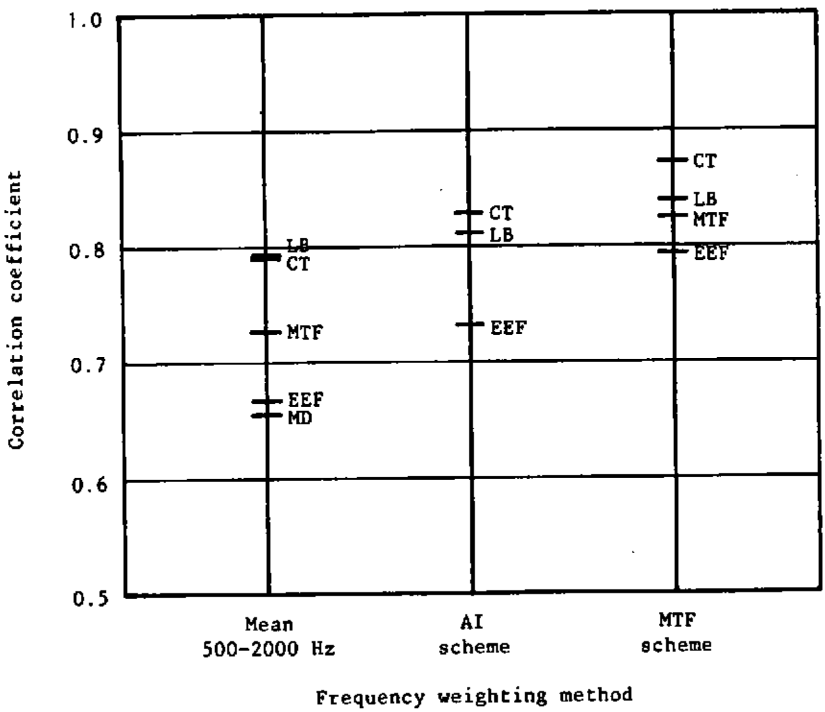

Figure 1. Correlation coefficients between subjective intelligibility and objective measures, as labelled, for three different weighting methods of objective signal frequency.

measurements in the 500 Hz and 2 kHz octaves only and is likely to be slightly less well correlated with intelligibility than the MTF in the first column (taken from 54 points on the frequency matrix). For the remaining columns in Figure 1, a weighting across the measurement octaves was used.

Two frequency weightings for speech intelligibility have been proposed. The first is associated with the Articulation Index (AI) scheme [8] and the second was developed by Houtgast, Steeneken and Plomp [9], therefore called here the MTF weighting. The weighting factors are listed in Table 2. Since no measurement results at 8 kHz were available, the 8 kHz octave weighting for the MTF weighting scheme has been redistributed over the adjacent octaves. The major difference between the two weightings is the greater weight given to bass frequencies in the MTF weighting.

Results in Figure 1 show improved correlations for each measure when five rather than three octaves are taken into consideration with the AI weighting scheme. Further improved correlations are found in all cases with the MTF weighting scheme. In this last situation Centre Time emerges as the best correlated with subjective results with the very respectable correlation

Table 2. Weighting factors applied to octave measurements

Octave frequency (Hz):	125	250	500	1k	2k	4k
AI weighting	0	.05	.15	.23	.32	.25
MTF weighting	.13	.14	.11	.15	.23	.24

coefficient of r=-0.87. Lochner and Burger's measure comes second and is marginally the best when only the mean of 500 to 2000 Hz is taken. However even the worst correlation coefficient of r=0.65 is valuable for practical purposes.

CONCLUSIONS

This analysis has perhaps not been as extensive on all points as one would wish, but the conclusions in each case are relatively unambiguous. Signal-to-noise ratio emerges as only likely to be a problem in theatres with high background noise levels, say greater than NC23. In theatres with inherently good intelligibility, particularly proscenium theatres, even higher noise levels are unlikely to degrade intelligibility.

Analysis of subjective questionnaire results shows listeners' indifference to the perceptible attributes of 'Intimacy' and 'Reverberance', which both appear important for music listening. Responses on scales of intelligibility match preference very closely. The best objective correlates of subjective intelligibility are Centre Time and Lochner and Burger's measure. The correlation is better for all objective measures when a wide frequency range is used for objective measurements. Also consistent is the superiority of the frequency weighting scheme proposed for MTF measurements. This result implies that attention to behaviour at 125 and 250 Hz is important. The simplest design conclusion from this is that bass frequency rise in reverberation time is to be avoided for speech use.

If objective intelligibility measures have paramount importance for theatre design, their prediction from drawings and basic acoustic absorption data would offer a valuable design aid. A reliable prediction technique is not easily produced; prediction based on models of direct sound and an exponential decay is quite inadequate for theatres. While efforts at predicting the Early energy fraction [15] have been reasonably successful, the possibility of similar methods being developed for other measures appears remote.

References

1. H.G.LATHAM and P.I.NEWMAN 1980 Proceedings of the Institute of Acoustics, Spring meeting, Bedford College, London, p. 21-25. Scaling speech quality in theatre acoustics.
2. D.ROSTOLLAND 1982 Acustica **50**, 118-125. Acoustic features of shouted voice.
3. T.S.KORN 1954 J. Acoust. Soc. Am. **26**, 793-794. Effect of psychological feedback on converstional noise reduction in rooms.
4. L.CREMER and H.A.MÜLLER (translated by T.J.SCHULTZ) 1982 'Principles and applications of room acoustics, Vol.1'. Applied Science Publishers, London.
5. H.J.M.STEENEKEN and T.HOUTGAST 1978 Institute for Perception TNO. Soesterberg, The Netherlands. Report no. IZR 1978-22. Comparison of some methods for measuring speech levels.

SPEECH IN THEATRES : WHAT ARE THE IMPORTANT CONSIDERATIONS ?

References (contd)

6. L.L.BERANEK 1971 'Noise and vibration control'. McGraw-Hill, New York.
7. M.KLEINER 1980 Acustica **46**, 82-88. On the audience induced background noise level in auditoria.
8. K.D.KRYTER 1962 J. Acoust. Soc. Am. **34**, 1689. Methods for the calculation and use of the Articulation Index.
9. T.HOUTGAST, H.J.M.STEENEKEN and R.PLOMP 1980 Acustica **46**, 60-72. Predicting speech intelligibility in rooms from the Modulation Transfer Function,I: General room acoustics.
10. H.G.LATHAM and P.I.NEWMAN 1982 Proceedings of the Institute of Acoustics, Auditorium acoustics and Electro-acoustics meeting, Edinburgh. Subjective preference design criteria for evaluation acoustic quality.
11. J.P.A.LOCHNER and J.F.BURGER 1961 Acustica 11, 195-200. The intelligibility of speech under reverberant conditions.
12. M.KLEINER 1981 Acustica **47**, 55-71. Speech intelligibility in real and simulated sound fields.
13. R.J.HAWKES 1985 Fifth FASE symposium, Thessalonika, Greece. Criteria for noise criteria.
14. Brüel and Kjaer Technical Review 1985 No.3.
15. L.J.LEE and M.BARRON 1985 submitted to J. Acoust. Soc. Am. Objective measures of speech intelligibility in theatres.