From: Proc. 3rd Anglo-Spanish-Netherlands Symposium on Acoustics, 24-25 April 1975, Rotterdam. Published as Publ. nr 33 of Ned.Akoestisch Genootschap, Delft, 1975

NEDERLANDS AKOESTISCH GENOOTSCHAP

SOME EXPERIMENTS WITH THE MTF AS A PREDICTOR OF SPEECH INTELLIGIBILITY IN ENCLOSURES

T. Houtgast and H.J.M. Steeneken
Institute for Perception TNO, Soesterberg, The Netherlands

Summary

For a variety of sound transmission paths in enclosures three types of measurements were preformed: (1) determination of the Modulation Transfer Function MTF, (2) intelligibility measurements, and (3) a traditional determination of reverberation time T. The relation between MTF's and intelligibility scores is found to be very similar to the one determined previously in the laboratory. The relation between MTF and reverberation time T may deviate considerably from the theoretical relation derived for an ideal exponential reverberation curve. This is interpreted as an indication that T does not adequately reflect the extreme importance of the top of the reverberation curve. Finally, it is shown that in case of an ideal exponential reverberation curve different procedures lead to very similar results when the effect of reverberation is expressed in terms of an equivalent S/N ratio.

Introduction

MTF stands for Modulation Transfer Function. It finds its application in several fields, in particular for testing the quality of optical systems. In general terms, the function specifies to what extent the modulation depth of a sinusoidal pattern of high and low intensity at the "input" of a system is transferred, or reproduced at the "output", as a function of modulation frequency. The potential of the MTF approach in the field of room acoustics was shown before (Houtgast and Steeneken, 1973). Fig. 1 illustrates the basic principle. The test signal is wide-band noise of which the intensity is 100% sine-wave modulated with frequency F. During transmission in a room, the signal is received by a microphone. The modulation depth m is determined (after octave-band filtering of the received signal, if desired) and expressed as Modulation Transfer MT in dB. This quantity as a function of modulation frequency F constitutes the MTF. It was shown that in the case of reverberation, echos or interfering noise

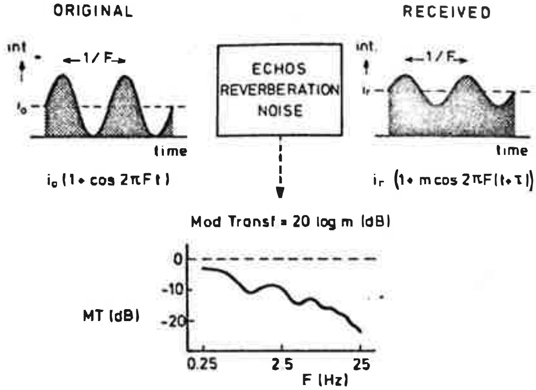

Fig. 1. Principle and definitions underlying the Modulation Transfer Function.

the MTF allows for the estimation of a quality index of a transmission path which correlates well with intelligibility scores measured with talkers and listeners, at least for a variety of conditions set up in the laboratory.

The present study is concerned with some further practical questions related to the MTF approach: (1) experience with the MTF outside the laboratory, and (2) the relation of the MTF with more traditional measures.

Field application of the MTF

The first verification of the validity of the MTF approach, as mentioned above, was carried out in the laboratory with a small-size reverberation room (65 m^3). This section is concerned with the question whether the application to a number of practical situations would raise any unexpected problems.

The measurements were performed in five very different enclosures, with volumes ranging from 250 to 40.000 m^3. An artificial head was used as the sound source, placed at the speaker's position. Signals were delivered by a tape recorder. A microphone was placed at an arbitrary position in the room, and the signal received by the microphone was recorded on tape. In some rooms two conditions were considered, one including the amplification system present in the room (in that case, the microphone belonging to that system was placed in front of the artificial head). In total, eight conditions were considered. For each of these conditions two types of signals were transmitted successively: (1) the test signal for measuring the MTF, and (2) four PB-word lists of fifty words each. The number of conditions was doubled by adding continuous pink noise to the recorded signals. For each of the sixteen conditions thus obtained, the MTF was determined and from this the single value "weighted MTF" was derived in the usual way (ref. 2). The PB-word scores for the sixteen conditions were determined with four listeners.

The results are presented in Fig. 2. For each of the sixteen conditions the weighted MTF and the PB-word score are plotted in a graph which also contains the older "calibration data" based on laboratory measurements. The figure indicates that the present data fit reasonably to the calibration curve. It may be concluded that the field application of the MTF approach does not raise any serious problems.

Theoretical relations between MTF and other measures

The MTF reflects the effects of reverberation and interfering noise in one single curve. There are other procedures for combining the influence of reverberation and noise by which, essentially, a reverberation time is first converted into an equivalent S/N ratio. Then, this equivalent noise

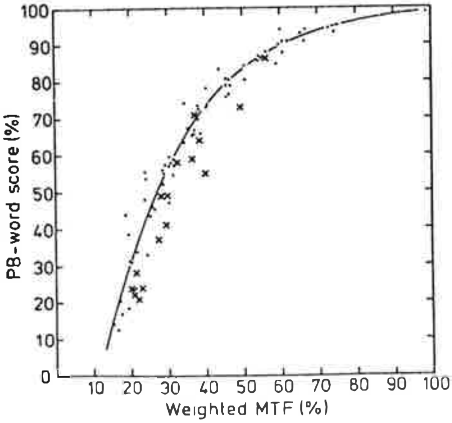

Fig. 2. Relation between the physical index, derived from the MTF, and the intelligibility scores. The sixteen crosses refer to the present study, the curve and the points refer to older laboratory measurements.

is added to the existing interfering noise, which results in a S/N ratio reflecting the effects of both reverberation and noise. To compare different procedures, let us consider the relation between the reverberation time and the equivalent S/N ratio, as predicted by each of the procedures for the theoretical case of an ideal exponential reverberation curve.

MTF procedure. For an ideal exponential reverberation curve, the remaining modulation depth (m in Fig. 1) is a function of the product of modulation frequency F and reverberation time T (ref. 2)

$$m = (1 + 0{,}22F^2T^2)^{-\frac{1}{2}} \qquad (1)$$

If we take F = 4 Hz as the frequency most pronounced in the envelope of running speech (ref. 1), the relation is

$$m = (1 + 3{,}5\, T^2)^{-\frac{1}{2}} \qquad (2)$$

For interfering noise, the remaining modulation depth is a function of the ratio of the noise intensity i_n and the signal intensity i_o

$$m = (1 + i_n/i_o)^{-1} \qquad (3)$$

In combining equations (2) and (3), it follows that the equivalent S/N ratio (i_o/i_n in dB) is equal to:

$$(S/N)_{eq} = -10 \log |(1 + 3{,}5T^2)^{\frac{1}{2}} - 1| \text{ dB} \qquad (4)$$

This relation is plotted in Fig. 3.

The 50-msec boundary. A common way to interpret reverberation in terms of a S/N ratio is to divide the echogram in two parts: all reflections arriving within the first 50 msec are considered as "useful" signal, and all reflections arriving after that interval are considered as "detrimental" noise. For an ideal exponential reverberation curve with reverberation time T, the S/N ratio thus defined is equal to

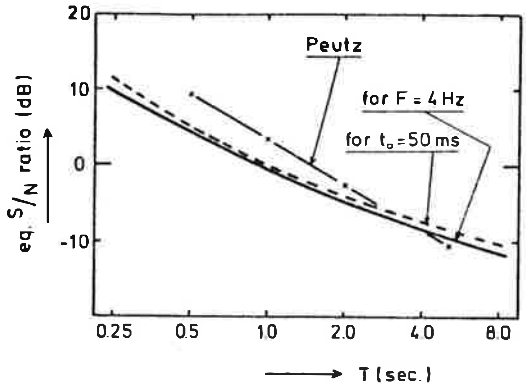

Fig. 3. For the theoretical case of an ideal exponential reverberation curve, the three curves present the relation between reverberation time and equivalent S/N ratio as predicted by three different procedures.

$$(S/N)_{eq} = 10 \log |10^{0,3/T} - 1|$$

This relation is plotted in Fig. 3.

The Peutz procedure. Many data obtained by Peutz on intelligibility scores in conditions of reverberation and interfering noise have been combined in a single graph (Peutz en Klein, 1973). From that, the equivalent S/N ratio can be specified by considering the two limit cases of (1) reverberation without interfering noise and (2) interfering noise without reverberation. The relation thus obtained is plotted in Fig. 3. (The S/N ratios are taken 7 dB lower than specified in the graph; in their figure the speech level refers to an average of the peak deflections of the measuring instrument, whereas in Fig. 3 the level refers to the long-term average intensity of the speech.)

Fig. 3 indicates that, when the effect of reverberation is converted into an equivalent S/N ratio, the three different procedures lead to very similar relationships. Hence, for this theoretical case of an ideal exponential reverberation curve, the MTF approach is similar to more familiar procedures. Of course, the MTF has the attractive feature that in practical situations the different effects are combined almost "automatically", without the need of defining a S/N ratio for running speech subjected to reverberation, or a reverberation time in case of a non-ideal reverberation curve.

T and non-ideal reverberation

The relative weakness of the reverberation time as a single measure for quantifying the effect of reverberation as encountered in practical conditions is illustrated by Fig. 4. The figure presents individual data points obtained in the eight conditions mentioned before (no additional noise). For each of the eight conditions, the five octave bands with center frequencies from 250 up to 4000 Hz were considered individually. For each octave band two types of data were obtained: (1) the reverberation time, determined in the traditional way from recordings of the reverbera-

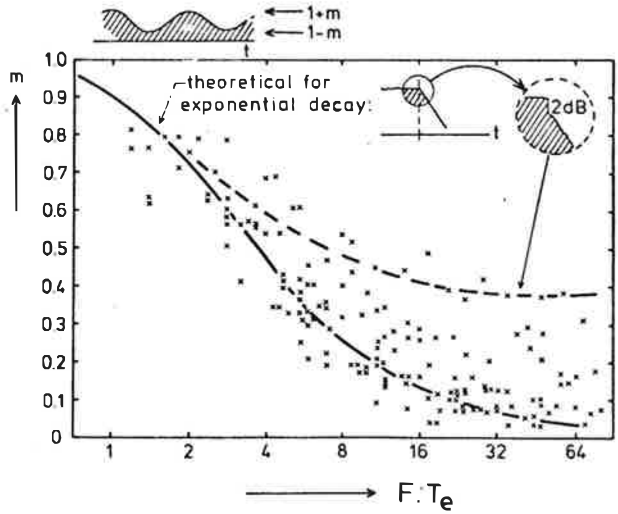

Fig. 4. Each point represents a measurement (for one specific condition) of the reverberation time T_e (estimated from the slope of traditional reverberation curves) and the modulation depth m for modulation frequency F. The uninterrupted curve gives the theoretical relation for an ideal exponential reverberation curve, whereas the interrupted curve refers to the hypothetical case of a reverberation curve with an initial drop of 2 dB.

tion curve (indicated as the estimated reverberation time T_e), and (2) the remaining modulation depth m for a range of modulation frequencies F. The theoretical relation between m and the product FT_e, as given by eq. (1), is represented in Fig. 4 by the uninterrupted curve. The actual data points show a considerable deviation from this theoretical curve. An example of only a slight perturbation of the ideal exponential reverberation curve, which leads to differences of the order of magnitude as observed, is given in Fig. 4 as well. For the theoretical case of a reverberation curve which first shows a initial drop of only 2 dB, the relation between m and FT_e is given by the interrupted curve. This indicates that the first few dB's of the reverberation curve are extremely important in terms of modulation transfer and, accordingly, probably also with respect to speech intelligibility. This is not reflected in the value of T which is, traditionally, based on the slope of the reverberation curve in the region from about -5 to -20 dB.

Conclusions

Field experiments with the MTF did not encounter serious unexpected problems.

When the effect of reverberation is expressed in terms of an equivalent S/N ratio, it is shown that different procedures lead to very similar relations, at least for the theoretical case of an ideal exponential reverberation curve. The results of many measurements in field conditions illustrate that in case of non-ideal reverberation curves the effect of reverberation is not always expressed adequately by the traditional reverberation time.

References

1. Houtgast, T. and Steeneken, H.J.M. (1972). "Envelope Spectrum and Intelligibility of Speech in Enclosures", Proceedings of IEEE-AFCRL 1972, Speech Conference.
2. Houtgast, T. and Steeneken, H.J.M. (1973). "The Modulation Transfer Function in Room Acoustics as a Predictor of Speech Intelligibility", Acustica 28, 66-73.
3. Houtgast, T. and Steeneken, H.J.M. (1973). "Intelligibility in Rooms Derived from Physical Measurements", Proceedings of FASE Symposium on Speech Intelligibility, Liège, 1973.
4. Klein, W. and Peutz, V.A.M. (1973). "Articulation Loss of Consonants Influenced by Noise, Reverberation and Echo", Proceedings of FASE Symposium on Speech Intelligibility, Liège, 1973.