vdt 1 – 2018

▲ Picture 1: Sound reinforcement with high intelligibility in a gallery despite adverse acoustical conditions

SPEECH INTELLIGIBILITY IN PA SYSTEMS

Text & Translation Karl M. Slavik and Adrian Slavik

"Sorry, I didn't catch that." If a thought like that is going through a listener's head, chances are that the sound reinforcement is not working as intended – regardless if it´s in a concert hall or at a railway station. The measurement of voice intelligibility does not only provide distinct information about the current state of a PA system, it also gives important clues for optimizing it.

On December 24th, 1915, the world's first large-scale PA system was used to amplify speech and music at a public event in San Francisco. Approximately 100,000 listeners were provided with clear, intelligible speech information. One should think that the last 100 years led to the ultimate perfection of sound reinforcement systems and that poor voice intelligibility is a relic of days long gone, but daily PA practice shows different. Hard to understand speech and a generally nasty sound are still part of the repertoire of some installed or touring sound systems. It´s annoying when you paid full price to see your favourite band or play and the vocals are barely intelligible, but it's clearly dangerous and even illegal if spoken warnings are not comprehensible.

WHAT IS SPEECH INTELLIGIBILITY?

The term Speech Intelligibility describes how much of the clearness and distinctiveness of spoken information is preserved when transmitted over a PA system. Early in the history of audio manually conducted tests were used to rate the speech intelligibility of sound systems: A speaker at a lectern read meaningless words and syllables, which were then written down by the audience as correct as possible. The result was given as a percentage, where 100 percent would have been a perfect score. In the 1940ies Bell Laboratories were developing first electronic measurement methods for speech intelligibility. The ALCons (Articulation Loss of Consonants) refers to the loss of pronunciation of consonants, where the percentage stands for the incorrectly understood words or consonants. An ALCons value of 0.00% would thus mean error-free transmission (no lost consonants).

Modern measurement methods, such as STI (Speech Transmission Index) or STIPA (Speech Transmission Index of Public Address Systems) try to represent the result with a single number, with as many interference factors as possible (ergo the reality) to be included. The result is expressed as a numerical value (STI) between 0.00 (no intelligibility) and 1.00 (perfect speech intelligibility). Good sound reinforcement achieves results between 0.45 and 0.65, even in acoustically unfavourable rooms and 0.70 to 0.90 in acoustically good rooms. For comparison: A good studio transmission chain accomplishes typical STI values between 0.90 and 0.97 (as measured by the author) between microphone in the voice booth and the studio monitor in the control room.

FRIENDS AND ENEMIES OF SPEECH INTELLIGIBILITY

Technically speaking, good speech intelligibility is always given if the modulation depth of the speech or test signal is maintained without a change. The signal is transmitted without aberrations and without masking of the amplitude, in the spectrum and on the time axis.

Good speech intelligibility has many enemies, but also powerful friends. Here is an overview of the most common pitfalls, from which implicitly the solution can be derived.

Room acoustics is one of the most important influencing factors. Long reverberation times (RT60) as well as late reflections (around 80 to 150 ms) act like a „filler“ that clogs the useful modulation depth of the speech signal by covering it, thus making the recognition of information more difficult.

Badly positioned loudspeakers with inappropriate directional characteristics might provide sound for the room, such as to opposite walls and glass fronts, but not directly to the listener. As a result, the directly radiated sound from the speakers at the position of the audience might be proportionally lower than the noise level (reflections, interferences), speech intelligibility decreases.

Undersized public-address systems that are unable to top the noise of a loud audience or drown in the noise of an incoming train are another problem. Again, the signal-to-noise ratio is worsened and speech intelligibility decreased. Sufficiently sized sound systems should be able to provide at least 10 to 15 dB more level than the loudest background noise, with natural limits (hearing protection) and disturbing masking effects to consider.

Psychoacoustic masking effects also contribute to the loss of speech intelligibility. Not only external noise sources can cover the spoken word, but also speech itself, especially at very high voice levels. Low-frequency components in speech can mask quieter, higher-frequency sounds and thus prevent their perception.

Another disturbance are **linear and non-linear** distortions of the transmission path. These disturbances are not only caused by clipping amplifiers and bad speakers, but often by well-intentioned, but badly implemented signal processing. Too much compression of dynamics, overdriven limiters, unnecessary boosts or cuts in the frequency range and also a too stingy data reduction can worsen speech intelligibility.

So beware! Musically accompanied readings in former aircraft hangars, improvisational theatre in emptied swimming pools, platform announcements during the arrival of a train, in all these cases very precise planning and a greatly increased technical effort are needed to achieve reasonable or at least tolerable speech intelligibility.

LEGAL REQUIREMENTS

Whether it‘s concert sound, airport announcements or speech at an auditorium: Speech intelligibility is one of the most important and decisive parameters for the evaluation of all types of public address systems. If regular PA systems are also used as voice alarm systems,

THE AUTHOR

Karl M. Slavik is engaged in the field of professional sound and video technology since the 1980ies and is owner of ARTECAST Audio Video Broadcast since 2005. He works for customers domestically and abroad, including the ARD-ZDF Media Academy, Dolby Labs, ORF, Puls 4, Junger Audio, Special Olympics, Projekt Spielberg (Red Bull Ring), Radio Klassik Stephansdom, SRG SSR and many more.

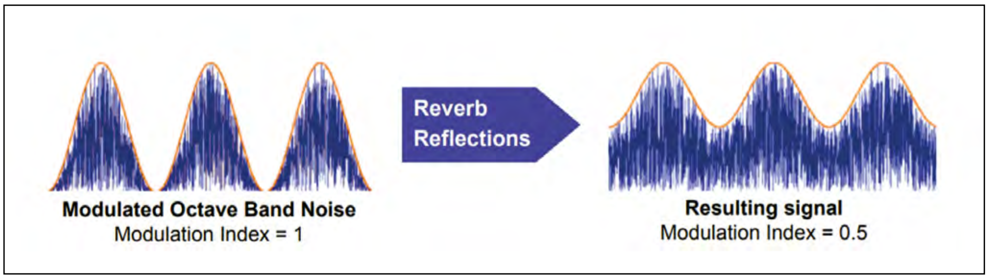
STI-Value	Quality1)	Intelligibility of syllables in %	Intelligibility of words in %	Intelligibility of sentences in %
0 – 0.3	Bad	0 – 34	0 – 67	0 – 89
0.3 – 0.45	Poor	34 – 48	67 – 78	89 – 92
0.45 – 0.6	Fair	48 – 67	78 – 87	92 – 95
0.6 – 0.75	Good	67 – 90	87 – 94	95 – 96
0.75 – 1	Excellent	90 – 96	94 – 96	96 – 100

1) A more detailed classification of speech intelligibility (STI) into 12 categories between A+ (better than 0.76) and U (worse than 0.36) can be found in EN 60268-16 Annex F and Annex G.

◄ Table 1 Meaning of STI-values in practice

© 2018 by Karl M. Slavik

▲ Picture 2: Measurement of speech intelligibility with NTi Audio XL2



Grafic: NTI-Audio

▲ Picture 3: Reverberation and other disturbances reduce the depth of modulation

particularly strict guidelines apply. In Germany, the requirements for voice alarm systems are regulated in DIN VDE 0833-4, in Austria the fire protection guidelines TRVB 123 and 158 regulate that matter. In the UK standard BS5839-8 applies for "VA" (Voice Alarm) or "VACIE" (Voice Alarm Control and Indicating Equipment). For all EU countries EN 50849 applies for „electroacoustic emergency warning systems". The details for speech intelligibility measurement are specified in EN 60268-16.
But when does a public-address system qualify as a „voice alarm system" that is subject to the mentioned standards, especially regarding its speech intelligibility (and other criteria)?
Many permanent installations, for example in theatres, hotels and conference centres, are a combination of high-performance sound reinforcement (for lectures, audio-visual presentations, concerts, parties) and voice alarm systems. In these cases, the specification for „electroacoustic emergency warning systems" according to EN 50849 applies to the entire (!) system. This standard defines the performance requirements for such systems and thus guarantees that availability and operational safety are ensured for the highest demands. In the case of an alarm the running performance must be interrupted automatically and a usually automated announcement with sufficient speech intelligibility should be transmitted. For most systems, a speech intelligibility of STI = 0.5 or 0.55 is mandatory during normal operation, and if system components fail, at least STI = 0.45 must be achieved (see, for example, DIN VDE 0833-4 or TRVB 158).

▼ Picture 4: Acoustics analyser set with STIPA functionality: NTi Audio XL2 analyser with measurement microphone and generator MR-PRO

© 2018 by Karl M. Slavik

At large open-air concerts and sporting events, but also in smaller venues, public address systems can also serve as an emergency voice alarm system. The authority that approves the event can therefore impose similar operational safety and speech intelligibility requirements as with fixed installations. Even without special regulatory requirements, it is urgently recommended to plan systems according to the current state of the art and to ensure maximum speech intelligibility.
For example, in the event of injuries following a mass panic or due to bad weather, it is generally checked by the authorities whether announcements by the organizers were well audible and effectively used for traffic guidance (general diligence and operational safety obligations).

MEASUREMENT OF SPEECH INTELLIGIBILITY

Although there are several methods of measuring speech intelligibility today, STIPA has become a fast and accurate method for many applications. The measurement of the „Speech Transmission Index for Public Address Systems" has been further developed by NTi Audio and is now standardized according to EN 60268-16. Several T&M manufacturers are offering this method. STIPA is based on the measurement of the MTF (Modulation Transfer Function), which measures the loss of modulation depth in different octave bands. As a measurement signal, specially modulated noise in the octave bands of 125, 250, 500, 1,000, 2,000, 4,000 and 8,000 Hz is used, which assesses the transmission quality in the baseband as well as the harmonics of the speech (formants) which is particularly important for intelligibility. In each octave band, the noise is modulated with two frequencies, resulting in an amplitude modulated noise signal. Spectrum and modulation correspond to the time-averaged signal of a male speaker.

The measurement of speech intelligibility can be realized in two ways: either by means of a signal generator (for example NTi Audio MR-PRO), which feeds the STIPA signal directly into the sound system under test, or by means of a „TalkBox" with an already built-in signal generator. This special loudspeaker serves as a „speaker replacement" and is positioned in front of the microphone of the sound system. As an analyser the NTi Audio XL2 with STIPA software and measuring microphone is used. The measurement itself takes about 15 seconds, a correction of the determined speech intelligibility with the spectrum of ambient noise is possible. During the measurement no impulsive noises or other short or speech-like sounds should occur. An automatic averaging process calculates the average and the statistical deviation of several measurements. The result is displayed as a speech intelligibility index in STI (Speech Transmission Index) or CIS (Common Intelligibility Scale, CIS = 1 + log (STI)).

▶ Table 2: Composition of the STIPA test signal

	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							✓							✓

References

- The Measurement of Speech Intelligibility. Herman J.M. Steeneken, TNO Human Factors, Soesterberg, the Netherlands.
- Manual XL2, NTi Audio, 2018 https://www.nti-audio.com/Portals/0/data/en/XL2-Manual.pdf
- Beschallungsanlagen in Räumen Teil 3 (Messtechnik und STI). Seminarunterlagen (PDF) von Anselm Goertz und Alfred Schmitz. DEGA Akademie Raumakustik und Beschallung, 2017 (German only)
- EN 60268-16, Sound system equipment - Part 16: Objective rating of speech intelligibility by speech transmission index (IEC 60268-16:2011)
- NTi Audio Application Note „Speech intelligibility", 2017