STIPA –
THE GOLDEN MEAN BETWEEN FULL STI AND RASTI

Ole-Herman Bjor, Norsonic AS

1. INTRODUCTION

The speech transmission index, STI, has shown to be a valuable tool for objective rating the speech intelligibility. From the first presentation in Acustica in 1971 [1], the method has been refined and diversified for various applications. This year the International Electrotechnical Commission, IEC, has launched the third revision of the International Standard specifying the method for calculating the index: IEC 60268-16 [2]. Central in the development is the work at TNO-Human Factors, The Nederlands, and the pioneers Tammo Houtgast and Herman Steeneken.

2. SPEECH TRANSMISSION INDEX

The basis for the index is that speech intelligibility to a large extent is based on the slowly modulation of the strength of the sound pressure signal acting as a carrier.

In the described method the carrier is a stationary gaussian noise signal divided in seven bands in octave steps ranging from 125 Hz to 8 kHz. The bandwidth of each band is one-half octave. Each of the bands may be modulated with 14 modulation frequencies. The modulation frequencies are selected in one-third octave steps from 0,63 Hz to 12,5 Hz. This gives a total of 96 combinations.

In this context, the square of the sound pressure is called intensity. The intensity is the quantity being modulated. A small loudspeaker acting as a speaker is normally used for sending out the modulated excitation signal.

The sound in the listener position is received with a microphone. The level and the degree of modulation in each octave band are used for obtaining the speech transmission index. Noise and reverberation in the room will reduce the observed degree of modulation. The method also considers the effect of the most common types of distortions like harmonic distortion and intermodulation. Some other forms of non-linearity, like frequency shifts and frequency multiplications, are not treated effectively.

In order to fully take care of the effects of non-linearity, it is important that the basic signal being modulated is a noise signal with a high crest-factor, a spectral distribution similar to the long-term speech spectrum, and that the main modulation frequency is selected one by one. The measurement of the full STI therefore has to be performed as a sequence of measurements. If each of the 96 combinations is measured in 10 seconds, the total measurement time will be about a quarter of an hour. Such a long measurement time in order to obtain the STI-value in one position of a room limits the applicability of the full STI-method.

The STI-method may be modified in different ways to reduce the measurement time. If the system to be measured is regarded as linear, a number of solutions exist. The excitation signal may be modulated with all modulation frequencies simultaneously and the components may be separated after reception by the use of filters or Fourier analysis. A more common method is to calculate the complex modulation transfer function from the impulse response of the room. M. Schroeder has shown [3] how the modulation transfer function may be calculated from the octave-band filtered impulse response. Douglas D. Rife has shown how maximum-length sequences may be used for measuring the modulation transfer function [4].

The STI value may be calculated from a noisefree impulse response together with information on the level of the received signal and background noise.

The modulation transfer function at frequency, F, may also be calculated from the reverberation time, T, and the effective signal-to-noise ratio S/N (exponential decay is assumed) [5]:

$$m(F) = \frac{1}{\sqrt{1+(2pF\,\frac{T}{13,8})^2}} \cdot \frac{1}{1+10^{(-S/N)/10}} \qquad (1)$$

As seen from this formula, a limited signal-to-noise ratio reduces the modulation transfer function for all frequencies. A long reverberation time reduces the modulation for the highest modulation frequencies.

The STI-value is a weighted average of the different modulation indexes. The last revision of the method also considers masking effects and the absolute threshold of hearing.

3. RASTI AND STIPA

In order to simplify the direct measurement, the RASTI-method (Room Acoustic Speech Transmission Index) was developed at TNO in 1979. Different instruments were developed for the measurement according to this standard. A typical measurement time was 10 to 15 seconds. The RASTI method only considers two octave bands 500 Hz and 2 kHz.

Due to the simplicity in use, the RASTI-instruments were used also for applications beyond the main design goal – room acoustics. The RASTI-value is often used for assessing the quality of public address systems, but comparisons with subjective measurements have shown that the deterioration of speech intelligibility is not handled correct if the PA-system is strongly non-linear or suffers from limited bandwidth.

In order to improve the accuracy in the intelligibility assessment, the STIPA-method was developed. It handles effects due to reverberation in the room and distortions commonly found in public address systems. It also performs well for room acoustics and can therefore in nearly all cases replace the RASTI-method and deliver results more closely to the values obtained by the full STI-method. The measurement time for a STIPA-measurement is similar to the RASTI-method: 10 – 15 sec.

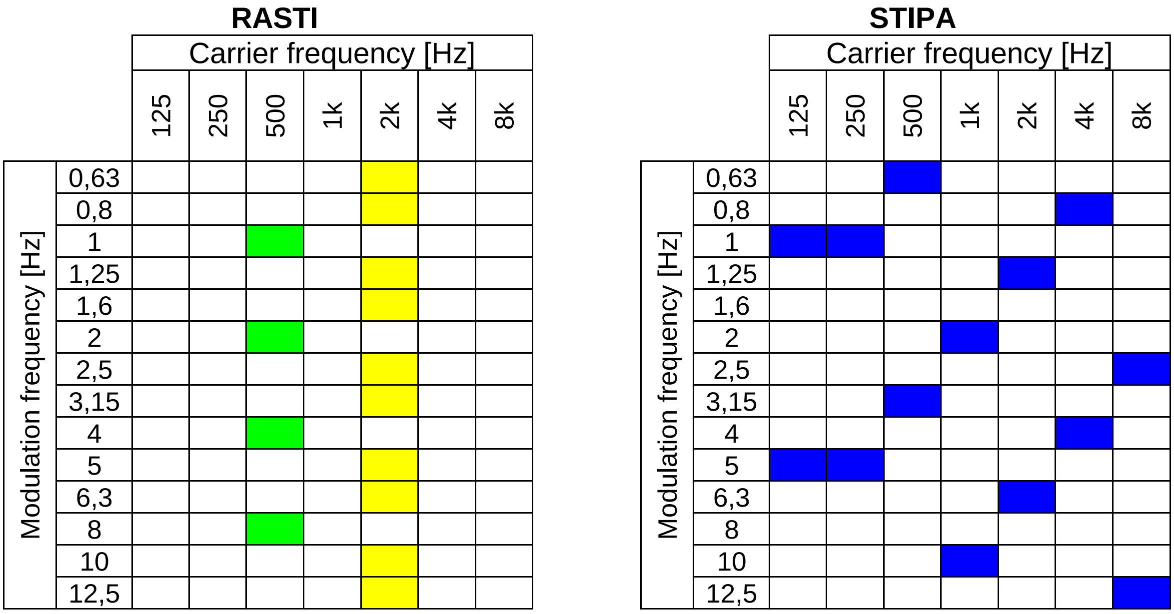

Fig. 1 *Carrier and modulation frequencies for RASTI and STIPA*

4. MEASUREMENT WITH SOUND LEVEL METER NOR118

The STIPA-method may be implemented in a standard sound level meter. In the sound level meter Norsonic Nor118 the method is implemented as a program option in the instrument software.

4.1 EXCITATION SIGNAL

A CD delivered with the instrument option contains the excitation signal. The signal is the sum of six bands of noise, each modulated with two frequencies as specified for the STIPA-method. The shape of the spectrum is specified in the standard. The sound level meter in the normal mode of operation may be used for verifying the spectral weighting. The excitation signal runs continuously and no synchronization between the excitation and the instrument is needed.

4.2 MEASUREMENT

After the mode of operation is selected, pressing the START button initiates a measurement. The measurement runs for about 13 seconds. At the end the STI-value calculated according to the STIPA-method is presented on the screen together with a statement of the assessment; “Excellent”, “Good”, “Fair”, “Poor” or “Bad”. Selecting the table display allows the mean level in each band to be displayed.

During the measurement the short time equivalent-level in each octave band is measured with a time-resolution of 5 milliseconds. The Figure 2 shows a typical level versus time diagram for the 500 Hz octave. The level is converted to squared sound pressure, Pa^2, or

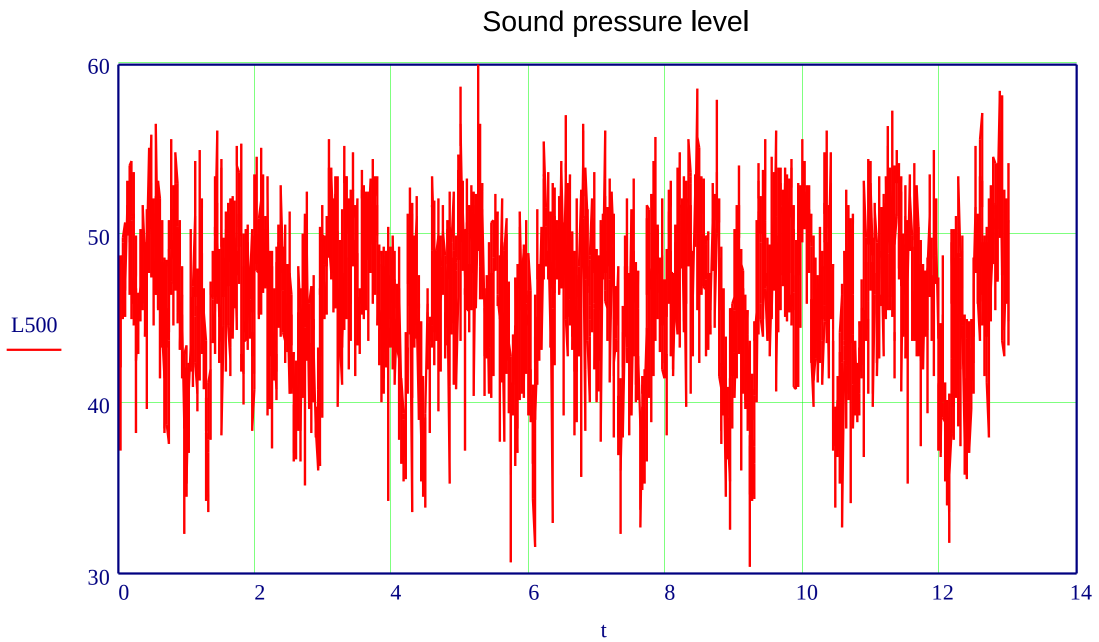

Fig. 2 *Typical sound pressure level in the 500 Hz band as function of time (seconds)*

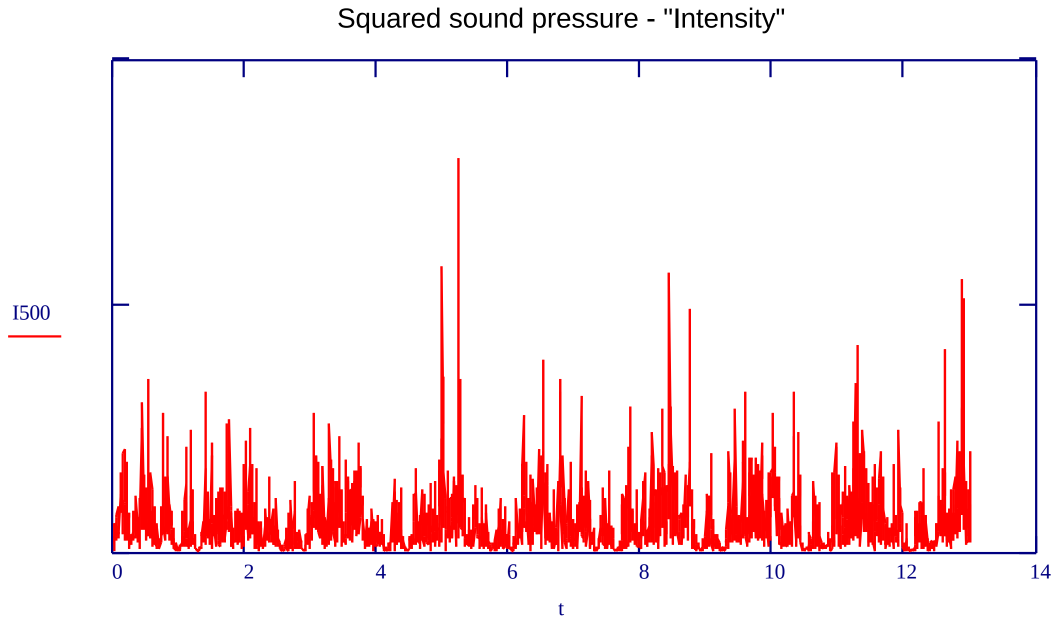

Fig. 3 *Typical sound "intensity" in the 500 Hz band as function of time (seconds)*

"intensity" in this context. See figure 3. The degree of modulation for each frequency is found by Fourier transformation. By comparing the measured degree of modulation with the degree of modulation in the excitation (55%), the value for the modulation transfer function is obtained. In total, twelve combinations of carrier/modulation frequencies are measured.

The measured modulation transfer function is corrected in order to compensate for the common threshold of hearing and for the masking effect in the human auditory organ.

Although not a part of the International Standard, the implementation in the instrument allows the addition of a specified background noise level – one level for each octave. The correction is done according to equation (2):

$$mc_{k,f} = m_{k,f} \frac{I_k}{I_k + Irs_k + Ino_k + Iam_k} \qquad (2)$$

where

$mc_{k,f}$ is the corrected modulation transfer function for octave band number k and modulation frequency *f*.

$m_{k,f}$ is the measured modulation transfer function.

I_k is the measured "intensity" in octave band number *k*

Irs_k is the "intensity" in octave band number *k* related to the hearing threshold

Ino_k is an optional "intensity in octave band number *k* corresponding to a specified background noise level. If not used this value is zero.

Iam_k is a "intensity" in octave band number k used to mimic the masking effect in the auditory organ. The value is a function of the level in the adjacent lower octave band.

Note that the noise correction is an extension from the method specified in IEC 60268-16. This allows a measurement in a situation without the normal background noise and recalculation of what the STI-value would have been with a specified noise level.

The results may be stored in the instrument and later downloaded to a computer for report generation.

5. CONCLUDING REMARKS

Implementation of the STIPA-measurement mode in an ordinary sound level meter allows an easy and fast measurement of speech intelligibility. Normally, the STI-value is measured with the normal background noise and the value indicates the speech intelligibility with this level of background noise. For situation where the measurement is performed in an untypical quiet situation a synthetic background noise may be added and the STI-value related to this situation is indicated. This feature is especially useful when assessing speech intelligibility in auditoriums without an audience.

Fig 4
Sound level meter Nor118

6. REFERENCES

[1] T. Houtgast and H.J.M. Steeneken, (1971). "Evaluation of Speech Transmission Channels by using Artificial Signals". Acustica vol. 25, pp355–367.

[2] IEC 60268-16 (2003-05) Sound system equipment - Part 16: Objective rating of speech intelligibility by speech transmission index.

[3] M.R. Schroeder(1981). "Modulation Transfer Functions: Definitions and measurements" Acustica, vol. 49, pp179–182

[4] D. D. Rife (1992). "Modulation Transfer Function Measurement with Maximum-Length Sequences". J. Audio Eng. Soc. Vol 40, No. 10 1992 October, pp779–789.

[5] T. Houtgast, H.J.M. Steeneken and R. Plomp (1980). "Predicting Speech Intelligibility in Rooms from the Modulation Transfer Function. I. General Room Acoustics". Acustica vol 46, pp60–72.