Proceedings of the Institute of Acoustics

INTERPRETATION OF STI MEASUREMENTS IN PRACTISE

Paul Doany

Dar Al-Handasah Consultants (Shair & Partners), London

1. INTRODUCTION

The speech transmission index (STI) is one of the well-recognised methods of speech intelligibility analysis, being implemented in various equipment by different manufacturers. This paper gives some examples of results obtained using Bruel & Kjaer/RASTI, and relates these to the theoretical expectation on the basis of simple derived expressions presented in an earlier paper by the Author, wherein the direct-to-reverberant ratio D/R is considered in the analysis, which is a slight departure from the original EDT approach proposed by Steeneken & Houtgast. Validation of the suggested formulae is demonstrated with practical measurements where the usefulness of their application in the post-processing of RASTI data is shown along with proposed modification of the modulation reduction curves normally used with RASTI. Comparative results obtained using Techron/TEF are also briefly given.

2. OVERVIEW OF THE STI METHOD

The STI method is based on analysis of the modulation transfer function (MTF), defining an index in the range 0 to 1, for no intelligibility to full intelligibility. This index is statistically related to the empirically derived Articulation Loss of Consonants ($\%AL_{cons}$), with a corresponding range of 100% to 0%.

While the "subjective" limitations of all such type of analyses is inherent in the nature of the object parameter of "speech intelligibility" itself, the STI method may be loosely considered to be more appealing, being related to a "physical" parameter, namely the MTF, so that statistical confidence in the predictive analysis of this index would appear to be less flawed than a method totally reliant on empirical techniques.

Whether or not this is only a "psychological" difference between these two methods, the Author has in fact illustrated recently that if indeed both methods are expressed using the same traditional parameters of reverberation time, T, direct-to-reverberant field ratio, D/R, and signal-to-noise ratio,

S/N, remarkable similarity is noted between the methods [1]. This is illustrated for a set of expressions developed by the Author for STI in comparison with the algorithm suggested by Peutz and implemented in the TEF analyser [2]. This has also been observed in several experiments reported using the TEF Analyser and B&K RASTI [2-4].

3. QUANTITATIVE EFFECT OF THE DIRECT FIELD

The expression for the MTF derived by the Author [1], from which the STI can be easily obtained in any band of interest is given below:

$$m(F) = \frac{\sqrt{\{1 + \beta/(1+\alpha^2)\}^2 + \{\alpha\beta/(1+\alpha^2)\}^2}}{(1 + \beta)}, \qquad (1)$$

where, $\alpha = 2\pi FT/13.8$,

$\beta = 10^{0.1(D/R)}$,

F being the modulating frequency, and T being the reverberation time. The above expression is based on a simplified model, assuming a specular direct field and exponential decay [1]. This paper gives examples of results obtained by the Author in a series of measurements carried out in 1988-89, where its use has been attempted in practise and its possible usefulness illustrated, mainly for RASTI, although its application is not limited thereto.

4. POST-PROCESSING ANALYSIS FOR THE RASTI

RASTI uses two amplitude modulated noise carriers in the bands 500Hz and 2000Hz, with modulating frequencies of 1.02, 2.03, 4.07, 8.14Hz, and 0.73, 1.45, 2.90, 5.81, 11.63Hz, respectively.

B&K documentation for the use of RASTI is very clear, and gives the user comprehensive instructions in addition to background information on the quantitative analysis implemented in the device itself [5]. To illustrate the procedure followed by the Author, examples of various measurements are included for two typical positions in the "far field", and in the "near field".

4.1 Operation in the Far-field

Figure 1 gives the results for a ceiling distributed system with operation, at distances of over 3.5-4 times the critical distance. Several measurements were carried out, and Figure 1 gives a typical example which illustrates the additional post-processing carried out on a specially developed PC-programme, which provides the following results:

(a) The records are printed in a table form as shown (Figure 1), including computations of the index against each record, in addition to the overall average modulation reduction values and the respective RASTI, with one table for each of the 500Hz and 2000Hz, combining the overall RASTI in the former, for convenience.

(b) The table also gives calculated values of EDT, on the basis of a limited observation of the modulation reduction values at the lower frequencies (as indeed suggested by B&K [6], but not actually implemented in the RASTI device itself). These may be compared with the Eq-EDT given by RASTI, on the basis of the equivalent value of EDT that yields the obtained index in the absence of noise; the measurements were indeed carried out in noiseless conditions (note LEQ values), so it is expected that these two EDT values should not be very different. The value of FIT is hence included, which gives the rms error between the observed values and the pure exponential case, illustrating useful near-exponential decay with computed values of T (or EDT, as they are the same in this case) = 2.89s (500Hz), and 2.48s, (2000Hz).

(c) A multi-record of the modulation reduction values is plotted on a common set of coordinates for the two bands, 500Hz and 2000Hz; this is useful for averaging out the statistical fluctuations inherent in a noise carrier, in addition to observing odd values which could for example be due to extraneous noise during the measurement, perhaps not noticed by the operator; the example given in Figure 1 is for 5 records.

(d) The "average" modulation reduction curve is then superimposed over the standard set provided with the B&K documentation, which give the case of pure exponential decay, and these pair are conveniently placed under the respective multi-record plots for ease of reference. Where these standard curves depart in shape from the recorded average, it illustrates that the shape of the decay curve is non-exponential, and that the effect of a direct field needs to be considered. Hence, the addition of another curve representing the theoretically expected plot derived for the application of Equation (1), with values of T, D/R and F being included accordingly; the values of T are obtained from a reverberant cut-off decay curve, as shown in Figure 2, which for this example are T = 3.2s (500Hz) and 2.3s (2000Hz) - note similarity with calculated EDT values mentioned earlier, namely 2.89s and 2.48s, respectively, illustrating the suitability of the Steeneken & Houtgast expressions [6], as implemented here.

(e) A plot of the measured RASTI records is also given, with an indication of the time average (0.361), the computed average (0.365), in addition to the standard deviation (0.016), along with other information about the actual measurement in the form of a title block.

Numerous measurements were taken in other areas, with reverberation times of 3.2s and 3.3s (2000Hz), where similar results to the sample given herein have been noted. Invariably, the computed EDT gave results within +/- 5% of the measured reverberation time as recorded from cut-off decays through the system under test, further validating the work of Steeneken and Houtgast under the stated conditions. Well over one hundred test positions were assessed, under different operating conditions.

This of course is to be expected, since under such operating conditions the near exponential decay condition applies, and D/R effect is negligible.

The need to investigate the case of operation under the effect of the direct field is considered next, with the objective of validating the work of the Author, as Steeneken & Houtgast do not provide any special formulae for such cases, relying on the EDT with assumed exponential decay beyond the 10 dB traditional drop. It will be appreciated that the end result in terms of STI is not affected by the analysis proposed here, but as the interest of the Author is to carry out post-processing of results with specific examination of T and D/R, this has been pursued further, in addition to the interest in extending these new formulae for predictive analysis as well, where use of these traditional parameters is considered to be an advantage for ease of quantitative handling [1].

4.2 Operation in the Near-Field

Figures 3-4 give the equivalent to the above analysis for another distributed system, where the D/R values are relatively high. The results indicate a RASTI of 0.64 (i.e. equivalent AL_{cons} of 5.3%), with reverberation times of 2.6s(500Hz) and 2.6s(2000Hz).

As expected the plots of average modulation reduction values onto the B&K family of curves are far from being purely exponential, at both 500Hz and 2000Hz. The calculated FIT values are clearly high, and it is noted that the program does not even attempt to calculate an equivalent EDT in most records.

The multi-record of Figure 3 at 500Hz shows one poor measurement of m3 (0.84), which under normal circumstances would be discarded prior to processing; this has been kept here to show that the resulting effect is small, when averaged, illustrating the usefulness of taking multi-records in practise.

For the case of 2000Hz, a D/R value of 1.1dB is noted which is equivalent to around 3.6dB dip in the cut-off decay curve shown in Figure 4. The expected modulation reduction values derived using Equation 1 are plotted in Figure 3, for T = 2.6s (2000Hz), for which good agreement with the measured value is noted; i.e., an expected STI2000 value of 0.602 as compared with 0.621 by measurement.

A similar plot is given for the 500Hz band, based upon the reverberant decay curve, with T = 2.6s and D/R = 3.5dB (dip of 5.1 dB in the decay curve of Figure 4). The resulting modulation reduction values derived using Equation 1 are plotted in Figure 3, for which good agreement with the measured values is noted; i.e., an expected STI500 value of 0.658 as compared with 0.667 by measurement.

Extensive tests were carried out here [4] for different positions and in various rooms with reverberation time in the range of 2.6 to 3.0s (2000Hz), for which very good agreement with the expected performance has been noted. The ideal model postulated in the analysis by the Author is not strictly applicable, given the effect of early reflections not considered in the model [1], but rather combined in the D/R factor as a specular field. This is the reason for the difference between the "shape" of the expected curve in comparison with the one measured, despite the coherence in the STI value. If viewed with this minor limitation in mind, the usefulness of the proposed formulae is apparent, especially when compared with the exponential decay curves by Steeneken & Houtgast as illustrated in Figure 3.

5. OTHER MEASUREMENTS CARRIED OUT

In addition to the above RASTI measurements carried out at the test site for which Figures 3-4 give RASTI records, other measurements were also carried out, using a TEF analyser in addition to subjective articulation tests, and the comprehensive results of these have already been reported [4]. The results specific to the test position being discussed in this paper are shown in Figure 5, including: ETC records (1/3 octave) at 500Hz and 2000Hz, yielding from the latter 5.3% AL_{cons}, D/R of 1.95 dB, and T of 1.69s. The STI analysis gave an overall index of 0.68. The subjective tests gave an AL_{cons} of 4.7%, with a standard deviation of 1.2. These results may be summarised as follows, including data on separate band analysis where relevant:

parameter	Full index	2000Hz	500Hz
RASTI/B&K	0.64 [5.3%]	0.62 [5.9%]	0.67 [4.5%]
STI/TEF	0.68 [4.3%]	0.68 [4.3%]	0.55 [8.7%]
AL_{cons}/*TEF*	[0.64] 5.3%	---	---
Subjective	[0.66] 4.7% *(σ=1.2)*	---	---

(values in brackets are conversions from the actually measured STI or %AL_{cons})

These results are included here to indicate the extent of ancillary measurements undertaken, and reference may be made to [4] for further details of these. Differences due to the use of 1/3 octave in the TEF/AL_{cons} are of course to be expected, but these are not significant for the present analysis.

6. SUMMARY AND CONCLUSION

It is concluded that the suggested formulae by the Author may be useful in the post-processing of RASTI (and STI) results, wherein the effect of the D/R value is taken into account, along with the reverberation time, these being parameters that are easier to handle than EDT for example, especially where simple type predictive analysis is required.

For this purpose the Author proposes use of a new modified set of curves for modulation reduction, giving D/R as a parameter, which can be easily generated for the actual T being experienced in any specific acoustical environment. Figures 6a/b give examples of these for T = 1,2,3,4s. These curves are presented in a compatible form to the B&K documentation provided for the users of RASTI, with the intent of supplementing the latter where it is appropriate to do so. It will be noted that as D/R tends to infinity, the curves approach the ideal set included in the B&K documentation, and for such cases it is not required to use the modified set.

While the Author suggests use of the new expressions as noted herein, it is to be appreciated that with the advent of computer-assisted acoustical simulation and modelling [1], use of a more accurate EDT/ETC approach is to be preferred, with the MTF being derived directly through appropriate Fourier Transforms, without the need for such simple type expressions. The intent of this paper is to present an example where it may be useful to use such simple procedures, where it is not justified to generate the modelling necessary for detailed analysis, thereby hoping to give the STI method more universal applications in practise.

7. REFERENCES

[1] P. Doany, "A Comparison of STI and AL_{cons} Applications in Speech Intelligibility Analysis", *Proc. Inst. Acoust.*, Vol. 12, pt. 8, pp 143-157, (1990, Nov.).

[2] D. Davis, C. Davis, "Applications of Speech Intelligibility to Sound Reinforcement", *J. Audio Eng. Soc.*, Vol. 37, No. 12, pp 1002-1019, (1989, Dec.).

[3] C. P. Davis, "Measurement of % AL_{cons}", *J. Audio Eng. Soc.*, Vol. 34, No.11, pp. 905-909, (1986).

[4] P. Mapp, P. Doany, "Speech Intelligibility Analysis and Measurement for a Distributed Sound System in a Reverberant Environment", *AES 87th Conv.*, New York City, (1989, Oct.).

[5] *"Speech Transmission Meter Type 3361"*, Instruction Manual, Bruel & Kjaer, (1986, Jan.).

[6] H.J.M. Steeneken, J. Houtgast, "RASTI, A Tool for Evaluating Auditoria", *B&K Technical Review*, No.3, pp 13-39 (inc.appendix), (1985).

INTERPRETATION OF STI MEASUREMENTS

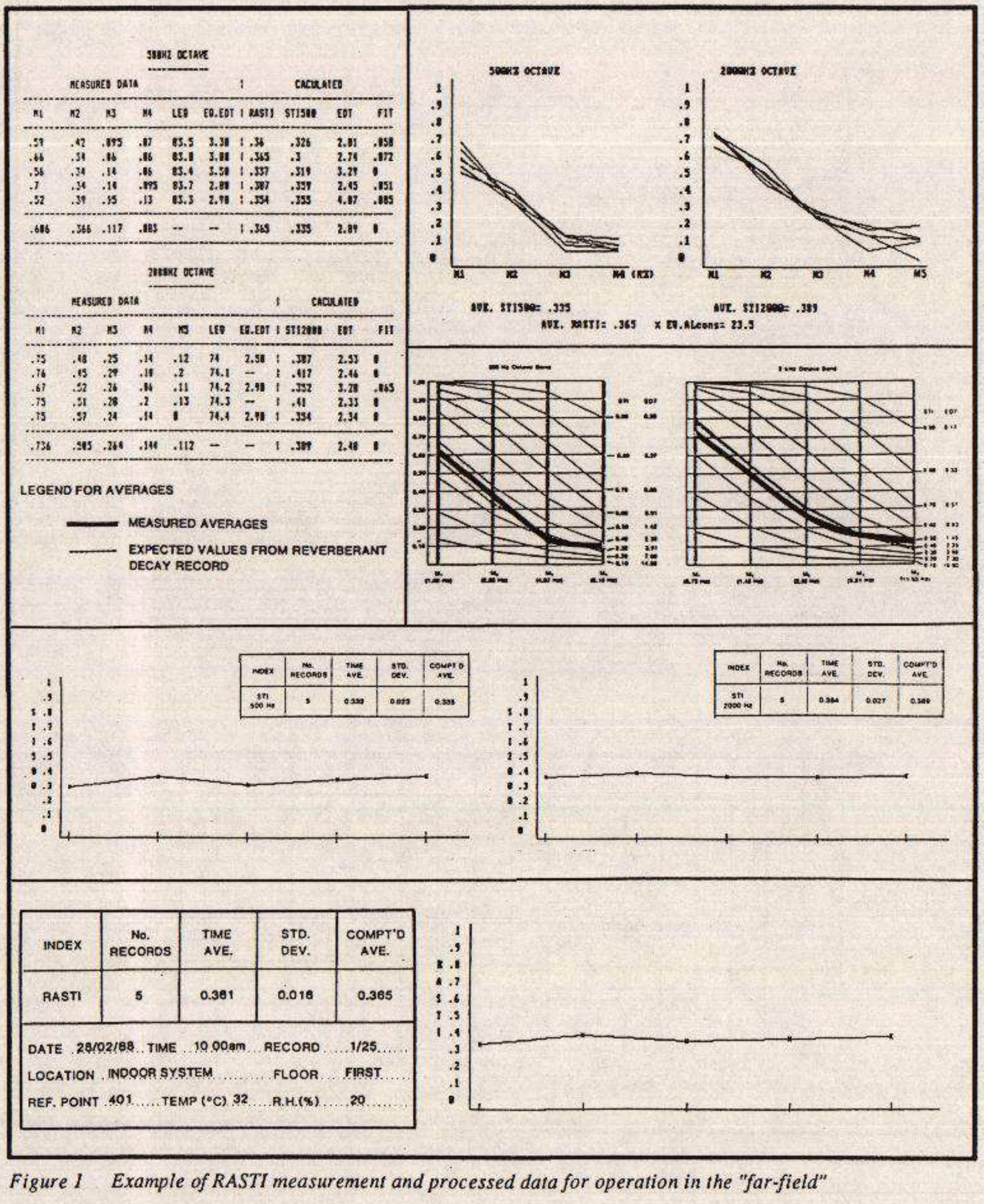


Figure 1 Example of RASTI measurement and processed data for operation in the "far-field"

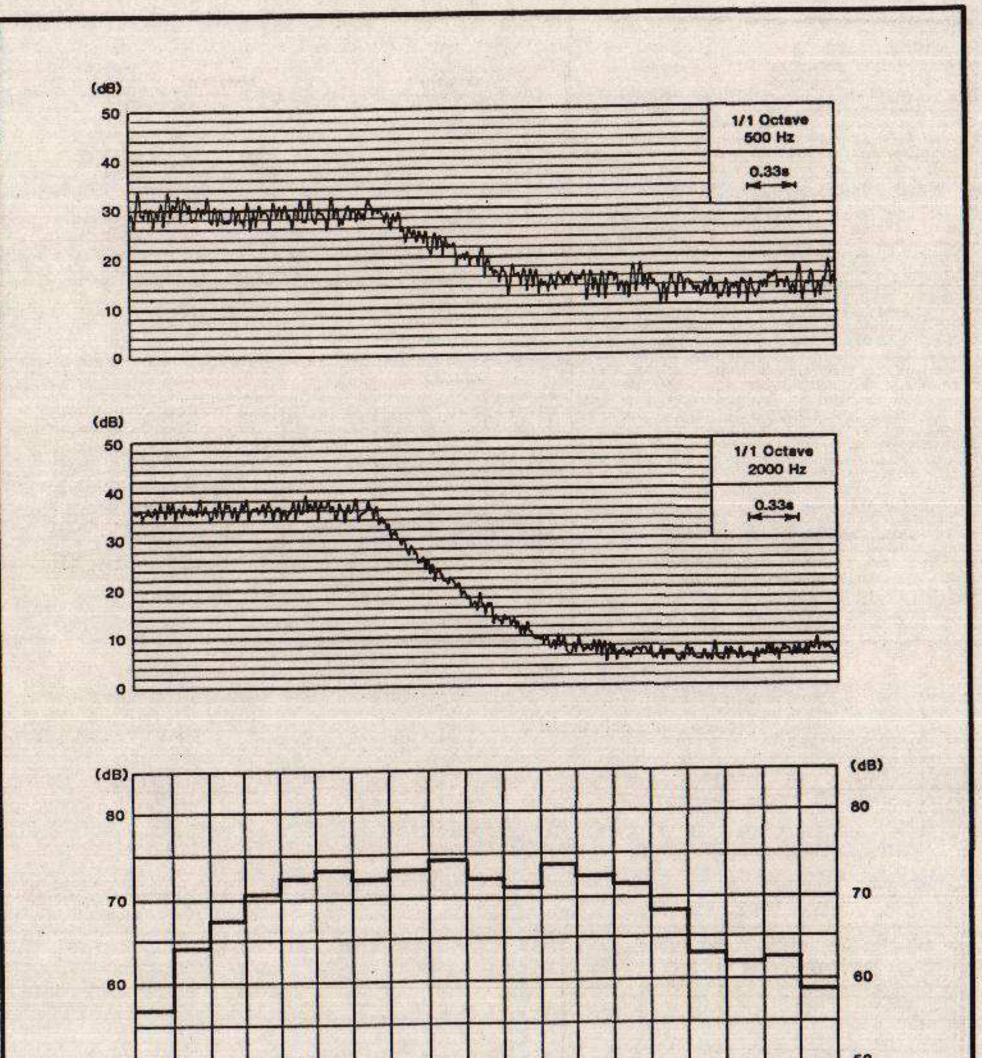


Figure 2 Ancillary measurements for the "far-field" position of Figure 1, giving reverberant cut-off decay and 1/3 octave frequency response.

INTERPRETATION OF STI MEASUREMENTS

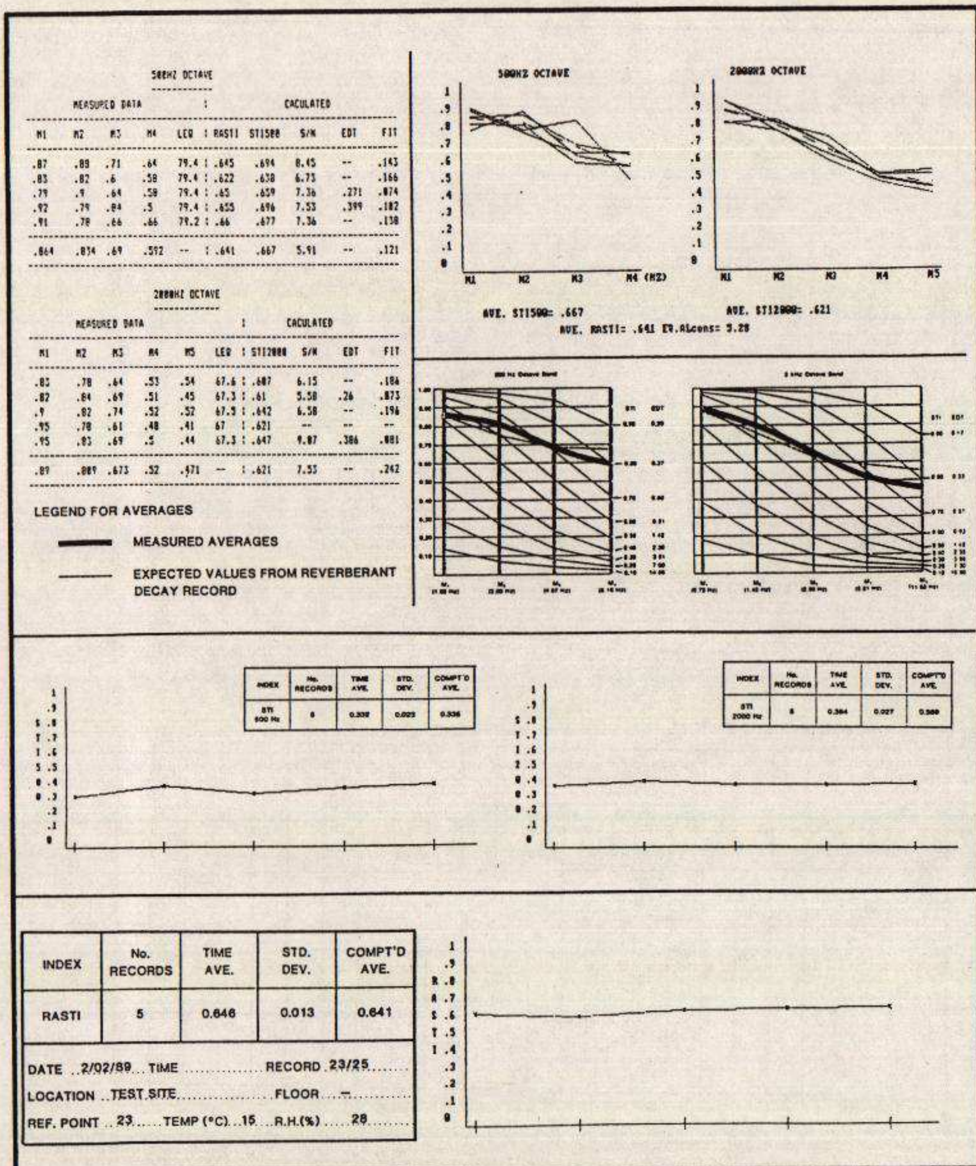


Figure 3 Example of RASTI measurements and processed data for operation in the "near-field"

INTERPRETATION OF STI MEASUREMENTS

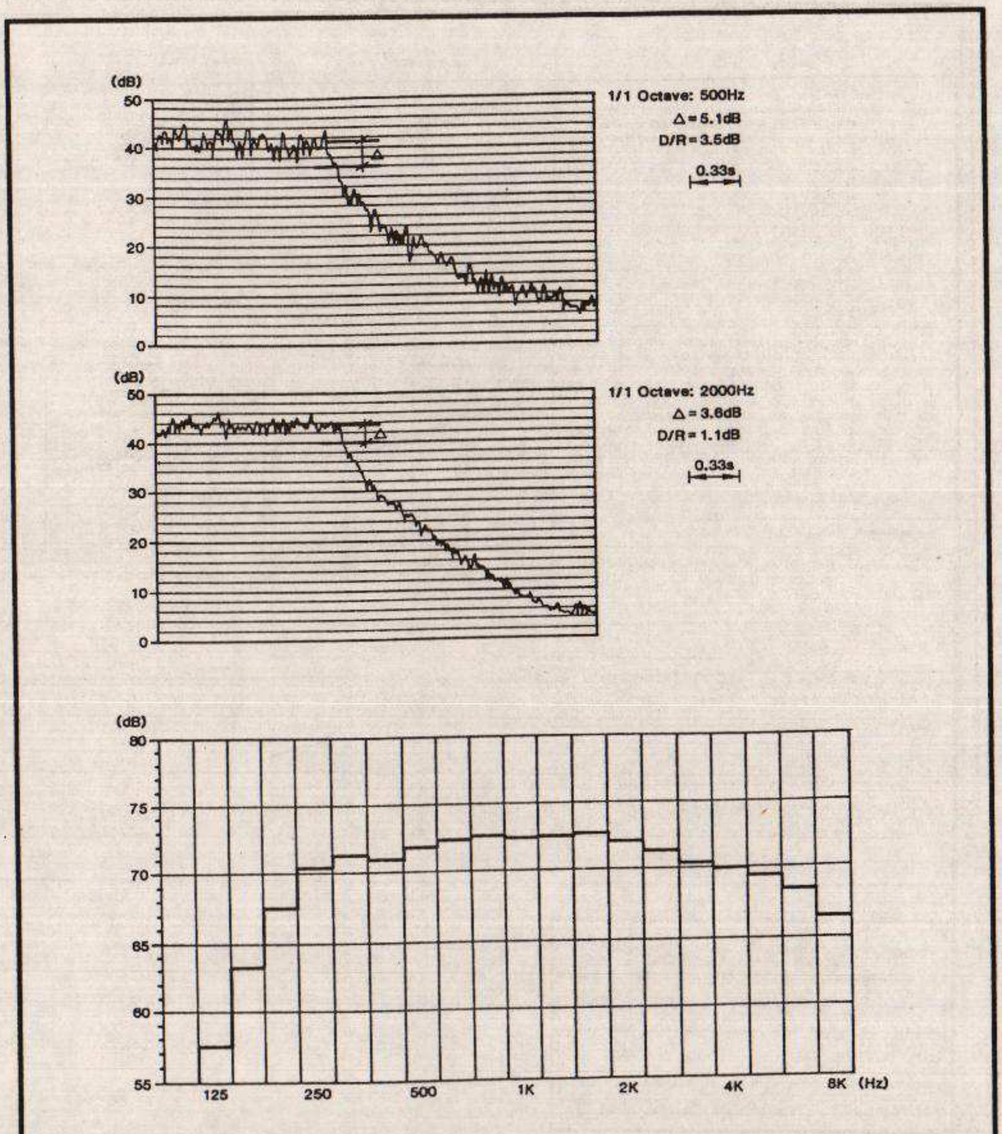
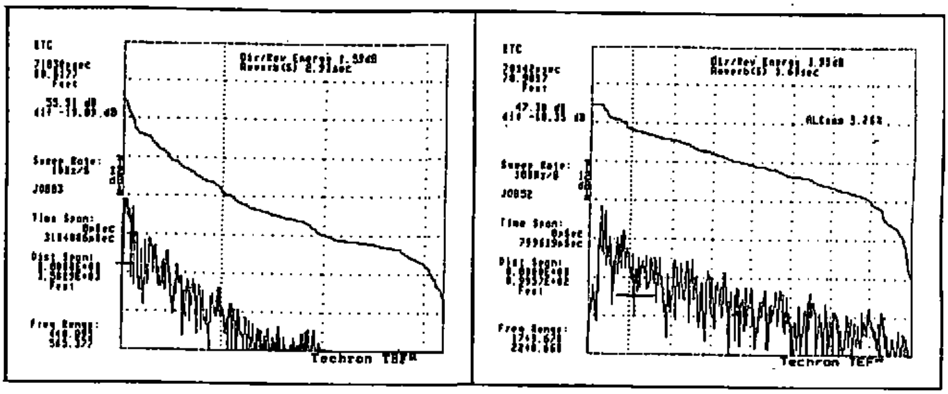


Figure 4 *Ancillary measurements for the "near-field" position of Figure 3, giving reverberant cut-off decay and 1/3 octave frequency response*

INTERPRETATION OF STI MEASUREMENTS



Parameter \ Band	125	250	500	1K	2K	4K	8K	Full Band
Eq-S/N	4.1dB	3.4dB	1.6dB	4.7dB	6.3dB	7.7dB	8.0dB	6.3dB
Eq-EDT	0.76s	0.66s	1.19s	0.70s	0.63s	0.42s	0.40s	0.63s
STI	0.64	0.61	0.55	0.66	0.66	0.76	0.77	0.68

Figure 5 TEF measurements of ETC, ALcons and STI at the "near-field" test position given in Figures 3-4

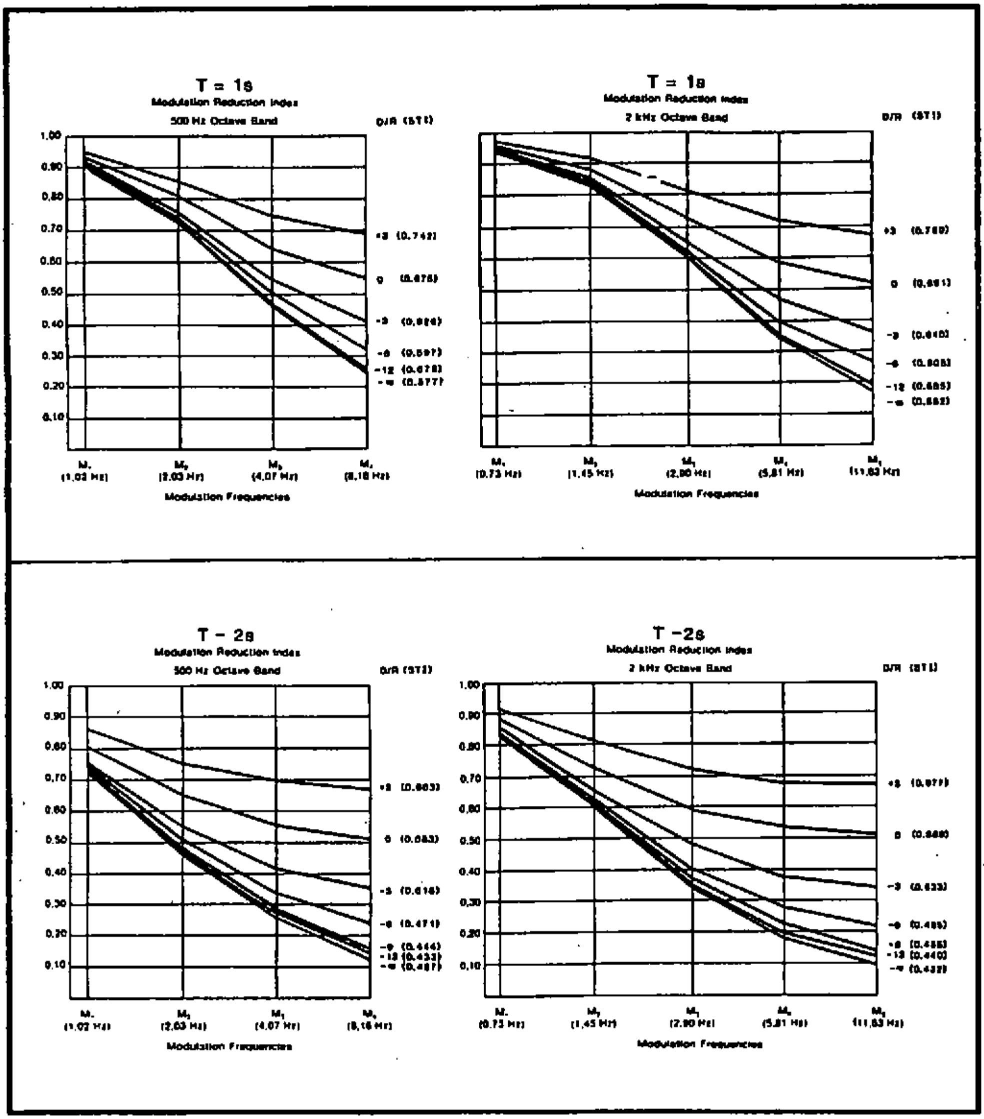


Figure 6a Modified modulation-reduction plots for various values of D/R (note that the plots for D/R of infinity correspond to the ideal exponential decay plots provided by B&K for users of RASTI): values of T of 1 and 2 seconds

INTERPRETATION OF STI MEASUREMENTS

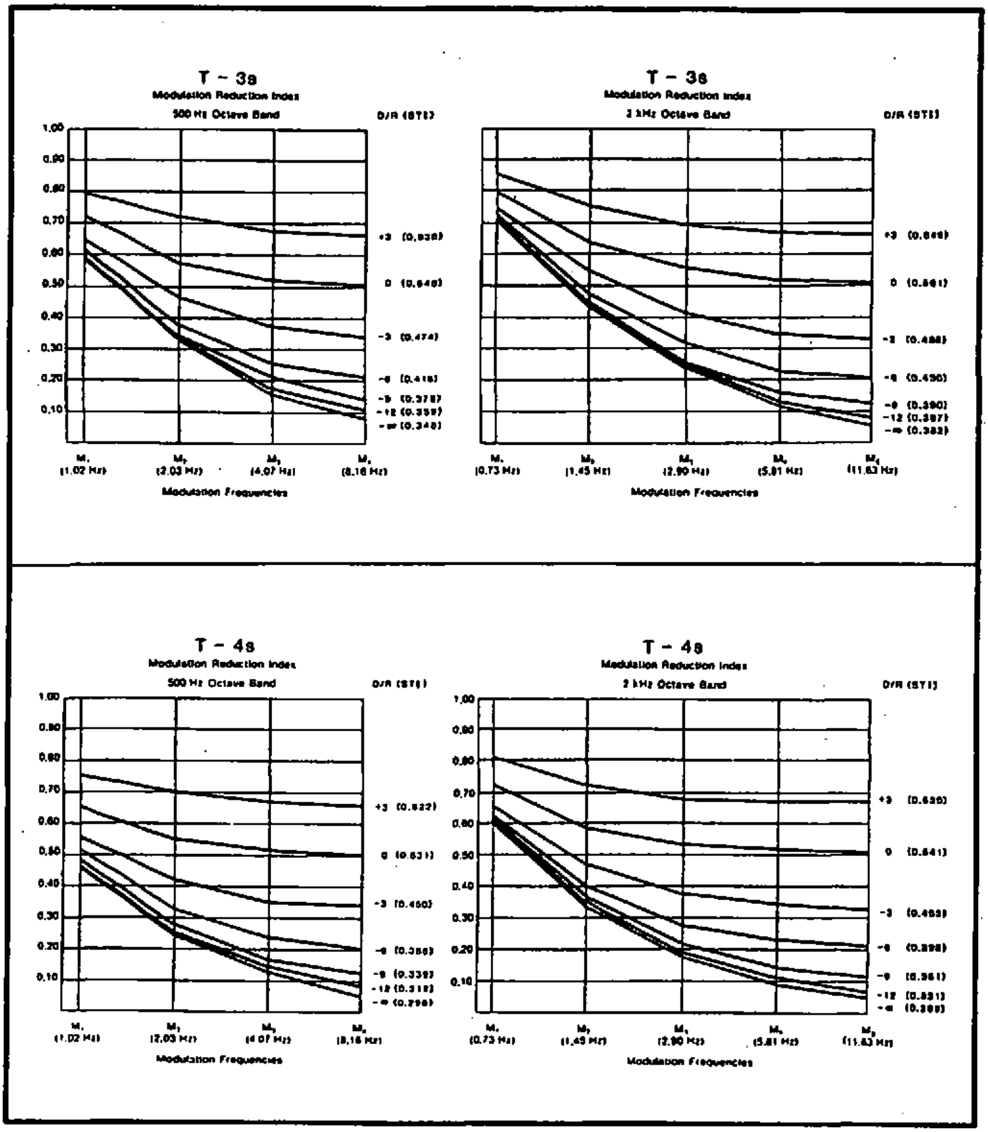


Figure 6b Modified modulation-reduction plots for various values of D/R (note that the plots for D/R of infinity correspond to the ideal exponential decay plots provided by B&K for users of RASTI): values of T of 3 and 4 seconds