Proceedings of the Institute of Acoustics

CONVERSION OF RASTI TO ALCONS; A ROBUST AND REVERSIBLE PROCESS?

Peter Barnett

AMS Acoustics, 43 Chase Side, Southgate, London, N14 5BP.

INTRODUCTION

The interest in conversion and translation between these two speech intelligibility indicators results from a number of factors but mainly due to the fact that RASTI is difficult to predict directly but relatively straight forward to measure and AI_{cons} is relatively straight forward to predict but difficult to measure.

Hence predict AI_{cons}; translate to RASTI which may then suffer a measurement audit.

The idea seems attractive but is, of course, sensitive to the relationship between the two indicators. We should also investigate if this or any relationship is robust and finally is it reversible?

The figure below provides an indication that there may be a relationship.

Family Tree of Intelligibility Indicators

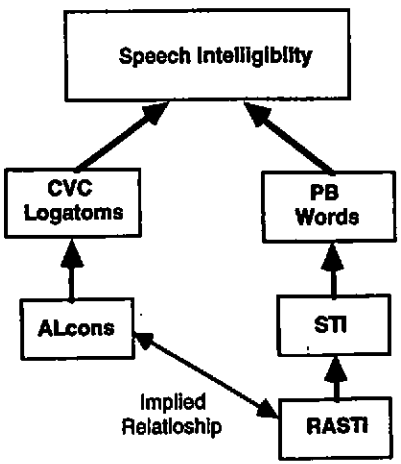

Fig. 1

CONVERSION OF RASTI TO ALCONS

If we define the final parameter under test as speech intelligibility (Si) then

$Si = f_1$ (RASTI) - via PB Word Scores
$Si = f_2$ (AL_{cons}) - via CVC Logotoms

then for identical conditions we may in a very simplistic way write

$$f_1 (RASTI) = f_2 (AL_{cons})$$

Now of course f_1 and f_2 may be extremely complex and non-linear with other dependent and independent variables.

EXISTING RELATIONSHIPS

There are two existing proposed relationships to convert RASTI to AL_{cons} and vice versa - they are:

1. Attributable to Farrel Becker

$$\%AL_{cons} = 170.5405e^{-5.419STI}$$

2. Attributable to Peutz

$$\%AL_{cons} = 10^{\frac{STI-0.09}{-0.425}}$$

The figure overleaf gives a graphical representation of the relationship.

CONVERSION OF RASTI TO ALCONS

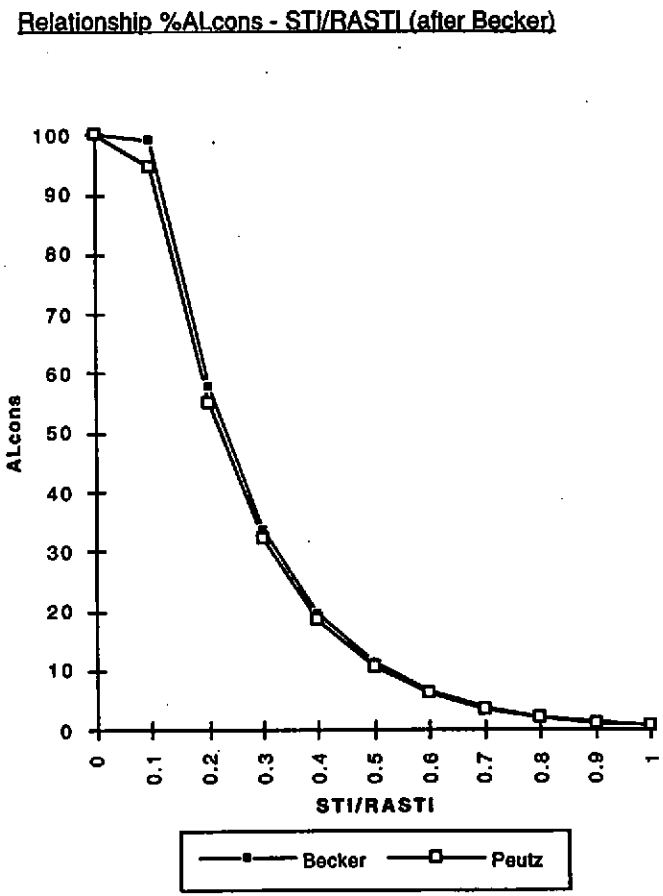

Fig. 2

I am not aware that either have been published and therefore unable to provide any background information.

CONVERSION OF RASTI TO ALCONS

WHICH $\%AL_{CONS}$?

$\%AL_{cons}$ has appeared in a wide variety of formants vis:

Format 1 - Simple Expression - $\%AL_{cons} = \dfrac{200r^2T^2}{VQ}$

where: r = Distance from source (m)
T = Reverberation Time (sec.)
V = Volume of space (m^3)
Q = Directivity of source.

and

Format 2 - Complex Expression - $\%AL_{cons} = 100\left(10^{-2(A+BC-ABC)} + 0.015\right)$

where: $A = -0.32Lg\left[\dfrac{L_R + L_N}{10L_D + L_R + L_N}\right]$

$B = -0.32Lg\left[\dfrac{L_N}{10L_R + L_N}\right]$

$C = -0.5Lg\left(\dfrac{T}{12}\right)$

where: $L_D = 10^{\frac{Lpd}{10}}$, $L_R = 10^{\frac{Lpr}{10}}$, $L_N = 10^{\frac{LpN}{10}}$
and T = Reverberation Time (secs.)

where: L_{pd} = direct sound pressure level dB re 20μPa
L_{pr} = reverberant sound pressure level dB re 20μPa
L_{pN} = Noise sound pressure level dB re 20μPa.

Both equations may be rewritten in terms of direct-to-reverberant ratio and reverberation time vis:

Format 1 - $\%AL_{cons} = \dfrac{2}{\pi}.T.10^{-\left(\frac{D/R}{10}\right)}$

CONVERSION OF RASTI TO ALCONS

Format 2 - $\%AL_{cons} = 100\left(10^{-2(A+BC-ABC)} + 0 - .015\right)$

where: $A = -0.32Lg\left[\frac{L_{DR} + L_{DN}}{L_{DR} + 10L_{DR}L_{DN} + L_{DN}}\right]$

$B = -0.32Lg\left[\frac{L_{DR}}{10L_{DN} + L_{DR}}\right]$

$C = -0.5Lg\left(\frac{T}{12}\right)$

where: T = Reverberation Time (secs).

$L_{DR} = 10^{\left(\frac{D/R}{10}\right)}$

$L_{DN} = 10^{\left(\frac{S/N}{10}\right)}$

where: D/R = Direct-to-Reverberant Ratio (dB)
S/N = Signal-to-Noise Ratio (dB).

It is therefore possible to plot each relationship against D/R (S/N is assumed >25dB).

These are shown in the figures below:

Graph of %ALcons (Format 1) vis D/R for Various Reverberation Times

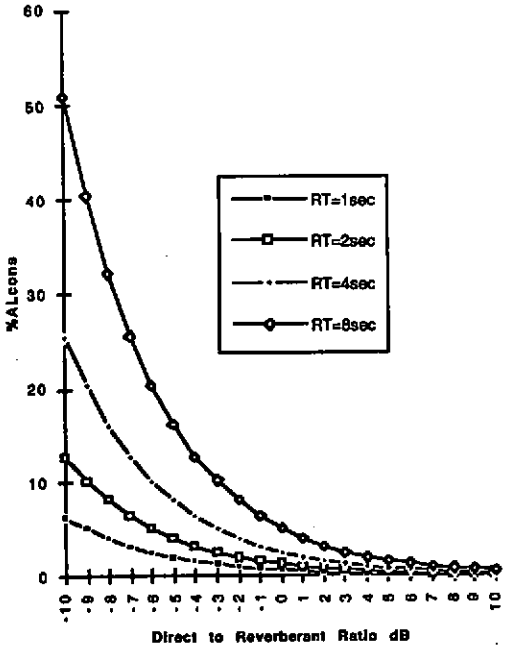

Fig. 3

Graph of %ALcons (Format 2) vis D/R
for Variation Reverberation Times

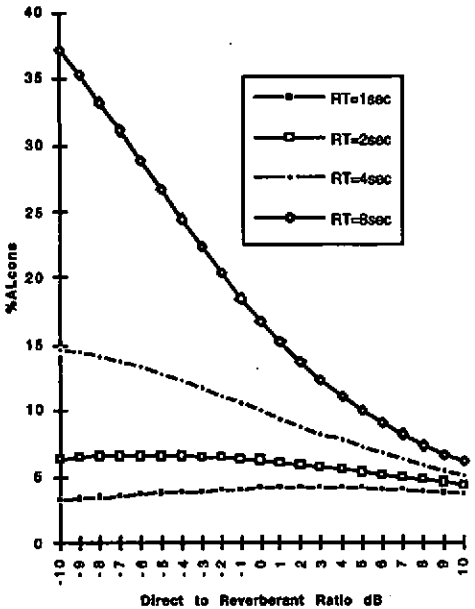

Fig. 4

It can be seen that the two formats are quite different, for ease of comparison they are plotted together on the following figure (1 sec. and 4 secs. only).

Graph of %ALcons (both Formats) vis D/R
for RT = 1 sec. and 4 sec.

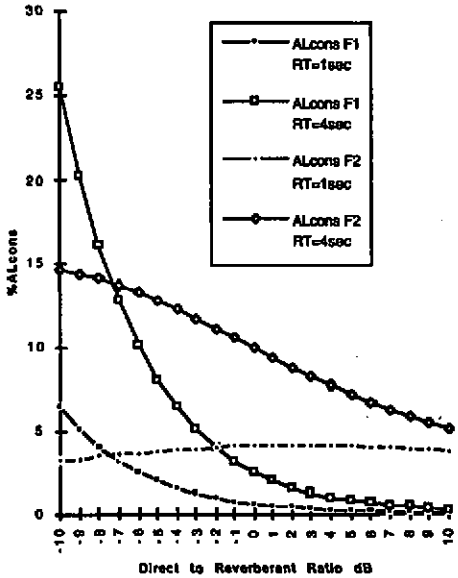

Fig. 5

Clearly there is a major difference which I am unable to explain but it does beg the question which algorithm is applicable.

STI/RASTI

In common with $\%AL_{cons}$ STI/RASTI may be presented graphically against D/R. Computations were performed on the assumption that the acoustic field was a homogenous continuum.

Graph of STI/RASTI vis D/R for Various Reverberation Times

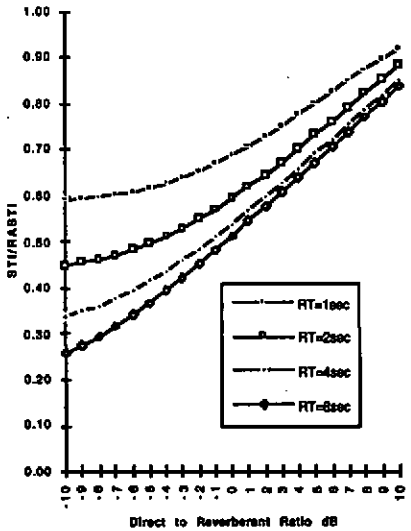

Fig. 6

The graph is as expected being asymptotic to $m(F) = \frac{1}{\sqrt{1+\left(\frac{2\pi FT}{13.8}\right)^2}}$ and produces an improvement with increasing D/R ratio.

COMPARISON OF AL_{CONS} AND RASTI

Since both indicators are now referenced to direct-to-reverberant ratio, a direct comparison may be made. For convenience only RT = 2 sec. and RT = 4 sec. are displayed and the $\%AL_{cons}$ axis has been reversed such that both indicators show the same trend.

CONVERSION OF RASTI TO ALCONS

The figure below demonstrates a comparison using the simple Format 1 relationship.

Graph of STI/RASTI and Format 1 %ALcons (axis reversed) vis D/R for RT 2 secs. and 4 secs.

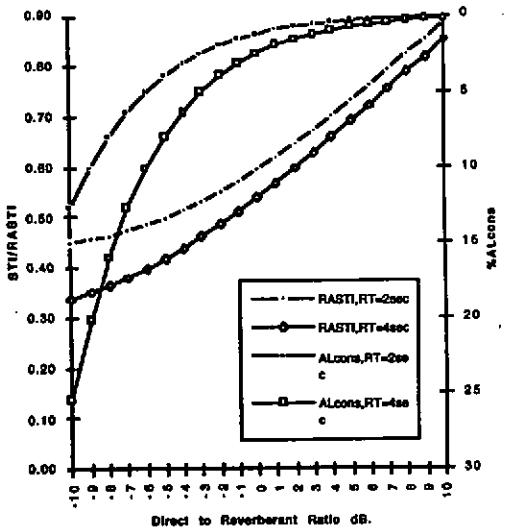

Fig. 7

The figure below displays a comparison using the complex Format 2 relationship.

Graph of STI/RASTI and %ALcons (axis reversed) vis D/R for RT 2 secs. and 4 secs.

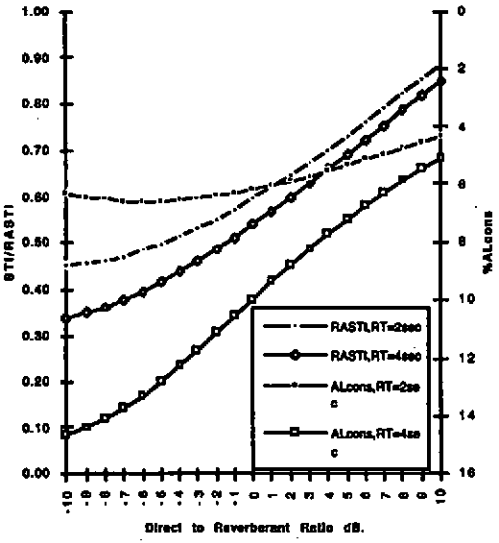

Fig. 8

The potential for a definable relationship may be deduced from the two foregoing figures. Fig. 7 has been reproduced overleaf without the legend but with major grid lines present.

CONVERSION OF RASTI TO ALCONS

Fig. 7 Reproduced with Grid Lines

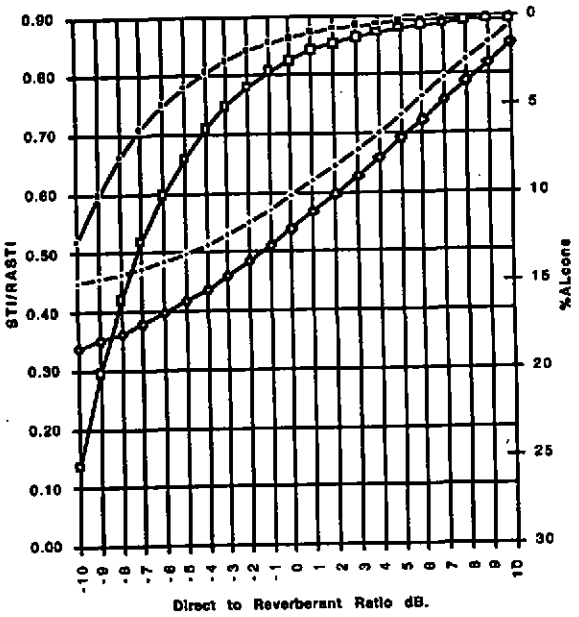

Fig. 9

Now ordinates are lines of iso-position in a space and abscissa are lines of iso-intelligibility. Hence we may deduce that for RASTI = 0.5 with RT = 2 sec. the direct-to-reverberant ratio shall be approximately -5dB.

Therefore in the same position in space with the same RT the $\%AL_{cons}$ may be determined as approximately 5%. Now we may hypothesise that if we were able to immediately change the RT in the space from 2 secs. to 4 secs. then by moving closer to the source, a new position may be found which also gives 0.5 RASTI. This condition may be determined traversing the 0.5 RASTI abscissa to 4 secs. RT RASTI curve we find that the new position to return a RASTI of 0.5 with 4 secs. RT is at D/R = approximately -1dB. Now at the new position with the increased RT the $\%AL_{cons}$ should also be as before in the region 5%, examination of fig. 7 by tracing the -1dB D/R ordinate to the 4 sec. $AL_{con\ s}$ curve gives a returned value of around 3%. A small difference may be observed.

If however, the same exercise is carried out at RASTI = 0.45 then the returned $\%AL_{cons}$ of 2 sec. and 4 sec. RT are 13% and 5% respectively. A large difference.

Hence we may deduce that any conversion relationship between RASTI and AL_{cons} must be a factor of reverberation time which neither the Becker or Peutz algorithms are.

The same process may be repeated for the Format 2 complex relationship and this results in even greater differences. See the figure below:

Fig. 8 with Major Grid Lines

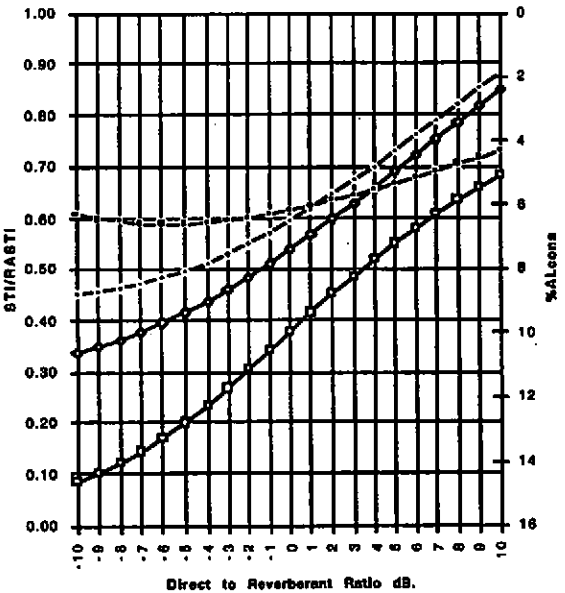

Fig. 10

CONVERSION OF RASTI TO ALCONS

The table below summarises the results for both sets of data at various RASTI values:

RASTI	Format 1 AL_{cons}		Format 2 AL_{cons}	
	RT = 2 secs.	RT = 4 secs.	RT = 2 secs.	RT = 4 secs.
0.45	13%	3%	6%	12%
0.55	2%	2%	6%	10%
0.65	≈1%	≈1%	5%	8%
0.75	≈1%	≈1%	5%	7%

Hence it can be seen that the conversion of both formats appears to be RT dependent but with differing sensitivities.

THE EFFECT OF NOISE

RASTI of course allows for the addition of noise using the m(F) multiplicand $\frac{1}{1+10^{\frac{-S/N}{10}}}$ where S/N = signal-to-noise ratio dB. Hence if at a given position for a given RT a RASTI value is returned which may be translated into a $\%AL_{cons}$ values, then the addition of noise at that position such as to reduce the RASTI, should also translate to a corresponding $\%AL_{cons}$.

To test the hypothesis that noise affects both $\%AL_{cons}$ and RASTI in the same manner, we may look to the effect of noise on CVC logatoms and that of PB Word Scores.

Consideration of the relationship between CVC logatoms and PB Word Scores shown in figure 11.

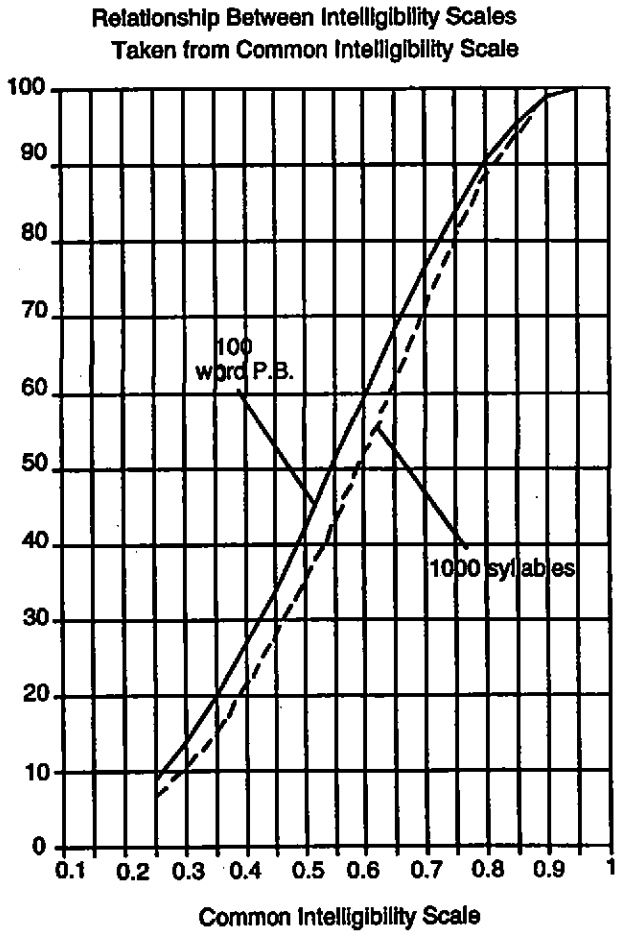

Fig. 11

suggests that there would be very little difference caused by the introduction of noise on these two subjective indicators although intuition suggests that since words may be deduced from speech that the introduction of noise may not produce a proportional response. The inverse may be true of CVC logotoms.

Notwithstanding the foregoing the complex Format 2 relationship does include for noise.

The figure below shows the variation of RASTI with S/N ratio for noiseless RASTI values 0.45, 0.55, 0.65 and 0.75 for RT's 2 secs. and 4 secs.

Graph of Effect of Noise on Identical RASTI
Data Sets a through d but with RT = 2 secs. and 4 secs.

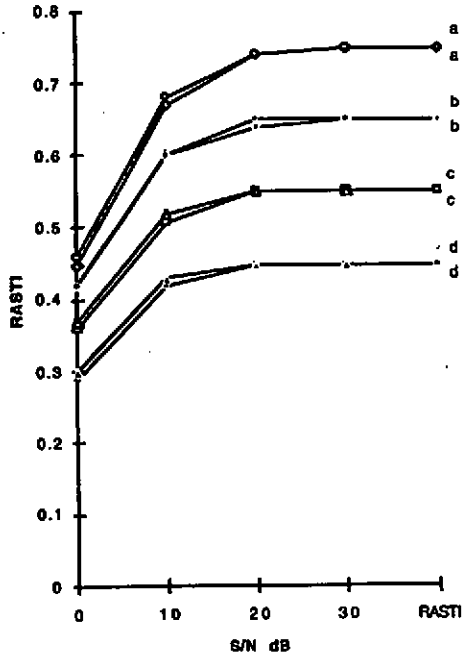

Fig. 12

Data pairs are annotated with the same letter.

It can be seen that there is very little variation between values for 2 secs. and 4 secs. and hence the effect of noise is roughly independent of RT (though not D/R).

CONVERSION OF RASTI TO ALCONS

The figure below shows the variation of AL_{cons} with S/N ratio for RASTI values 0.45, 0.55, 0.65 and 0.75 for RT's 2 secs. and 4 secs.

Effect of Noise on Identical %ALcons Data Sets a through d but with RT = 2 secs. and 4 secs.

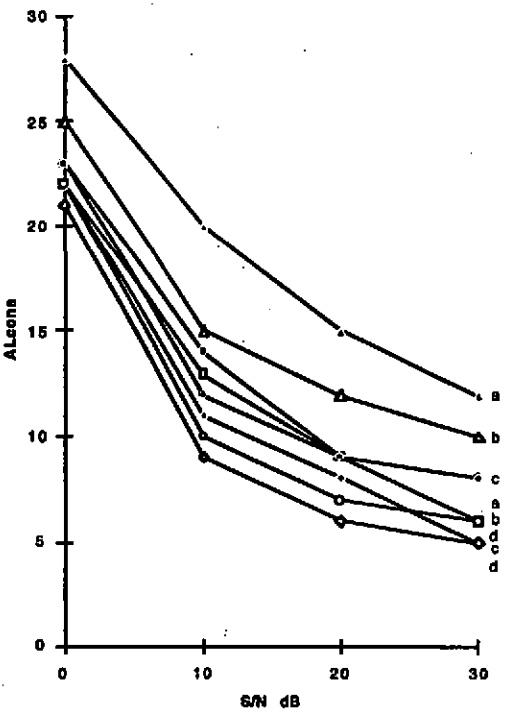

Fig. 13

Data pairs are annotated with the same letter.

It can be seen that there is a considerable difference between equivalent sets of data with the different RT and hence the effect of noise is a function of RT as well as S/N and D/R.

REVERSIBLE PROCESS

So far we have examined the conversion of RASTI to $\%AL_{cons}$, the expressions for conversion suggest that the process is reversible.

Embodied within the RASTI concept is the fact that to be a valid measurement the frequency response of the system shall be uniform over the frequency range 200Hz - 6kHz and hence we may assume if a RASTI value is quoted that the response of the system meets this simple criteria.

Since $\%AL_{cons}$ is determined at 2kHz and brings with it no qualification regarding frequency response, reversible translation cannot take place unless the same response criteria is applied as for RASTI.

Finally, since STI carries with it no qualification in regard of response or spectral content, translation between this scale is almost impossible.

DISCUSSION AND CONCLUSIONS

I believe the foregoing has demonstrated that there must be considerable doubt in a single conversion process between RASTI and AL_{cons} and vice versa. If a complete translation is possible then it should include reverberation time.

There also remains the problem of the dual formats for $\%AL_{cons}$ and the lack of published information and data on this subject. Consideration of the original work of Peutz and the present expression suggests that considerable metamorphosis has taken place. Equipment now apparently exists to measure $\%AL_{cons}$ yet in truth the concept is subjectively biased and hence the only real test is a subjective CVC audit.

STI/RASTI on the other hand is based on an objective measurement procedure.

Another perplexing fact has emerged, consideration of fig. 11 suggests that there is little difference between the trends and sensitivities of CVC logotoms and PB (1000) Words for varying speech intelligibility. Now both $\%AL_{cons}$ and RASTI purport to describe these trends in terms of perceived speech intelligibility. Yet when they are displayed together against a common variable (D/R) they demonstrate quite different trends (see figs. 7 & 8).

From the foregoing I have reached the following tentative conclusions:

1. Translation from RASTI to AL_{cons} should be RT dependent.

2. Translation becomes uncertain if noise is present.,

3. The process is not reversible i.e. RASTI should not be deduced from AL_{cons} unless qualifiers are in place.

This simple investigation also raises a number of concerns.

A. How could this situation have existed for so long especially given the potential consequences?

B. Why are relationships and concepts so readily accepted without robust evidence?

C. How could it be that two expressions both purporting to calculate the same indicator give different results?