Speech Intelligibility

Measurement with XL2 Analyzer

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.

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. This application note will explain the principles behind these methods.

Index

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.

Other national or local regulatory bodies implement recommendations or requirements to conduct these measurements for maintaining minimum speech intelligibility.

Standard: IEC 60268-16, Objective rating of speech intelligibility by speech transmission index

National standards:	
ISO 7240	Fire detection and alarm systems, section 16 & 19
NFPA 72	National Fire Alarm Code (2010 edition)
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

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?

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).

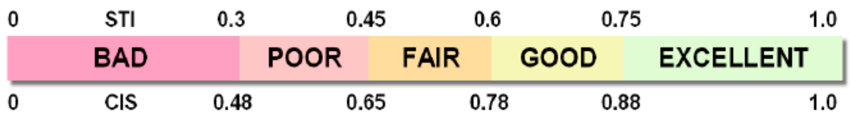

Figure 2: Speech Intelligibility may be expressed by a single number value. Two scales are most commonly used: STI and CIS (Common Intelligibility Scale)

Speech model

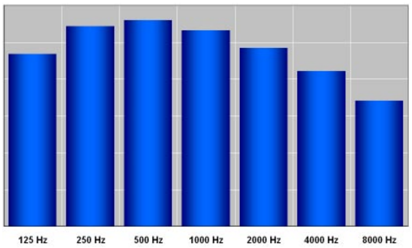

Figure 3: Average octave band spectrum of a male speaker

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 long-term speech frequency spectrum. Evolution in time is modulated as a set of intensity modulation.

Time modulation

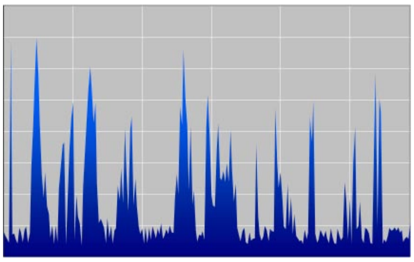

Figure 4: Envelope of a speech signal (250 Hz band).

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.

Frequency spectrum

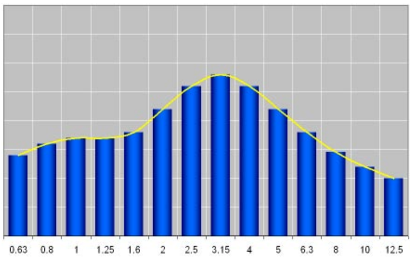

Figure 5: Frequency spectrum of the envelope (250 Hz band).

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

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 and 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 long-term 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.

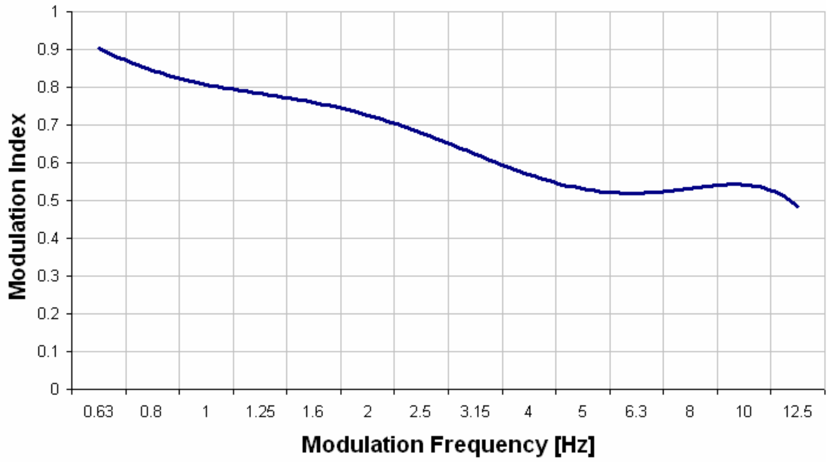

Figure 7: Modulation Transfer Function for one octave band

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 measurement, 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 speaking into a small room without echo’s, but not for electro-acoustic systems.

In order to cope with the long test time required for each STI measurement, a faster method called RASTI was developed. But this in turn weakens its ability for comprehensive testing and heavily compromises its resistance against outside interference. This also leads to poor correlation between subjectively evaluated STI and RASTI. However, RASTI used to be the only method to measure the quality of speech transmission with a portable instrument for a long time, thus it has been utilized in the aviation industry to measure public announcement systems, disregarding the above mentioned development restrictions of RASTI.

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.

Octave Bands \ 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
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 frequency

The practical application of RASTI is mainly restricted to quantify 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

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.

Octave Bands \ 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
125 Hz / 250 Hz			✓							✓				
500 Hz	✓							✓						
1 kHz						✓							✓	
2 kHz				✓							✓			
4 kHz		✓							✓					
8 kHz							✓							✓

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.

Octave Bands	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
125 Hz					✓							✓		
250 Hz			✓							✓				
500 Hz	✓							✓						
1 kHz						✓							✓	
2 kHz				✓							✓			
4 kHz		✓							✓					
8 kHz							✓							✓

Figure 11: The NTi Audio-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.

How does STI-PA compare to STI and RASTI

STI measured in public address systems has been very time consuming. A complete set of 98 measurements of modulation transfer functions (MTF) has to be obtained and summed. Due to the complex nature and the time required almost no really useful STI measurement systems were available for years. With the appearance of MLS based systems, STI was more often obtained, as it can be calculated out of the transfer function, as long as the entire system is strictly linear and synchronous, i.e. there must not be any non-linear processing or conditions, including compressors or limiters and close to zero wind, which is a rather rare situation. Microphone and speaker aren't allowed for movements during the measurements, which prohibits employment of handheld instruments. Thus it doesn't make sense to support MLS measurements in handheld instruments. Alternatively by using the dedicated STI-PA test signal, measurements can be accomplished with handheld instruments.

STI-PA, a derivative of STI, has been developed specifically to cope with the non-linear processing environment common to advanced sound systems, and to reduce the measurement time required to a practical level.

Calculation of % Alcons from STI-PA Measurement

Alcons (%) = 10^((1-STI)/0.45)

The calculation of STI-PA based on Alcon measurements is not reasonable due to the difference in the measurement principles.

Now STI-PA

A rising awareness for security issues, new technology, and the shortcomings of RASTI together triggered the speaker manufacturer Bose and the research institute TNO to develop a new method for measuring the quality of speech transmission of PA installations. The result of these efforts is STI-PA, which supports fast and accurate tests with portable instruments. STI-PA stands for Speech Transmission Index for Public Address systems.

Portable STI-PA analyzers, e.g. NTi Audio's XL2 Audio and Acoustic Analyzer, are able to evaluate speech intelligibility within 15 seconds per room position and are thus well suited for wide-area measurements and high productivity.

Who can and should conduct STI-PA measurements?

Even though the background of the STI-PA method is complex, the operation of STI-PA using the XL2 Audio and Acoustic Analyzer is very simple. Operators with a basic acoustic knowledge can easily conduct these measurements. The instrument's internal storage functionality also simplifies the measurements in larger buildings, where many measurements at many locations must be taken. The detailed access to the measured MTF (Modulation Transfer Function) matrix enables experts to post-process all measurement data.

Is this a research product or widely used?

STI-PA is the standardized result of extensive scientific research. But unlike RASTI, within two years as many as four international test instrument manufacturers have implemented the STI-PA approach and offer varying solutions. It is therefore perfectly valid to say that STI-PA is the widely accepted standard for measurements of the quality of speech transmission, combining the accuracy and advantages of full STI measurements with the benefit of extremely short measurement time of only 15 seconds per location.

Who is TNO?

TNO is a research and certification institute in the Netherlands, Europe. They are focusing on research around defense, security and safety and they have originally developed the STI as well as the STI-PA technology. Their knowledge is freely useable and not patent protected.

Compatibility?

Thanks to TNO, acting as a certification body, it is guaranteed that instruments from all certified test and measurement vendors will provide compatible measurement results. Three vendors are currently certified by TNO. Studies and comparisons conducted by Peter Mapp Associates, Colchester, Essex UK, confirmed that all certified vendors provide stable and comparable measurement results. Details of the comparison may be found in the AES publication titled "Is STI-PA a robust measure of speech intelligibility performance?"

Patent protected?

Even though the research part of STI and STI-PA is not patent protected, BOSE of America has been recently been granted the US-Patent 6,792,404B2 for the idea to implement STI or STI-PA onto a hand held analyzer. NTi Audio maintains a license agreement with BOSE for this patent and is therefore able to market the STI-PA implementation available for the XL2 Audio and Acoustic Analyzer in the US.

Can I buy STI-PA for my XL2?

Yes, STI-PA is an optional function for the XL2 Audio and Acoustic Analyzer. Any XL2 user may obtain a STI-PA license. With the key of the license he may request the activation key for his XL2 Audio and Acoustic Analyzer and full functionality is then activated.

What is a TalkBox? / Do I need a TalkBox for STI-PA?

The TalkBox is NTi Audio's calibrated acoustical sound source with built in digital solid state signal generator. You don't necessarily need a TalkBox if you are testing only the portion of the system beyond the microphone. But the use of the NTi Audio TalkBox as speaker substitute is advisable if:

- Regulations require a complete end-to-end system check in-

NTI TalkBox

cluding the microphone. This is the most realistic system check in any event.

- No electrical input is available to induct the electrical test signal.
- The level of the test signal is not clearly defined
- The characteristics of the speaker's acoustical environment are not negligible and flat.
- The characteristics, sensitivity and frequency response of the speaker's microphone is not known but needs to be considered.
- As above, if for any other reason it is desirable to test the entire signal chain under real conditions.
- The TalkBox is also capable of delivering white and pink noise and other special signals, and so is a very useful overall tool for system tuning and testing.

STI-PA Test Result

- The intelligibility index is measured in the range from 0 to 1, whereby 1 is perfect and the minimum requirement including measurement uncertainty and variation is >= 0.5.
- The variation of STI-PA test results shall be smaller than 0.03 STI at one test position, thus to fulfill measurement conditions for ambient noise characteristics. The actual variation shall be measured at a representative location.
- Best intelligibility is achieved at message levels in the range of 70-80 dBSPL. At higher sound pressure levels the self-protection of the ear comes into action, which is reflected in a reduced intelligibility index, such as a STI of 1 at 70 dBSPL may be down to 0.7 STI at higher sound pressure level.
- Based on the random STI-PA test signal the typical variations of the measurements is 0.01 – 0.03 STI. Thus at applications with STI-PA values < 0.63 STI the measurement has to be repeated twice and the arithmetical average of all 3 measurements calculated.
- In case variations are higher than 0.03 STI, further 3 measurement shall be carried out and all 6 readings arithmetical averaged.
- In case variations at the same test location are higher than 0.05 STI, causes for these discrepancies shall be detected, eliminated and the measurement repeated.

STI-PA Measurement Hints

- Any background noise has to be sufficiently static during the measurement, e.g. pink noise fulfills this requirement.
- Verify the environmental conditions prior testing. Complete STI-PA measurements without any test signal. The results shall be < 0.20 STI.
- Impulsive background noise during the measurement, such as speech, cause severe measurement errors. The STI-PA result is usually too high.
- In case such an impulsive noise cannot be prevented, the measurements might be shifted e.g. to night time, and afterwards corrected with the averaged daily background noise, using external post processing.
- Any CD-Players used to reproduce the STI-PA test signal have to be accurate as only limited time-shifts (+/- 200 ppm) are allowed to ensure reliable STI-PA test results. Pitch control and shock protection shall be disabled. Thus only professional players shall be used. You may verify the time shift of your CD-Player with a 1 kHz test signal:
- Insert the NTi Audio Test CD into the CD player and start track 1, which is the 1 kHz test signal.
- Connect the XL2 directly to the audio output and measure the signal frequency in the RMS/THD mode. The displayed frequency shall be in the range from 0.9998 kHz to 1.0002 kHz
- STI-PA test signals of other test system manufacturers may sound similar but are not compatible. Only the NTi Audio STI-PA test signal CD V1.1 or higher shall be used in combination with the XL2.
- STI-PA measurement of alarm systems should be carried out at emergency conditions (same sound pressure level and all components are activated).
- At locations with varying conditions e.g. some public areas with few people; others with crowds – the worst case STI-PA results should be measured. Consult your local regulations (e.g. in the U.S., the NFPA code) for specific directives concerning measurement locations and number of required total STI-PA measurements under different circumstances.
- Select typical locations based upon such regulations, or typically position the microphone at 1 - 1.2 meters above ground in sitting areas or 1.5 - 1.8 meters in standing areas (typical measurement positions are normally not directly in front of the speakers).

- The person taking the measurements should be out of the acoustic field, so not affecting the measurement results. For this purpose the measurement microphone can be mounted on a microphone stand and connected with the ASD-Cable to the XL2.
- Low STI-PA readings can be caused by
- Excessive sound reverberation, echoes or reflections
- Poor speaker directivity or speaker coverage
- Speaker power setting not in order (e.g. low signal-to-noise ratio)

STI-PA Post Processing

Measuring the speech intelligibility index under realistic environment conditions is often not applicable, e.g. playing the test signal in a railway station at emergency levels during peak hours will irritate passengers. Additionally at rush hours the characteristics of back ground noise might be highly impulsive. But a pre-requisite for accurate STI-PA measurements is a negligible impulsivity in the background noise.

Under such circumstances the STI-PA measurement should be shifted to a more suitable time of the day, e.g. night time. Such STI-PA measurements taken at untypical background noise conditions have to be post-processed. Post processing combines the STI-PA measurement data taken at quasi noise-free ambient conditions with the unweighted time-averaged octave band noise levels (Leq) taken e.g. during day time, at realistic environmental conditions.

The NTi Audio STI-PA Post Processing Software is tailored for this application e.g. to combine the night and day-time measurement.

"NTi_Audio_STI-PA_PostProcessing.xlt" is available for download at "www.nti-audio.com/XL2"

(Enable all macros at opening the document.)

What to do if Impulsive Noise is permanently present?

In a 24 hours factory or on a highway impulsive noise may be permanently present, thus STI-PA measurements shall not be carried out. In such instances the onsite conditions have to be simulated in a laboratory for STI-PA testing:

- The real noise spectrum shall be measured e.g. with the XL2 Audio and Acoustic Analyzer SLM function, averaging over a sufficient amount of time.
- A diffuse sound field of non-impulsive noise with same frequency shape and octave band levels as measured has to be generated in the laboratory.
- The real speaker - listener situation has to be reproduced acoustically in the laboratory as close as possible.
- Then the actual STI-PA measurement can be carried out. No post processing is required.

This approach will also be mandatory for systems including automated gain control (AGC), if they can't be tested in the original environment due to annoyance of people exposed to the test signal due to the impulsive background noise.

Further Information:

For further information please visit www.nti-audio.com.

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.