PRACTICAL ASSESSMENT OF SPEECH INTELLIBILITY IN AUDIO SYSTEMS

P.W. Barnett

AMS Acoustics, 52 Chase Side, Southgate, London, N14 5PA.

INTRODUCTION

The reliable prediction of speech intelligibility remains an uncertain science. Many problems still remain, not least the imperfect correlation between objective and subjective assessments. For we should remember that this assumed correlation is man-made. A major problem however remains in the prediction of either $\%AL_{cons}$ or STI. They are both open to varied interpretation and so with the uncertainty of prediction and the uncertainty of correlation between the predicted values and reality, we should ask ourselves 'is the expenditure of effort really worthwhile?' My perception is that it is, since without the mandatory initial and tentative steps, we cannot expect to achieve a good understanding of the subject. We must however be cognisant of the limitations and shortcomings of our models and methods employed lest we should fall foul of misunderstanding the significance of our predictions.

SCOPE

The scope of this Paper is limited to a simple mathematical analysis of the predictive model in current common use.

The mathematical model attributed to Peutz is as follows:

$$\%AL_{cons} = 100\,\{10^{-2[(A+BC)-ABC]} + 0.015\}$$

where:

$$A = -0.32 Log_{10}\left[\frac{(L_R+L_N)}{(10L_D+L_R+L_N)}\right]$$

$$B = -0.32 Log_{10}\left[\frac{L_N}{(10L_R+L_N)}\right]$$

$$C = 0.5 Log_{10}\left[\frac{RT}{12}\right]$$

where: L_R = Absolute reverberant sound pressure level
L_D = Absolute direct sound pressure level
and L_N = Absolute ambient noise level.

$$\textit{Note}:\ L_R = 10^{\frac{Spl_R}{10}}$$

where: Spl_R = Reverberant sound pressure level (dB re $2\times10^{-5}Nm^{-2}$).
Similarly for L_D and L_N.

This replaces the more familiar:

$$\%AL_{cons} = \frac{200RT^2D^2(N+1)}{VQm}$$

where:
RT = Reverberation time (secs.)
D = Distance from source (m)
N = Number of like sources
V = Volume of space (m^3)
Q = Directivity of source
m = Modifier to account for primary absorption.

The latest model includes ambient noise as a variable which the original did not . The inclusion of ambient noise is becoming increasingly important in venues with high ambient noise levels (e.g. sports stadia and grandstands) since it is often impracticable to maintain high signal-to-noise ratios.

LIMITATIONS

There are two basic simple limitations:

Expression $A \leq 1$ and

Expression $B \leq 1$.

For A≤1 then the following condition must be met:

$$Spl_R + Spl_N \leq SplD - 21\ dB$$

where:
Spl_D = Direct sound pressure level
Spl_R = Reverberant sound pressure level
Spl_N = Noise sound pressure level.

That is to say that the logarithmic addition of the reverberant sound pressure level and the ambient noise sound pressure level shall be not less than -21 dB reference to the direct sound pressure level.

For B≤1 then the following condition must be met:

$$Spl_N \leq Spl_R - 21\ dB$$

which is self-explanatory.

It is important therefore when using the model that if $Spl_N \leq SplR - 21$ dB then a value of $Spl_N = Spl_R - 21$ dB eg. B = 1 must be inserted in the prediction model which then becomes:

$$\%AL_{cons} = 100\ \{10^{-2[(A+C)-AC]} + 0.015\}$$

where:

$$A = -0.32 Log_{10}\left[\frac{(L_R)}{(10L_D+L_R)}\right]$$

$$C = 0.5 Log_{10}\left[\frac{RT}{12}\right]$$

where:
L_R = Absolute reverberant sound pressure level
L_D = Absolute direct sound pressure level
Note: Must still be A≤1.

It is worth noting that when $Spl_D >> Spl_R$

Hence: A tends to unity and $Spl_R >> Spl_N$ and B tends to unity and RT is small hence C is approximately unity.

Then $\%AL_{cons}$ 2.5%

Hence the minimum value of $\%AL_{cons}$ that can be predicted using this model within its limitations is 2.5% which if compared with the table given in Appendix A implies that using this model a subjective rating of excellent cannot be achieved.

PRACTICE

In practice audio system designers are able to control the difference between Spl_D and Spl_N and indeed the space imposes the relationship between Spl_D and Spl_R.

Hence the problem is better approached by introducing direct-to-reverberant ratio and system signal-to-noise ratio.

We may then substitute Spl_R as Spl_D -DR
where DR = Direct-to-reverberant ratio in dB

and Spl_N - Spl_D -DN
where DN = Signal-to-noise ratio in dB.

Hence: $L_R \frac{L_D}{L_{DR}}$ and $L_N = \frac{L_D}{L_{DN}}$

Therefore: $$A = -0.32\text{Log}_{10}\left[\frac{L_{DR} + L_{DN}}{L_{DR} + 10L_{DR}L_{DN} + L_{DN}}\right]$$

and B becomes: $$B = -0.32\text{Log}_{10}\left[\frac{L_{DR}}{10L_{DN} + L_{DR}}\right]$$

C remains unchanged as $C = 0.5\text{Log}_{10}\left[\frac{RT}{12}\right]$

By using the revised coefficients A and B together with C in the following expression

$$\%AL_{cons} = 100\,\{10^{-2[(A+C)-AC]} + 0.015\}$$

a series of graphs may be produced.

Figs. 1 and 2 are two examples showing the relationship between AL_{cons} and direct-to-reverberant ratio for varying RT but with fixed signal-to-noise ratio.

SIGNAL-TO-NOISE RATIO = 0dB

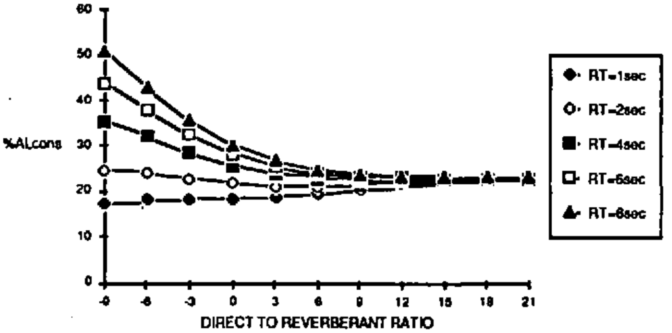

FIG 1

SIGNAL-TO-NOISE RATIO = 3dB

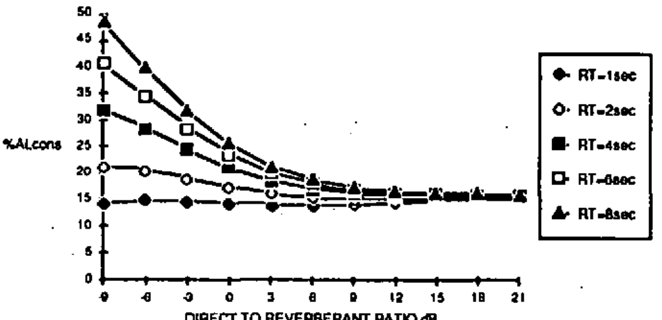

FIG 2

It can be seen that at low reverberation times and high direct-to-reverberant ratios, the achievable intelligibility reaches a minimum, fig. 3 shows these minimum values. It can be seen that for a $\%AL_{cons}$ of 10% (considered fair-to-good) the minimum signal-to-noise ratio required is 6 dB. From the constant signal-to-noise curves it can be seen that reducing the reverberation time further or improving the direct-to-reverberant ratio beyond 10 dB does little to improve the situation.

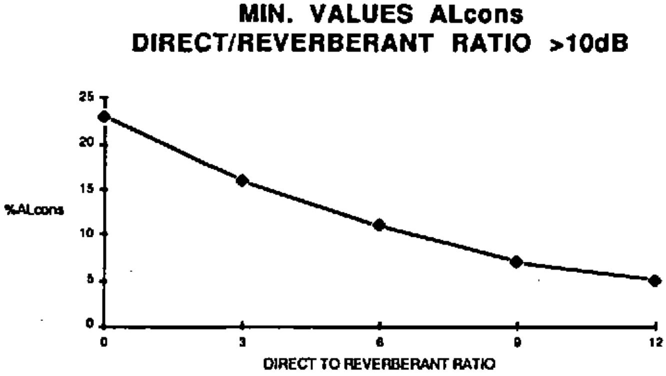

FIG 3

Figs. 4 and 5 show the predicted intelligibility for constant reverberation time. Fig. 4 demonstrates that for low RT and poor signal-to-noise ratio, the model gives unreasonable results and care should be exercised in its use in these regions.

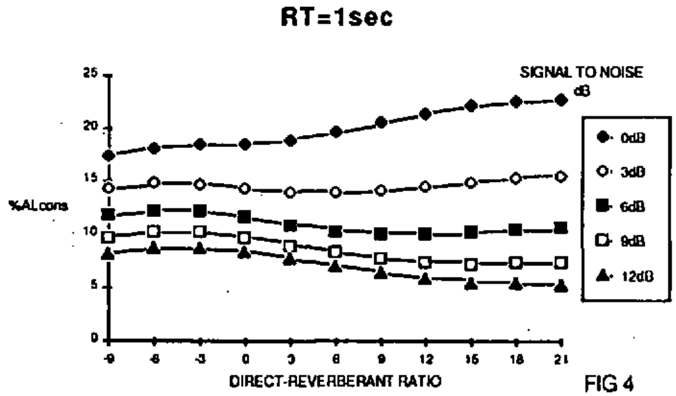

FIG 4

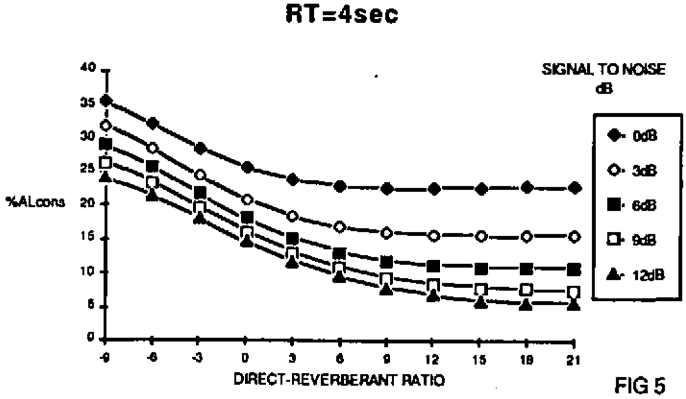

FIG 5

As one might expect at low values of RT the predicted values of intelligibility are insensitive to changes in either direct-to-reverberant ratio or signal-to-noise ratio. At larger RT's however predicted intelligibility become extremely sensitive to changes in direct-to-reverberant ratio especially in the region of 0 dB. The table below gives a rough indication of the rate of change of AL_{cons} with direct-to-reverberant ratio in the range -6 dB to +6 dB and for +12 dB signal-to-noise ratio.

RT (secs.)	$\%AL_{cons}$/Direct-to-reverberant ratio $\frac{\delta AL_{cons}}{\delta dB}$
1	0.13
2	0.4
4	1
6	1.5
8	2

It can be seen that for reverberation times in excess of 2 secs. the predicted intelligibility is sensitive to changes in direct-to-reverberant ratio and hence it is extremely important to correctly predict the relationship between the direct and reverberant sound fields.

PRACTICAL ASSESSMENT OF SPEECH INTELLIGIBILITY

Apart from the normal considerations, it is important to additionally consider the overall dispersion characteristics of the transducers used and to take proper account of the initial reflection from the primary surface.

CONCLUSIONS

I am concerned on two counts, firstly that this model does not give good agreement with reality and secondly that the model is open to misinterpretation.

On the first count I would refer to fig. 1. It would be expected that with an RT of 1 sec. a direct-to-reverberant ratio of +12 dB (or greater) and a signal-to-noise ratio of +9 dB the resulting subjective assessment would be in the range good where as for a $\%AL_{cons}$ of 7, the subjective rating is only fair.

The misinterpretation is, however I believe, a larger problem. A spot calculation for a reverberation time of 1 sec., direct-to-reverberant ratio of 12 dB and a signal-to-noise ratio of 0 dB would result in a calculated value of $\%AL_{cons}$ of 21%. A glance at fig. 1 would suggest that the calculated value was suspect since it is part of a series of values which are in opposition to what is sensible.

Finally, although figures should be quoted for reference sake, we must be aware and preferably make others aware of the accuracy of predicted intelligibility. One of the greatest uncertainties will be the direct-to-reverberant ratio. Given the error in predicting the reverberation time and the difficulty in determining the reverberant field, an error in direct-to-reverberant ratio in the region 3 - 6 dB would not seem unreasonable.

This would result in an error in prediction value in the region 3% - 6% for a 4 sec. RT and it must be remembered that this error is in addition to those inherent in the model.

Hence I would advocate the use of the subjective descriptive terms - Good, Fair and Poor - and more truthfully these should be applied not to an assessment of intelligibility rather they should reflect the probability of the success of the project.

References: Sound System Engineering 2nd Edition - Don Davis & Carolyn Davis
STI Measurements on Simulated Acoustic Environments - P.W.Barnett & P.H. Scarbrough - Proceedings Institute of Acoustics Vol.11 Pt. 7 (1989).

Appendix A

Subjective Assessment	RASTI	%ALcons	Subjective Assessment	RASTI	%ALcons
	0.20	57.7		0.60	6.6
	0.22	51.8		0.62	5.9
BAD	0.24	46.5		0.64	5.3
	0.26	41.7	GOOD	0.66	4.8
	0.28	37.4		0.68	4.3
	0.30	33.6		0.70	3.8
	0.32	30.1		0.72	3.4
	0.34	27.0		0.74	3.1
POOR	0.36	24.2		0.76	2.8
	0.38	21.8		0.78	2.5
	0.40	19.5		0.80	2.2
	0.42	17.5		0.82	2.0
	0.44	15.7		0.84	1.8
	0.46	14.1		0.86	1.6
	0.48	12.7	EXCELLENT	0.88	1.4
	0.50	11.4		0.90	1.3
FAIR	0.52	10.2		0.92	1.2
	0.54	9.1		0.94	1.0
	0.56	8.2		0.96	0.9
	0.58	7.4		0.98	0.8
				1.00	0.0

$\%AL_{cons} = 170.5405 \exp. (-5.419 * STI)$ & $STI = 0.1845 \ln \%AL_{cons} + 0.9482$.