Less noise • More sound

Introducing Speech Intelligibility

Public address systems in building complexes have to inform persons about escape directions in case of emergency. Such public buildings include airports, railway stations, shopping centers or concert halls. However if such announcements are misunderstood due to poor system quality, tragic consequences may result. Therefore, it is essential to design, install and verify sound reinforcement systems properly for intelligibility. In addition, a variety of other applications such as legal and medical applications may require intelligibility verification.

National standards - see Figure 1 - require the verification of electro acoustic sound systems for emergency purposes under realistic circumstances to ascertain a minimum level of speech intelligibility in case of an emergency. Thereby, speech intelligibility from a regulatory view is not a subjective measurement, but can be verified with several, more or less complex methods that have been standardized in IEC 60268-16.

Subjective analysis methods

Although frequency response, reverberation, distortion, signal-to-noise ratio or loudness are related to intelligibility, the conventional measurements of these parameters together only marginally relate to intelligibility. When added issues, such as directionality of drivers and the environment conditions are taken into consideration, the question is: How well a spoken message can be understood at different locations?

National standards:	
ISO 7240	*Fire detection and alarm systems, section 16 & 19*
NFPA 72	*National Fire Alarm Code 2002 (2002 edition, section 7.4.1.4)*
BS 5839-8	*Fire detection and alarm systems for buildings. Code of practice for the design, installation and servicing of voice alarm systems*
DIN 60849	*System regulation with application regulation DIN VDE 0833-4*

Figure 1: *A minimum of speech intelligibility is standardized*

The fundamental approach measuring intelligibility is to let a trained human speaker read a number of existing or synthetic words, whereas a representative number of listeners individually write down what they believe of having understood. The statistical analysis of these notes results in a value representing the percentage of words being understood correctly. Standardized procedures according to this method are PB-words, CVC or SRT (Speech Reception Threshold). However, conducting such tests is rather time consuming and costly, as well in some hazardous locations even impossible. Therefore, these methods are mainly used to verify alternate measurement methods.

Technical Methods

Already back in 1940, Bell Laboratories started to develop measurement technologies to determine the speech intelligibility. Nowadays, highly developed algorithms as SII (Speech Intelligibility Index) and various forms of the STI (Speech Transmission Index) allow measuring speech intelligibility. These techniques take care of many parameters which are important for intelligibility such as:

- Speech level
- Background noise level
- Reflections
- Reverberation
- Psychoacoustic effects (masking effects)

The basic idea of STI measurement consists in emitting a synthesized test signals instead of a human speakers voice. The speech intelligibility measurement acquires and evaluates this signal as perceived by the listeners ear. Extensive investigations have shown the relationship between the alteration of speech characteristics and the resulting speech intelligibility. These findings are incorporated into the speech intelligibility meter that is able to display the intelligibility result as a single number between 0 (unintelligible) and 1 (excellent intelligibility).

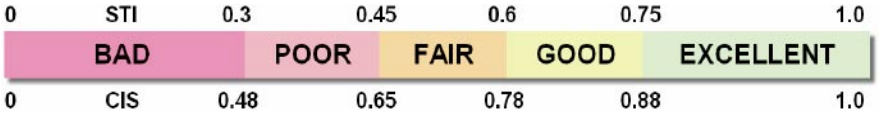

Figure 2: Speech Intelligibility may be expressed by a single number value. Two scales are most commonly used: STI and CIS (Common Intelligibility Scale)

STI, RASTI or STI-PA are the most established methods for measuring speech intelligibility. All of them basically apply the same principle, whereby RASTI and STI-PA are a simplified version of STI. Further on the principles behind these methods are explained.

Speech model

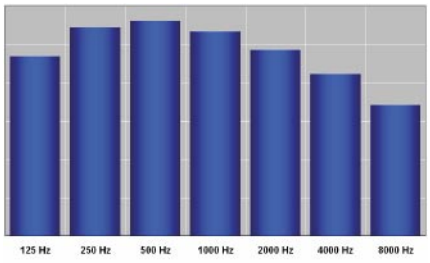

Figure 3: *Average octave band spectrum of a male speaker*

First of all, measuring the speech intelligibility requires a model for speech signals. For instance, speech may be described as a time evolution of various spectra. Superposition of all spectras defines the longterm speech frequency spectrum. Evolution in time is modulated as a set of intensitiy modulation.

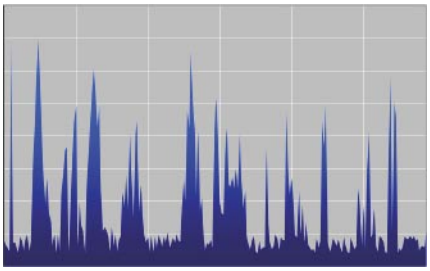

Figure 4: *Envelope of a speech signal (250 Hz band).*

Frequency spectrum

The spectral analysis of a male voice averaged over a longer time results in a typical characteristic as shown in Figure 3.

Time modulation

Level of frequency components varies, i.e. is "modulated" by the speaker. Figure 4 shows the envelope of a speech signal in the 250 Hz octave band. The shape of the envelope is given by averaging the time evolution of the speech contents.

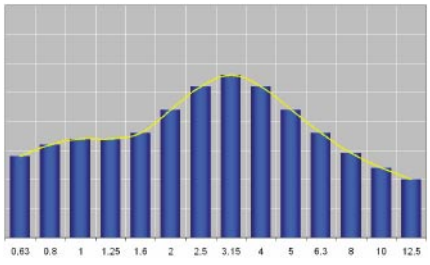

Figure 5: *Frequency spectrum of the envelope (250 Hz band).*

Analyzing the spectra of time modulation intensity shows that a speaker modulates the speech spectra with frequencies in the range from 0.1 to 24 Hz. A set of modulation frequencies from 0.63Hz to 12.5Hz sufficiently represents these modulations.

Modulation Transfer Function (MTF)

High speech intelligibility needs the spectral intensity modulation and the overall spectrum being preserved at reaching the listeners ears. Therefore, the three core intelligibility measurement methods STI, RASTI und STI-PA are based on measuring the MTFs (Modulation Transfer Functions) in 7 octave bands. For each octave band is one MTF quantifying the preservation degree of the intensity modulations in this band. These functions quantify how much the intensity modulations are preserved in 7 octave bands covering the longterm speech spectrum.

Figure 6: *Reverberation, background noise and reflection are responsible for degrading of the modulation index.*

Figure 7 shows the MFT of one octave band. This is derived from measuring the 1/3rd octave modulation frequencies, thus resulting in 14 frequencies between 0.63 and 12.5 Hz. Each modulation transfer function determines how well the modulations are preserved in the associated octave band.

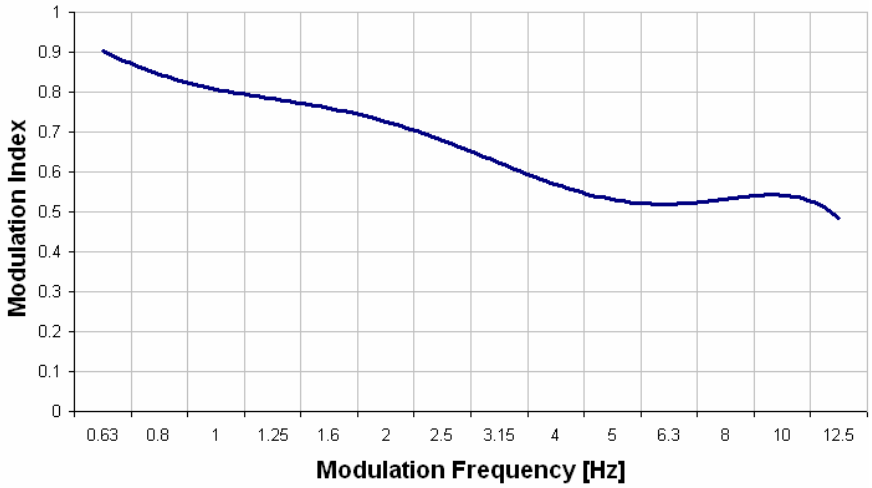

Figure 7: *Modulation Transefer Function for one octave band*

Based on the MTF results as well as sound pressure level, octave band depending hearing threshold, frequency response and psycho acoustic effects (masking effects) it becomes possible to reliably determine the preservation of speech intelligibility from speaker to listener. The calculations are based on extensive and profound evaluations and comparisons with subjective methods.

Measuring the complete MTF – as required for STI – can become rather time consuming. For instance, 14 * 7 = 98 individual measurements must be executed, thus resulting in a total acquisition duration of 15 minutes. Therefore, different approaches have been developed to reduce test duration and to enable speech intelligibility measurements with portable instruments.

STI - Speech Transmission Index

The STI result is based on the full set of 98 measurements. Since this approach requires a rather long test period, it is less frequently applied in practice. However, STI represents the most detailed method to measure the preservation of speech intelligibility during transmission and is mostly used if alternative approaches don't provide reliable results due to unfavorable environmental conditions.

	Modulation Frequencies													
	0.63 Hz	0.8 Hz	1 Hz	1.25 Hz	1.6 Hz	2 Hz	2.5 Hz	3.15 Hz	4 Hz	5 Hz	6,3 Hz	8 Hz	10 Hz	12.5 Hz
Octave Bands 125 Hz	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
250 Hz	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
500 Hz	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
1 kHz	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
2 kHz	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
4 kHz	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
8 kHz	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

Figure 8: *STI considers all 14 modulation frequencies and all 7 octave bands resulting in 98 modulation index results.*

In practice, the STI result is mostly calculated from the impulse response (MLSA) that has been acquired e.g. with a PC-based system. This approach is much quicker, but requires post-processing with spectral frequency weighting and lot of experience. The measurement assumes a linear behavior of the setup, i.e. there must be no non-linear processing or conditions, including compressors or limiters and close to zero wind speeds, which is a rather rare situation. Microphone and speakers aren't allowed for movements during measurement. As handheld instruments aren't fixed during measurment, it doesn't make sense to support MLS testing in handheld instruments.

RASTI- Room Acoustics Speech Transmission Index

RASTI has been developed for special cases, such as a human lecturer speaks into a small room with neglectabe echoes, without electro-acoustic systems. RASTI acquires only few segments of a complete MTF table, which obviously represents an extreme simplification of STI. Therefore, tight restrictions must be met to acquire reliable speech intelligibility results with RASTI. Furthermore, the RASTI result does not consider significant parameters such as the frequency response, echoes or frequency-dependant reverberation times. For a RASTI measurement, only two simultaneously generated frequency bands are considered, i.e. the 500 Hz and the 2 kHz band which then are modulated with four and five frequencies respectively.

	Modulation Frequencies												
Octave Bands	0.63 Hz	0.8 Hz	1 Hz	1.25 Hz	1.6 Hz	2 Hz	2.5 Hz	3.15 Hz	4 Hz	5 Hz	6,3 Hz	8 Hz	10 Hz	12.5 Hz
125 Hz														
250 Hz														
500 Hz			✓			✓			✓			✓		
1 kHz														
2 kHz	✓			✓			✓			✓			✓	
4 kHz														
8 kHz														

Figure 9: *RASTI uses 9 different modulation frequencies in 2 octave bands. The yellow marked octave bands and modulation frequencies are not considered by RASTI.*

The practical application of RASTI is mainly restricted to qunatify the intelligibility index of the channel between two persons. However, RASTI used to be the only method to measure the speech intelligibility with a portable instrument for a long time, thus it has been utilized in the aviation industry to measure public announcement systems.

STI-PA - Speech Transmission Index for Public Address Systems

A rising awareness for security issues, new technological means and the shortcomings of RASTI triggered the speaker manufacturer Bose and the research institute TNO to develop a new method for speech intelligibility measurements of PA installations. The result of these efforts is STI-PA, which allows quick and accurate tests with portable instruments.

Like RASTI, STI-PA applies a simplified procedure to calculate the MTF. But STI-PA determines one MTF by analyzing all seven frequency bands, whereby each band is modulated with two frequencies.

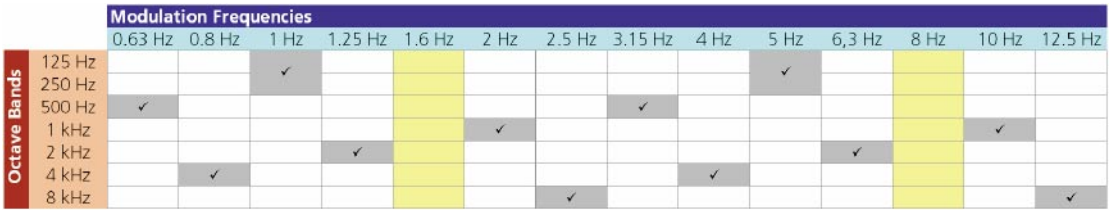

Figure 10: *The IEC60268-16 describes a STI-PA method where the 125Hz band and 250Hz band are combined and the yellow marked modulation frequencies are not considered.*

Supposing that no severe impulsive background noise is present and that no massive non-linear distortions occur, STI-PA provides results as accurate as STI. If however impulsive background noise is present during the normal system operation hours, it is usually possible to mitigate the effects by also acquiring a measurement at a more favorable time e.g. under slightly different conditions in the area, or during the night time - and to calculate an unbiased overall measurement by using the results of both test cycles.

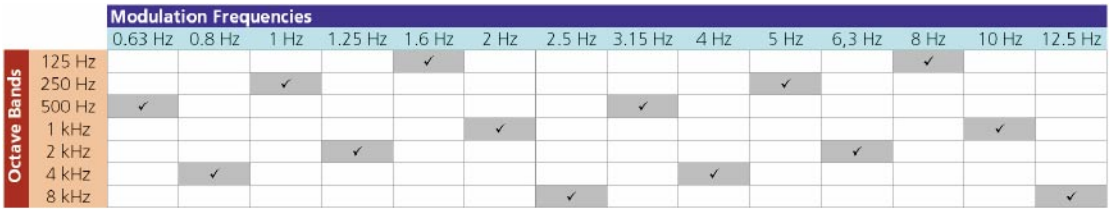

Figure 11: *The NTI-STI-PA method (verified by TNO) considers all 7 octave bands and all 14 modulation frequencies resulting in slightly more accurate results than the IEC STI-PA method.*

Further Information:

Detailed information on speech intelligibility measurements are contained in the IEC60268-16 (2003-5) standard, which also describes the test procedures and the requirements in practice.