# The application of the STI in speech communication

Wolfgang Probst, Michael Böhm, Datakustik GmbH, Dornierstr. 4, D-82205 Gilching
January 2017

## *Summary*

The intelligibility of speech is an important criterium in the assessment of the acoustic quality of work places or other whereabouts. While in some areas with people working together good speech intelligibility is the target, it would be disturbing in other cases where the unwished understanding of spoken information is disturbing and may even hurt the confidence of exchanged information by others. The first step in the assessment of such speech-induced effects and quasi the "atom" of an assessment of scenarios of any complexity in working environments is the quantification of speech intelligibility by calculating or measuring the Speech Transmission Index STI for a defined position of source and receiver. It includes all the major effects and influences like room response, background noise, the limitation of the dynamic range by the absolute speech reception threshold level with low signal levels and the auditory masking with high signal levels. As this assessment method that can easily be applied with existing technique is often not applied in favour of the limitation of reverberation time, the basic concept is summarized and demonstrated with simple examples in the following.

## *Introduction*

The human language is a very special type of noise in order to ensure an optimal acoustic environment in workplaces. While the task with regard to the noise of machines and other noise-relevant technical devices is a minimization of the noise levels to avoid the disturbance of the concentration or even the danger hearing loss with speech other effects like good intelligibility or - on the contrary - bad intelligibility for reasons of privacy may be important. The understanding of what is said is, however, not only dependent on the proportional sound pressure level, but is determined by a series of further influences which lead to a reduction in the modulation depth which is important for understanding. These include, for example, the level of a background noise which is always present, the distance of the signal level from the hearing threshold, the masking of the modulation in a frequency band by the level of the adjacent lower frequency band, and especially the "smearing" of the modulation by the time delay due to different propagation paths of a sound signal inside a room.
All these aspects are compacted with the Speech Transmission Index STI according to IEC 60268-16 / 1 / to a characteristic value which, despite its complex "design", is extremely easy to understand and to use. The STI describes the quality of the transmission of voice information from the location of the voice source - for example, a speaker or a loudspeaker - to a fixed receiver position. It quantifies with its value

$$0 \leq \mathrm{STI} \leq 1$$

this quality and taking into account the sound power level of the speech it allows to assess the intelligibility at the position of a listener. For a STI < 0.2, less than 30%, for STI > 0.5, more than 80% of the spoken is understood - which of course are roughly average values.

## *The "design-concept" of the STI and the implementation in the acoustic simulation*

For the calculation of the STI, a speech signal emitted by the source is described by its A-weighted sound power level according to Table 1 and by a standardized frequency spectrum according to Table 2

Table 1 - A-weighted sound power levels of speaking people as a function of speech effort

| Speech effort | $L_{WA}$ dB |
|---|---|
| whispering | 47 |
| quietly | 53 |
| unstressed | 59 |
| normal, relaxed | 65 |
| normal, raised | 71 |
| raised | 77 |
| loud | 83 |
| very loud | 89 |
| shouting | 95 |

Table 2 - Linear band levels of speech normalized to an A-weighted overall level of 0 dB (after / 1 /)

| frequency Hz | 125 | 250 | 500 | 1000 | 2000 | 4000 | 8000 |
|---|---|---|---|---|---|---|---|
| Δ dB (male) | 2,9 | 2,9 | -0,8 | -6,8 | -12,8 | -18,8 | -24,8 |
| Δ dB (female) | - | 5,3 | -1,9 | -9,1 | -15,8 | -16,7 | -18,0 |

For the determination of the sound power level frequency spectrum, the value of Table 1, which is applicable for the voice effort, is added in all 7 columns of the corresponding line in Table 2.
According to Table 2, the sound carrying the speech information comprises the octave frequency bands from 125 Hz (k = 1) to 8000 Hz (k = 7). However, by the 7 level values only the average sound intensity in each frequency band $f_k$ is described. A temporally constant sound signal with this level would be perceived as noise and not transport any information. Only through the articulation by the human speech apparatus impressing a temporal fluctuation or a modulation on the sound signal leads to the perception of syllables and sentences and to an understanding of the information transmitted thereby. This variation or modulation takes place in a temporal sequence, which can be described for the evaluation of the quality of speech transmission by the 14 1/3-octave-frequencies $f_m$ from 0.63 Hz to 12.5 Hz.

This is illustrated by the principle diagram in Fig.1. The speech sound-signal (B) emitted by the speaker (A) and fluctuating in time depending on the modulation depth propagates and, due to different influences, has an often smaller fluctuation or a lower modulation depth arriving at the listener.
This effect can be described by separately analyzing this reduction in the modulation depth for each of the 7 frequency bands from Table 2. For this purpose, a modulation of the frequency $f_m$ - for example 1 Hz - is imparted to the sound signal in the octave band at the relevant frequency $f_k$ - for example, 1000 Hz - as is shown in line C and expressed mathematically in line D. The modulation depth of the sound with intensity $I_S$ emitted by the speaker is in the formalism of line D expressed by the amplitude or the modulation index $m_s$.

This sound signal impacts at the receiver with an intensity $I_R$ and with a modulation index $m_R$, which is generally reduced. The reduction of the modulation depth from the speaker position to the receiver position is then calculated for this frequency band $f_k$ (in the example 1000 Hz) and at this modulation frequency $f_m$ (in the example 1 Hz) by the quotient of the two modulation indices, the so-called modulation transfer factor MTF.
This mathematically not further developed concept shows the essential core of the assessment of the speech transmission by the reduction of the modulation depth. The numerous above-mentioned phenomena influencing the speech transmission, such as the reverberation of a room or additional background noise, are assessed how they reduce the 98 modulation transfer factors (7 octave frequency bands with 14 modulation frequencies each).

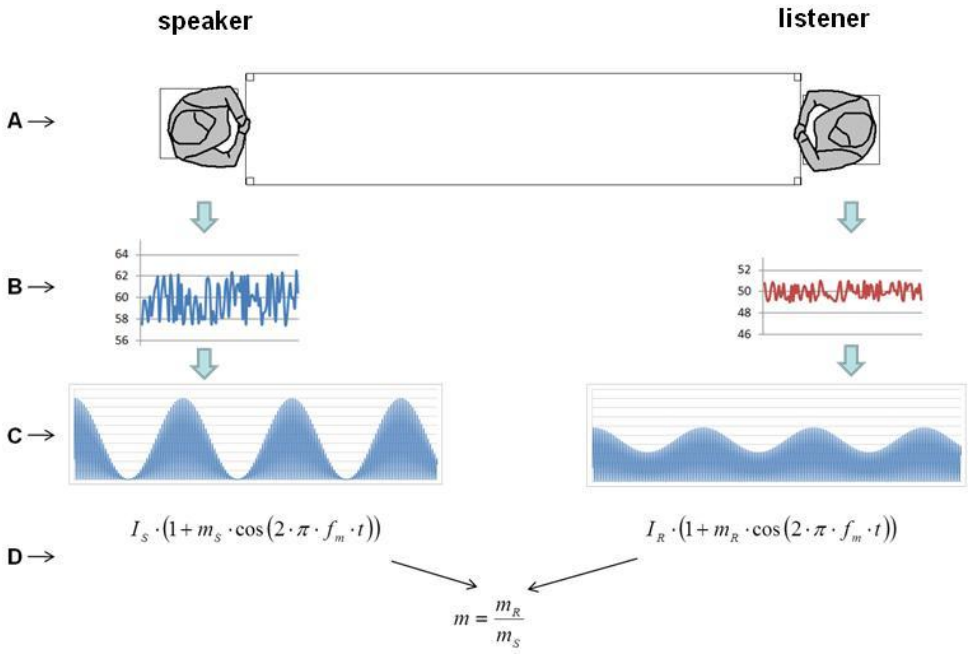


Figure 1 - Schematic presentation of the modulation transfer index m

In the following, a somewhat simplified illustration will be given of how the reverberation - or better the energetic impulse response - of a room affects the deterioration of the speech intelligibility by the reduction of the modulation depth and how this can be predicted with simulation calculations.

In the simulation, sound particles or rays are emitted in all directions from the source, as it is shown in figure 2 on the left, and their paths are calculated. The sound energy associated with a particle is weakened during each reflection, corresponding to the degree of absorption of the reflecting surface. The sound pressure level is calculated from the number of particles passing through a small counting volume centered around the receiver summing up the sound energy transported by them.
Since the length of each individual path of a particle is known up to the arrival at the counting volume, the time elapsed since the emission can also be determined taking into account the sound velocity. In this way, the sound pulses emitted by a source and arriving at the receiver can be subdivided into time classes and displayed as energy-related impulse responses according to the diagram on the right in figure 2.

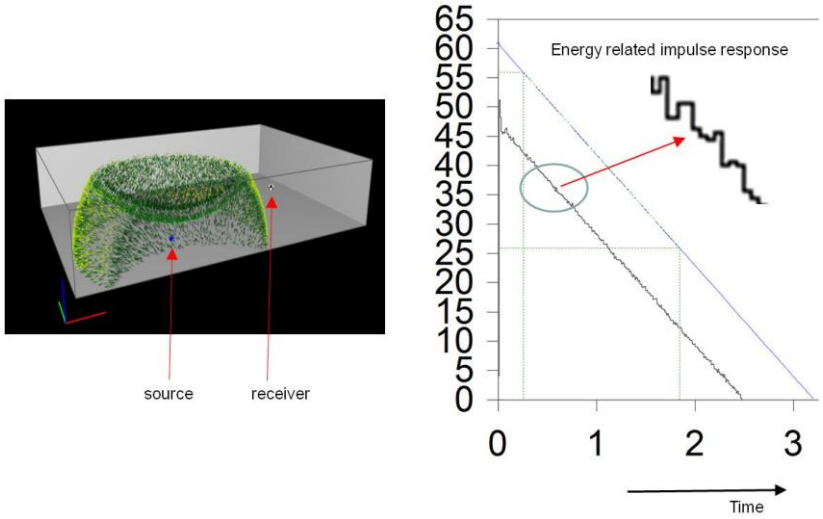


Figure 2 - Emission of sound particles (left) during simulation and temporal classification for determining the impulse response (right)

For the further understanding, it is essential that the zero point of the time axis on the right in figure 2 refers to the time of the emission of the sound particles irrespective of the time sequence of the individual paths calculated. The above decay curve, which is parallel to the impulse response, is produced by backward integration or by summation of the sound energy from the relevant point in time to the end of the time scale. This curve corresponds to the level decay after switching off a stationary sound source.

These contexts also apply to speech sounds - with the picture in figure 3, an attempt was made to present this in a somewhat simplistic manner. In the upper right corner of the picture are three ray paths from the telephone speaker to the - here involuntarily - listener represented. A syllable spoken at a given time comes as a direct sound, according to the impulse response, with a level of 68 dB, but also with 65 dB and a shift of 0.2 seconds as well as with 62 dB and a shift by 0.4 seconds on the listener (all values are exemplary and are for explanation only).

Although the sum of these three components has a higher level of 70.5 dB, the time-delayed superposition leads to a reduction in the modulation depth and thus to a poorer intelligibility. This also corresponds to the experience - in the hally room the perceived volume increases, but the intelligibility decreases because of the temporal "smearing" of the speech signal.

In the calculation of the above-mentioned modulation-transfer index for the octave band $f_k$ with the modulation frequency $f_m$, the function $F_s$ (picture 3 lower left) representing the sound signal is gradually shifted, weakened and reduced according to the temporal decay of the impulse response determined by measurement or simulation calculation (This corresponds mathematically to the convolution of function $F_s$ and impulse response). The corresponding modulation-transfer index m can then be determined with the remaining modulation index $m_R$.

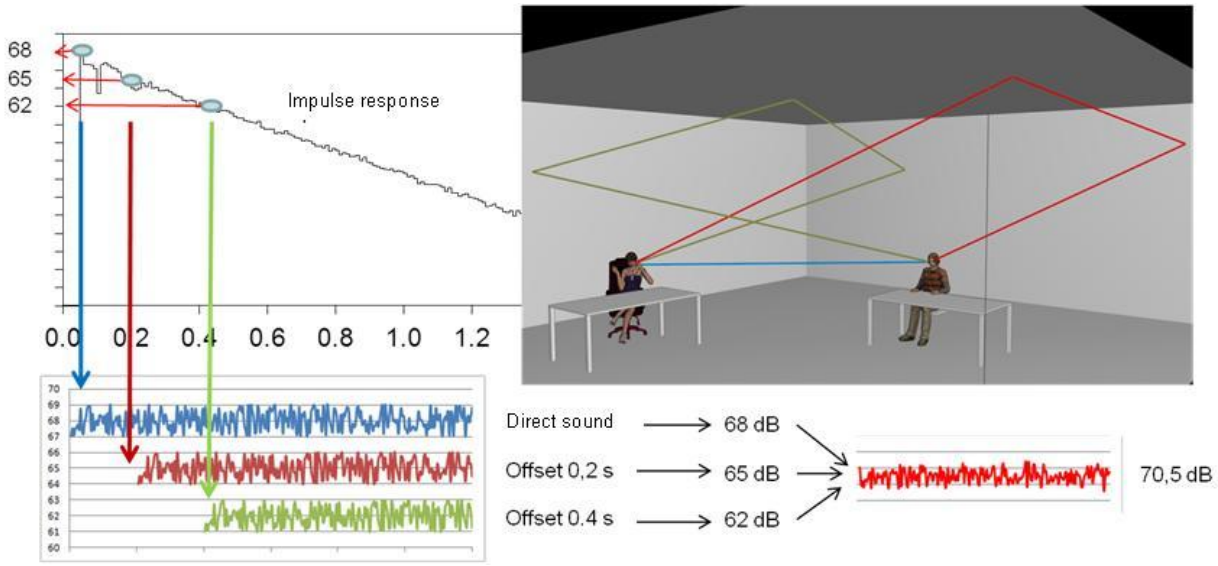


Figure 3 - Spread of speech sound (top right), impulse response (upper left), resulting time-shifted speech signals (lower left) and final total sound with reduced modulation

From the 98 modulation transfer indices, the ultimately interesting Speech Transmission Index STI is then calculated in several steps.

## *The influence of impulse response and reverberation time on the STI*

As an example of the influence of reverberation on the STI, this is calculated for the case where speakers and listeners are in the model of a reverberation chamber described in / 2 /. Figure 4 shows this model with the spreading particle cloud, and in Figure 5 the diagram of the decay curves determined from the energy-related impulse response (which is also referred to as an echogram) is shown here.

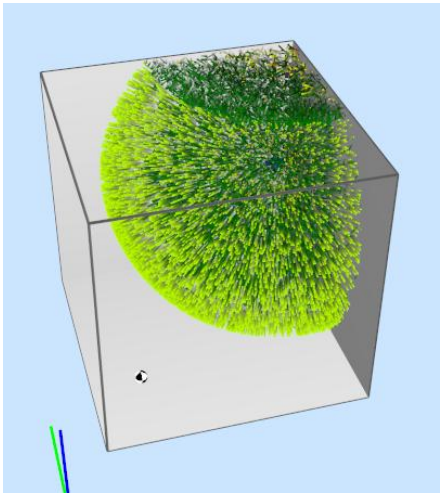

Figure 4 - Reverberation chamber with expanding cloud of sound particles

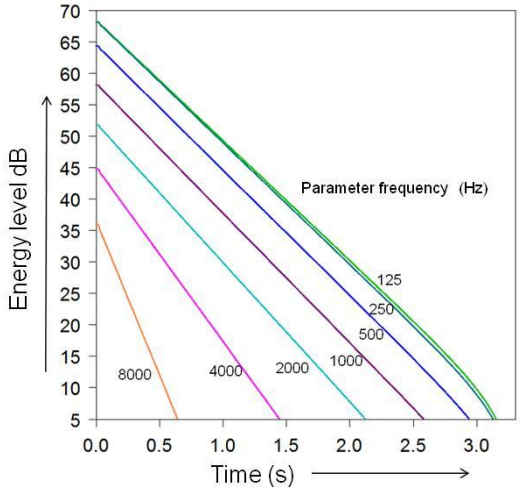


Figure 5 - Sound decay determined by simulation

In the diffuse sound field, the decay curves plotted with a logarithmic ordinate scale show a linear profile as shown in Figure 5, and each of these curves can be clearly described by a reverberation time. In this special case, the above calculation method leads to an analytically closed solution

$$m(f_m) = \frac{1}{\sqrt{1+\left(\frac{2\pi f_m T}{13{,}8}\right)^2}} \qquad (1)$$

mit

$f_m$ Modulation frequency in Hz
m modulation transfer index
T Reverberation time in s
$m(f_m)$ modulation transfer index at the modulation frequency $f_m$ and for the octave band center frequency to which the reverberation time T is related.

Figure 6 shows the reverberation times determined by simulation using the SERT method (see / 2 /), the 98 MTF indices calculated with (1) and finally the value of the STI calculated therefrom.

| $f_k$ (Hz) | 125 | 250 | 500 | 1000 | 2000 | 4000 | 8000 |
|---|---|---|---|---|---|---|---|
| $T_{60}$ (s) | 3.16 | 3.11 | 3.03 | 2.92 | 2.73 | 2.18 | 1.22 |

| fm | Octaveband-frequency (Hz) | | | | | | |
|---|---|---|---|---|---|---|---|
| (Hz) | 125 | 250 | 500 | 1000 | 2000 | 4000 | 8000 |
| **0.63** | 0.74 | 0.75 | 0.75 | 0.77 | 0.79 | 0.85 | 0.94 |
| **0.8** | 0.66 | 0.66 | 0.67 | 0.68 | 0.71 | 0.78 | 0.91 |
| **1** | 0.57 | 0.58 | 0.59 | 0.60 | 0.63 | 0.71 | 0.87 |
| **1.25** | 0.49 | 0.49 | 0.50 | 0.52 | 0.54 | 0.63 | 0.82 |
| **1.6** | 0.40 | 0.40 | 0.41 | 0.43 | 0.45 | 0.53 | 0.75 |
| **2** | 0.33 | 0.33 | 0.34 | 0.35 | 0.37 | 0.45 | 0.67 |
| **2.5** | 0.27 | 0.27 | 0.28 | 0.29 | 0.31 | 0.37 | 0.58 |
| **3.15** | 0.22 | 0.22 | 0.22 | 0.23 | 0.25 | 0.30 | 0.50 |
| **4** | 0.17 | 0.17 | 0.18 | 0.18 | 0.20 | 0.24 | 0.41 |
| **5** | 0.14 | 0.14 | 0.14 | 0.15 | 0.16 | 0.20 | 0.34 |
| **6.30** | 0.11 | 0.11 | 0.11 | 0.12 | 0.13 | 0.16 | 0.27 |
| **8** | 0.09 | 0.09 | 0.09 | 0.09 | 0.10 | 0.12 | 0.22 |
| **10** | 0.07 | 0.07 | 0.07 | 0.08 | 0.08 | 0.10 | 0.18 |
| **12.5** | 0.06 | 0.06 | 0.06 | 0.06 | 0.06 | 0.08 | 0.14 |

| STI | 0.41 |
|---|---|

Figure 6 - The determined reverberation times, modulation-transfer indices and the STI determined from these

This example is also suitable as a test case in the frame of quality assurance - with the echograms determined by the simulation the same value must be obtained.

Background noise provides an uncorrelated contribution to the speech signal - it also leads to a reduction in the modulation depth and thus the STI. Decisive is not the absolute sound pressure level, but the S / N value (signal to noise ratio in dB) or the difference between the levels of the speech signal and background noise. The calculation with variation of the S / N and the reverberation time T60 leads to the dependency shown in the diagrams figures 7 and 8.

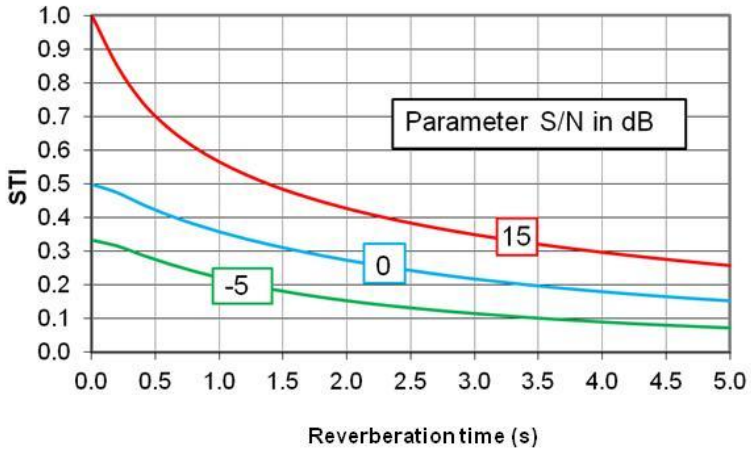


Figure 7 - The STI as a function of the reverberation time for 3 values of S / N in dB

The curve with an S / N of 15 dB - virtually no external noise - shows that even in this "ideal" case, reverberation times are required under approx. 1.5 s, in order to achieve an acceptable speech intelligibility.

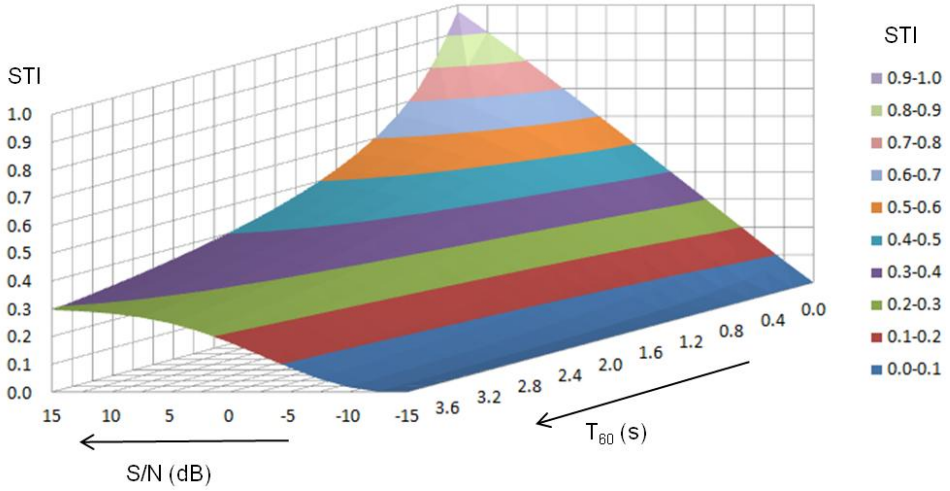


Figure 8 - The STI as a function of the S / N in dB and the reverberation time (only with diffuse sound field)

Bei der längsten in Bild 8 einbezogenen Nachhallzeit von 3,6 s kann, wie die entsprechende vordere Randkurve der STI-Fläche zeigt, auch bei fehlendem Hintergrundlärm mit S/N ≥ 15 dB kein höherer STI als 0,3 erreicht werden.
Alle diese Zusammenhänge beziehen sich darauf, dass der Sprach-Signalpegel beim Hörer mindestens 40 dB(A) und höchstens 80 dB(A) beträgt. Unterhalb von 40 dB(A) führt die Hörschwelle und oberhalb von 80 dB(A) der Maskierungseffekt zu einer Verringerung der Sprachverständlichkeit. Dies zeigt das Diagramm Bild 9, mit dem der STI bei den 3 Werten des Signalabstands in Abhängigkeit von dem beim Hörer vorhandenen Signalpegel in dB(A) dargestellt ist.
With the longest reverberation time of 3.6 s included in Fig. 8, the STI cannot exceed 0.3, even if background noise is missing due to a S/N ≥ 15 dB.
All these contexts refer to the fact that the speech signal level is at least 40 dB (A) and at most 80 dB (A) at the listener. Below 40 dB (A) the hearing threshold and above 80 dB (A) the masking effect leads to a reduction in the speech intelligibility. This is illustrated by the diagram figure 9 with which the STI is shown by 3 curves related to different S/N ratios as a function of the signal level in dB (A) at the receiver position.

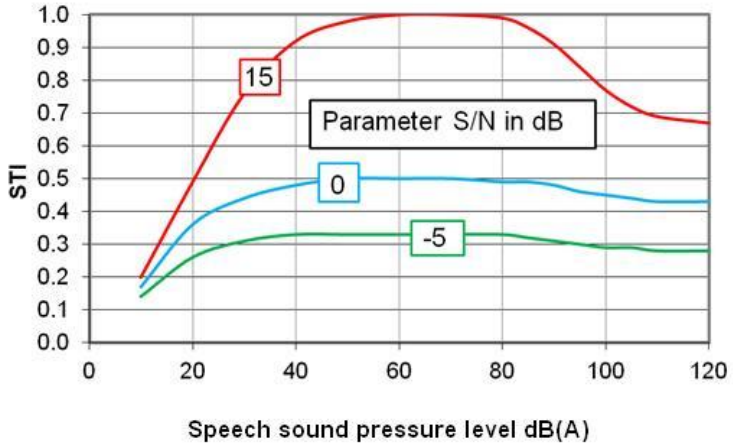


Figure 9 - STI dependence on the A-weighted signal level.

However, one should not overestimate the described dependency of the speech intelligibility from the reverberation time - it only applies if the listener is in a distance from the speaker where a diffuse sound field exists that can be described by a reverberation time.

This is shown by a simple calculation for the empty hall, shown in Fig. 10, with the dimensions 50 m x25 m x10 m - it is assumed to be very reflective with an average absorption coefficient of 0.1 at all surfaces. Due to a reverberation time of 4 s - 5 s, an STI of less than 0.3 would be expected without background noise.

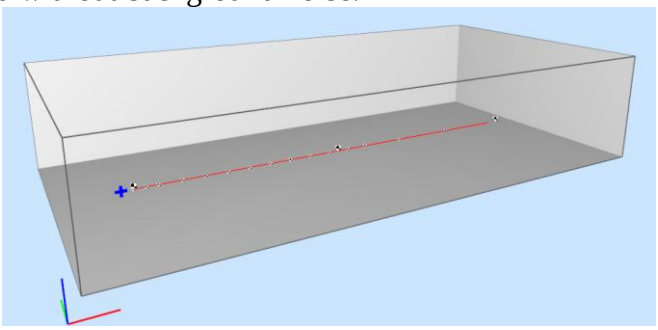

Fig. 10 - Empty hall 50 m x 25 m x 10 m with source and measuring path

In fact, from the echograms at the receivers along the straight path, the STI-values shown in the diagram figure 11 are calculated with simulation.

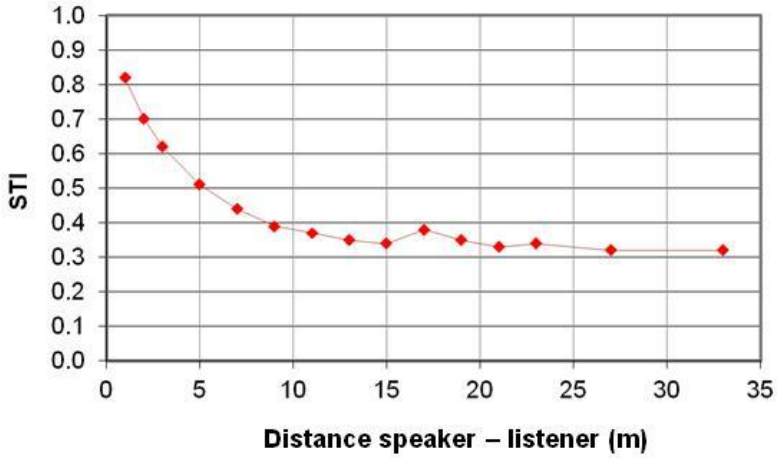


Figure 11 - The STI as a function of the distance of the listener from the speaker in the large hall

The STI of 0.5 to 0.8 near the speaker shows good speech intelligibility despite a high reverberation time of the room of more than 4 seconds. This is in line with experience - even in large and reverberant rooms like churches or the big aulas in some hotels or business buildings, there is no problem with the speech intelligibility at the usual speaker distance if the background level is low. The reason is shown by the echogram at a distance of 1 m from the loudspeaker - the level drops steeply by almost 30 dB, and then the reflection-induced decay process begins.

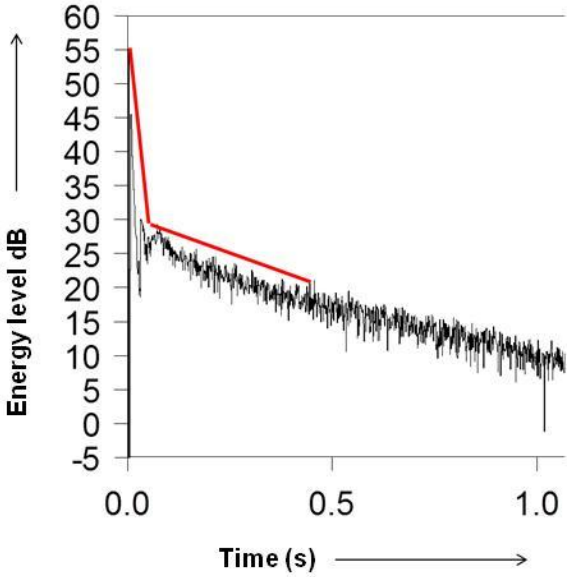


Figure 12 - Echogram at 1 m distance from the speaker (500 Hz band)

This is the great advantage of using the STI computed from the echogram to go far beyond the assessment of reverberation times, because it encompasses many phenomena affecting speech intelligibility and thus provides a much more reliable prognosis of speech-related effects.

## *The STI in environments with screening facilities*

Diffraction effects can often be neglected in the calculation of sound levels in rooms, because sound energy reaches the areas which are screened from the direct sound by many other unscreened propagation paths. With the method SERT / 2 / the sound particles passing near an edge are deflected in such a way that the diffraction field in accordance with the experimentally developed formalism by Maekawa / 3 / and also applied in ISO 9613-2 / 4 / is reproduced approximately.
According to Maekawa the sound pressure level caused by a sound source S at a receiver R in the case of free field propagation is reduced by the insertion of a wide screen by a frequency-dependent barrier attenuation $D_z$. If the length of the detour across the screen edge is greater than the length of the direct path - without a screen - between the source S and receiver R, the barrier attenuation is given by

$$D_z = 10 \cdot log(3 + 20 \cdot N) \ dB \qquad (2)$$

with

$$N = \frac{2}{\lambda} \cdot z$$

λ is the wavelength.

Mit dem einfachen Modellraum Bild 13 kann diese Beugungsrechnung geprüft werden - die Raumwände mit ihrem Absorptionsgrad von 1 ergeben eine reine Freifeldausbreitung und mit dem Schirm sollte sich an den Berechnungspunkten ein Schalldruckpegel entsprechend (2) ergeben. Bild 14 zeigt die Ablenkung der Teilchenbahnen in den abgeschirmten Bereich.

The diffraction calculation can be tested with the simple model room shown in figure 13 - the room walls with their absorption coefficient of 1 yield a pure free-field propagation and a sound pressure level corresponding to (2) should result at the calculation points behind the screen. Figure 14 shows the deflection of the particle paths into the shielded region.

M

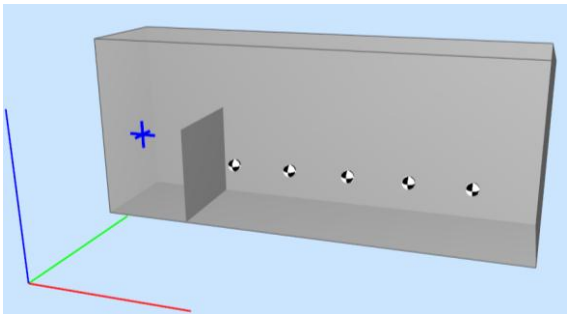

Figure 13 - Screen with source and receivers

M

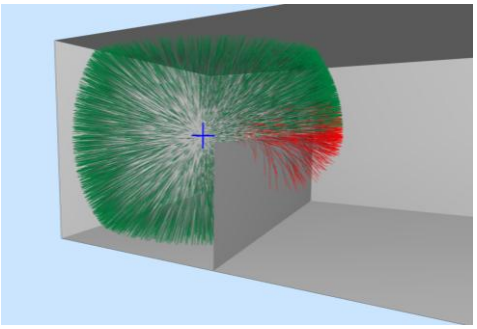

Figure 14 - Deflection of the sound particles to produce the diffraction field

The barrier attenuations $D_z$ were calculated from the sound pressure levels determined for the frequency bands 125 Hz to 8000 Hz by simulation with and without screen and are shown as dots in the diagram figure 15 with the curve calculated from equation (2).

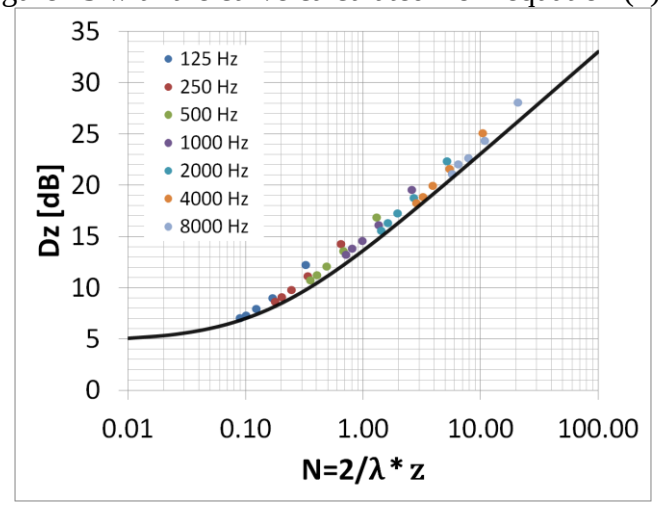


Figure 15 - The barrier attenuations determined for 7 frequency bands and 5 receiver positions in comparison to the Meakawa-curve according to (2)

Thus, even in cases where no other sources are relevant and where the sound level behind a screen is only determined by the diffracted sound, these sound levels and the STI values can be determined. This is of particular interest if partial screens, file cabinets or other screening devices provided between workstations in order to prevent disturbance by telephone calls or other speech communication.

The assessment of the effect of shields on the intelligibility of language may be demonstrated with the model figure 16. The room with the dimensions 20 m x 10 m x 6 m is divided by a 2.5 m high partition into a 6 m wide subspace with reflecting bounding surfaces and a 14 m wide subspace with highly absorbing boundary surfaces. In the first step, the sound pressure levels and the STI values without screen are calculated with a source spectrum according to the normally raised voice according to Tables 1 and 2 - as the results in Fig. 17 show, the speaker may be well understood throughout the room.

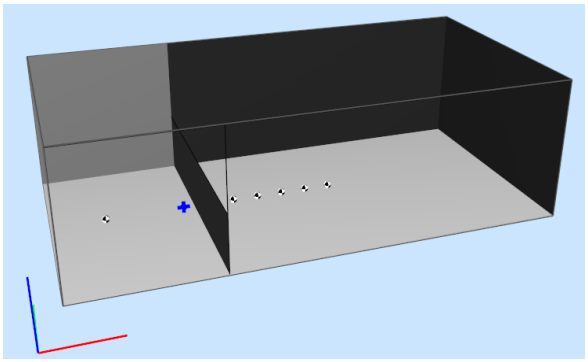

Figure 16 - Room with screen, source and receiver positions

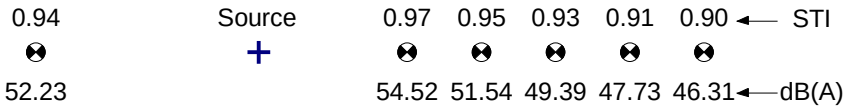


Bild 17 - Sound pressure values in dB(A) and values of the STI without screen

The result of the calculation with a screen are presented in figure 18.

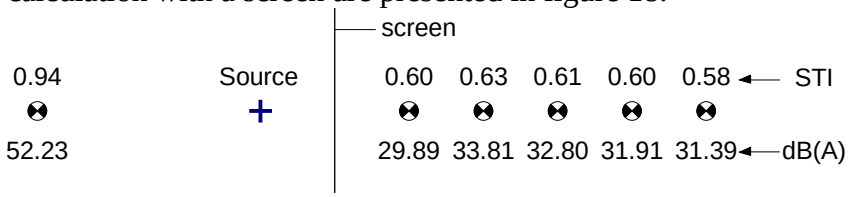


Figure 18 - Sound pressure levels calculated with particle diffraction and STI values with screen

The STI, which is clearly reduced with the screen, is the result of the fact that the levels are already so low that the hearing threshold contributes to the limitation of the remaining modulation depth. This can be shown by repeating this calculation with a sound power level of 100 dB (A) instead of 70 dB (A) as above. The result shows that the STI increases despite the same calculation process because the hearing threshold is not relevant with higher levels at the receivers.

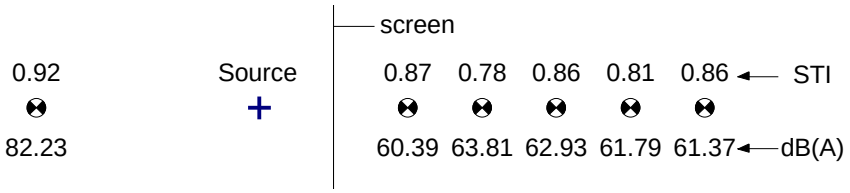


Figure 19 - With screen and with a sound power level of the source increased to 100 dB (A)

Finally, the calculation is repeated with a sound power level of 120 dB (A) - only possible with a loudspeaker as sound source.

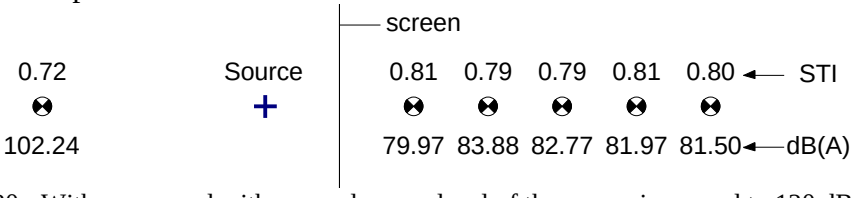


Figure 20 - With screen and with a sound power level of the source increased to 120 dB (A)

As the comparison of the results in Fig. 19 and Fig. 20 shows, the STI values in the shielded area remain essentially the same even with extremely high source sound power with an $L_{WA}$ of 120 dB - for the decrease of the modulation depth from the source to the receiver it plays no role whether the emitted and the received speech signal is increased or decreased by the

same level. However, similar to the "dipping" into the hearing threshold, an effect reducing the STI also occurs at very high levels upwards, as shown by the comparison of the STI values at the receiver position, which is not shielded from the source, at the far left. With a source sound power level of 120 dB (A), the masking effect of each frequency band for the adjacent higher frequency band leads to a deterioration of the perceived modulation depth and thus to a corresponding reduction of the STI. This is also plausible - a loudspeaker in a distance of 3 m, which is driven up to the sound power of 120 dB (A), leads to a lesser intelligibility.

Finally let us take the case of figure 18 with normal raised speech effort but with a background noise level typical for more silent working areas of 30 dB(A) - the result is shown in figure 21. This even low background level of 30 dB(A) leads to a reduction of the STI to values of 0.3 to 0.4.

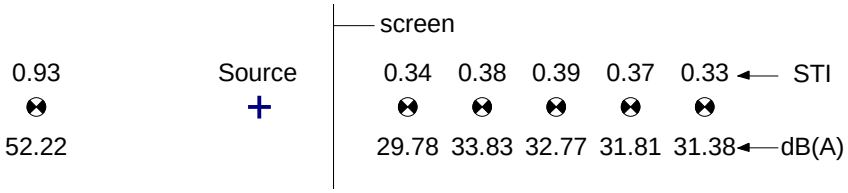


Figure 21 - With screen and normal raised speech effort, but an additional background noise of 30 dB(A)

And this is the essential "message" of the STI concept: shielding reduces the intensity of the speech signal, but this only reduces the absolute level and not the modulation depth. When a speaker is shielded from the listener in an absolutely quiet environment, the listener will only be misunderstood by the influence of the individual hearing threshold. In practice, however, not the hearing threshold, but the always present background noise forms the lower borderline.
This would also be the objective of an improved strategy - to assess the speech intelligibility between individual workplaces or workplace groups, it is necessary to include the typical background noise by classifying typical uses or by individual calculation of this background level. This would significantly improve the acoustic optimization of workplaces compared to the currently applied rigid upper and lower limits of the reverberation time.

## *Planning principles and examples*

The calculation of sound pressure levels and STI-values in the planning phase enables a targeted assessment and optimization of rooms in which communication by language plays an essential role.
Everyone knows this problem of restaurants, which one may only avoid because it is well known from experience that they are very loud if many people are there and a relaxed communication may not be possible. This is true for all room sizes from the small bistro with a narrow seating area for space reasons, to a large, light-flooded restaurant with the famous acoustic charm of station shacks.
Causes of the "acoustic pollution" are the conversations of the other guests - every person who speaks is a source of sound which must be taken into account in the planning calculation by an emission according to Table 1.
The loudness of a person depends on the level of the background noise to be drowned and the maximum distance of the people with whom you are communicating. In this sense, the beer and wine cellars with long tables in historical vaults form the negative end of the rating scale.

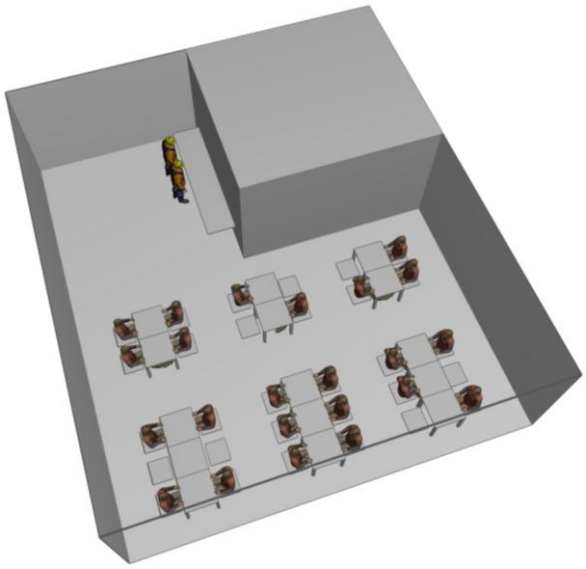

Figure 22 - Small restaurant with food bar

The principle of a planning process is explained with the model of a small restaurant with integrated food issue in figure 22. The starting point is the creation of a model of this kind, which, however, has to be done in a time-saving manner using modern software tools. The goal is that the conversation partners at their own table can be understood even when the assignment is full. In this case, the STI is a decisive parameter as described below.
First, a speaker position and a receiver position are fixed at each table in a typical position. Their distance is greater the longer or larger the tables are.
Then the sound pressure level defining the background level at the receiver position of a table is calculated by simulation with the speaker position at all other tables as sound sources.
Figure 23 shows this first step with the receiver position at the blue marked table with the speakers as sources of uncorrelated background noise marked in red. The sound pressure level calculated is taken into account as background level that is used to calculate the STI for the conversation at this table. (Further constantly existing external noise caused by air-conditioning systems or other external noise is taken into account, if necessary by energetic addition to increase the background noise level).
The STI characteristic for the conversation at this table is then calculated by simulation taking into account the background noise level determined in the first step and with speaker and most distant listener at this same table.

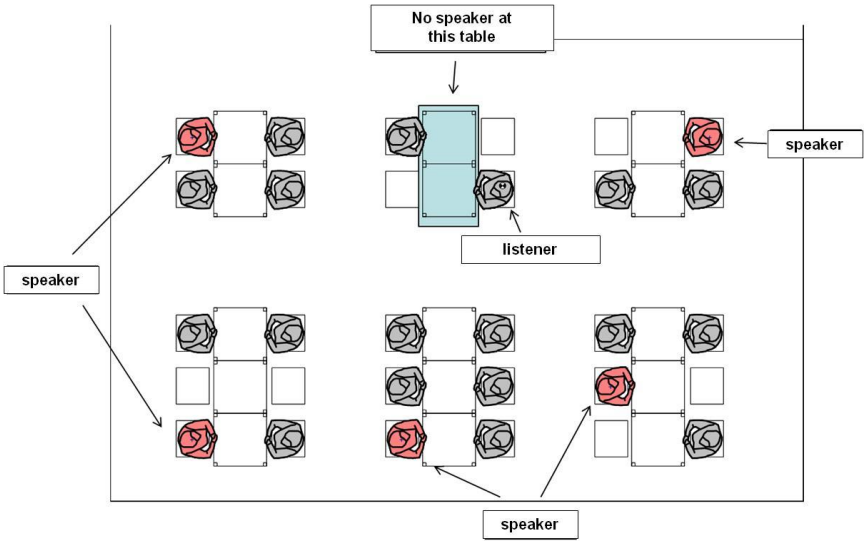


Figure 23 – Determination of the background level at the listener position at the table marked in blue with speakers as sources marked in red

This process is repeated for all tables, if possible automatically. Exemplary results for the characteristic STI values for the individual tables are shown in Fig. 24 for this example. The STI values are clearly less than 0.5 and indicate poor speech intelligibility.

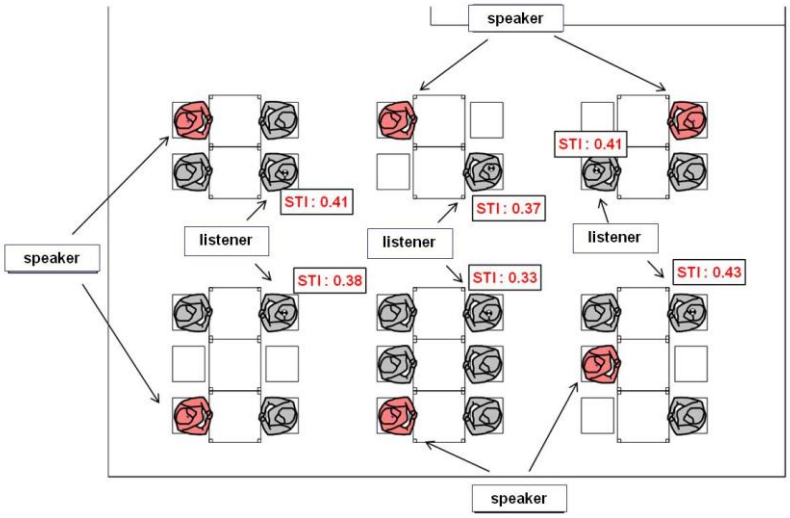


Figure 24 – STI-Values calculated for all tables

This is the starting point of a well known effect that everybody speaks with more effort to increase intelligibility in spite of the background noise. But a new calculation with this increased speech effort increases speech signal and background by the same amount of dB - the STI values will be the same.

The right solution is an acoustic planning with the well known toolbox of increased absorption, screens and possibly an improved layout. In the simple example, a certain acoustic separation is produced by absorbing screens in connection with the absorbing lining of two mutually perpendicular walls and the suspension of an absorbing baffle-system.

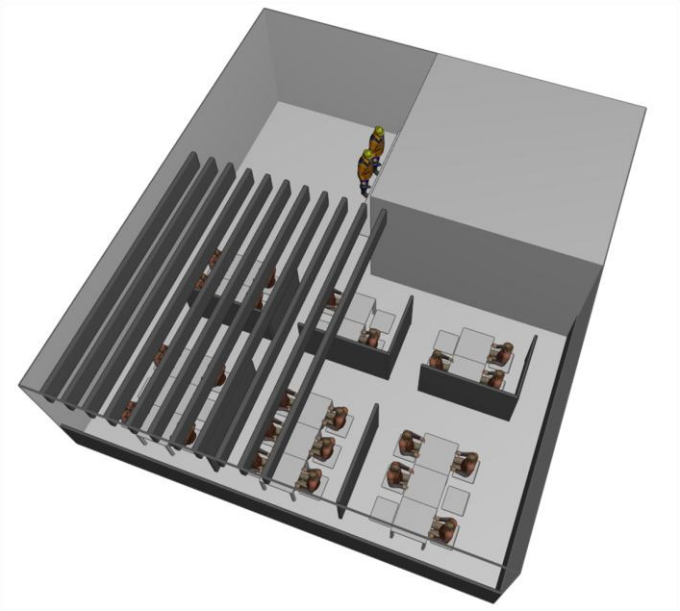

Figure 25 – Acoustic improvement by separating partial screens and an absorbing baffle system

The new simulation calculation according to the described procedure leads to the values of the STI shown in figure 26, thus indicating a substantial improvement in the acoustic atmosphere perceived by the guests.

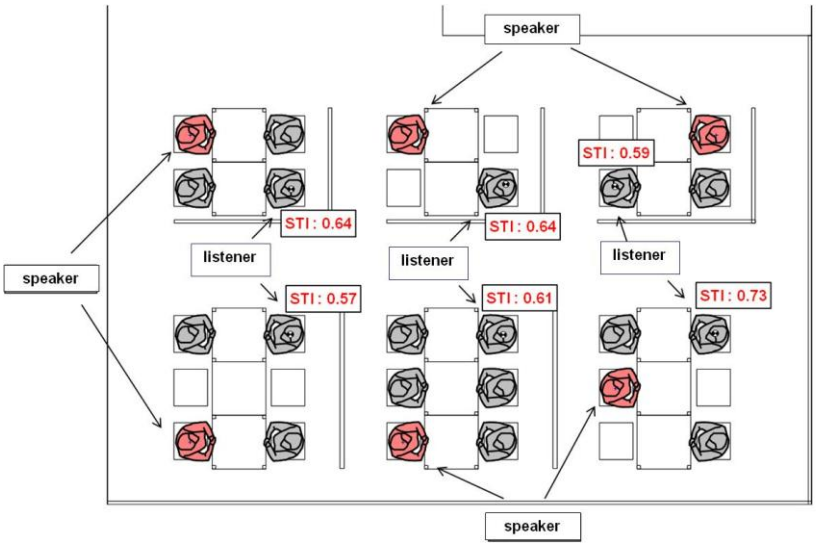


Figure 26 - STI values determined by simulation taking into account the planned measures

## *Conclusion*

The intelligibility of speech plays an important role in the assessment of the acoustic quality of lounges. From the disturbance of the ability to concentrate through involuntary monitoring of the communication of others in the multi-person office to the impossibility of a relaxed conversation with the own table neighbors in the restaurant, there are numerous aspects that can be "defused" by a predictive acoustic planning based on a quantification of the speech intelligibility by the STI. Due to its design, this standardized characteristic value takes into account the most important parameters, which can be influenced by acoustic planning in the sense of the desired optimization. The STI takes into account the energetic impulse response determined by a simulation calculation and thus - in contrast to the reverberation time - applies to any room shapes and equipment. The most important dependencies are covered by the inclusion of the inevitable background noise as well as the masking effects at very high levels (loudspeaker performance) or the frequency-dependent hearing threshold at very low levels. Because of this STI performance and the fact that the acoustic simulation is based on the same data basis as the usual calculation of sound pressure levels or reverberation times, it should be increasingly taken into account in standards and guidelines with requirements for workplaces and other areas.

## *Literature*

/1/ IEC 60268-16:2011 "Sound system equipment - Part 16: Objective rating of speech intelligibility by speech transmission index"

/2/ Probst, W.: Validierung eines energiebasierten Schallteilchen-Verfahrens bei der Berechnung der Schallausbreitung in Arbeitsräumen, Lärmbekämpfung Bd. 11, (2016) Nr. 2, S. 56 - 60, Springer-VDI-Verlag GmbH, Düsseldorf

/3/ Z. Maekawa: Noise Reduction by Screens, Applied Acoustics, 1, pp. 157-173

/4/ ISO 9613-2: 1996 "Acoustics - Attenuation of sound during propagation outdoors. Part 2: General method of calculation"