

# Die Anwendung des Speech Transmission Index (STI) zur Beurteilung von Sprachgeräuschen

Wolfgang Probst und Michael Böhm, Gilching

**Zusammenfassung** Die Verstehbarkeit von Sprachinformation ist ein wichtiges Kriterium bei der Beurteilung von Arbeitsplätzen oder anderen Aufenthaltsorten von Menschen. Während in manchen Bereichen, wie z. B. innerhalb von Arbeitsgruppen, eine gute Sprachverständlichkeit erwünscht ist, kann diese durch unerwünschtes Verstehen von Gesprächsinhalten störend sein und auch die Vertraulichkeit von besprochenen Inhalten verletzen. Der erste Schritt zur Beurteilung derartiger durch Sprache ausgelöster Wirkungen und quasi das „Atom“ der Beurteilung von beliebig komplexen Szenarien in Arbeitsstätten ist die Quantifizierung der Sprachverständlichkeit für eine gegebene Anordnung von Sprecher und Hörer durch Berechnung oder Messung des Speech Transmission Index STI. Er berücksichtigt die wesentlichen Einflüsse wie die der Raumerückwirkung, des Hintergrundgeräuschs, einer Verringerung der Modulationstiefe durch die Hörschwelle bei niedrigen und durch Maskierungseffekte bei hohen Signalpegeln. Da von der mit dem STI gegebenen und mit bestehender Technik leicht anwendbaren Beurteilungstechnik noch wenig Gebrauch gemacht wird, werden im Folgenden die zugrundeliegenden Konzepte zusammenfassend und mit einfachen Beispielen erläutert.

## The application of the STI in speech communication

**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 unwanted 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.

Die menschliche Sprache ist im Hinblick auf die Sicherstellung einer optimalen akustischen Umgebung in Arbeitsstätten eine sehr spezielle Geräuschart. Während die Aufgabe im Hinblick auf den Lärm von Maschinen und anderen lärmrelevanten technischen Einrichtungen stets darauf hinausläuft, die Beurteilungspegel an den Arbeitsplätzen so weitgehend wie möglich zu minimieren und so die Störung der Konzentration oder gar die Gefährdung des Gehörs durch Lärmeinwirkung zu vermeiden, kommt es bei der Sprachinformation neben dieser Minimierung im Falle von unerwünschter Einwirkung in vielen Fällen auch auf die Sicherstellung einer guten Verstehbarkeit an. Während Maschinenlärm durch den am Ohr des Betroffenen

einwirkenden Beurteilungspegel hinreichend beschrieben ist, spielt bei der Sprache neben dem von ihr verursachten anteiligen Beurteilungspegel auch die Verstehbarkeit der von ihr transportierten Information eine wesentliche Rolle. Das Verstehen dessen, was gesagt wird, ist aber nicht nur vom anteiligen Schalldruckpegel abhängig, sondern wird von einer Reihe weiterer Einflüsse bestimmt, die zu einer Verringerung der für das Verstehen wichtigen Modulationstiefe führen. Dazu zählt z. B. der Pegel eines stets vorhandenen Hintergrundgeräuschs, der Abstand des Signalpegels von der Hörschwelle, die Verdeckung der Modulation in einem Frequenzband durch den Pegel des benachbarten tieferen Frequenzbands und vor allem das „Verschmieren“ der für das Verstehen notwendigen Modulation durch das aufgrund unterschiedlich langer Laufwege zeitverzögerte Eintreffen eines Sprachsignals.

Alle genannten Aspekte sind mit dem Speech Transmission Index STI nach IEC 60268-16 [1] zu einer Kenngröße verdichtet, die trotz ihrer komplexen „Bauart“ extrem einfach zu verstehen und auch anzuwenden ist. Der STI beschreibt die Qualität der Übertragung von Sprachinformation vom Ort der Sprachquelle – dies kann beispielsweise eine sprechende Person oder ein Lautsprecher sein – bis zu einer festgelegten Hörerposition. Er quantifiziert mit seinem Wert

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

diese Qualität, wobei sich bei gegebenem Schalleistungspegel der Emission auch direkt die Sprachverständlichkeit an der Hörerposition beurteilen lässt. Bei einem  $\text{STI} < 0,2$  wird weniger als 30%, bei  $\text{STI} > 0,5$  wird mehr als 80% des Gesprochenen verstanden – wobei dies selbstverständlich grobe Mittelwertangaben sind.

## Das „Konstruktionsprinzip“ des STI und die Umsetzung in der akustischen Simulation

Für die Berechnung des STI wird ein von der Quelle emittiertes Sprachsignal durch seinen A-bewerteten Schalleistungspegel nach **Tabelle 1** sowie durch ein standardisiertes Frequenzspektrum nach **Tabelle 2** beschrieben.

Zur Ermittlung des Schalleistungspegel-Frequenzspektrums wird in allen sieben Spalten der entsprechenden Zeile in Tabelle 2 der für die Sprachanstrengung zutreffende Wert aus Tabelle 1 addiert.

Entsprechend Tabelle 2 umfasst der die Sprachinformation transportierende Schall die Oktav-Frequenzbänder von 125 Hz ( $k = 1$ ) bis 8 000 Hz ( $k = 7$ ). Durch die sieben Pegelwerte ist aber nur die mittlere Schallintensität in jedem Frequenzband  $f_k$  beschrieben – ein zeitlich konstantes Geräusch mit diesem Pegel würde als Rauschen empfunden und keinerlei Information transportieren. Erst über die Artikulation durch den menschlichen Sprachapparat wird dem Schallsignal – in den verschiedenen Frequenzbereichen unterschiedlich – eine zeitliche Schwankung bzw. eine Modulation aufgeprägt, die zur Wahr-

**Tabelle 1** A-bewerteter Schallleistungspegel sprechender Menschen in Abhängigkeit vom Sprechaufwand (nach Lazarus, 1986).

Sprachaufwand	LWA dB
flüsternd	47
leise	53
sehr entspannt	59
normal entspannt	65
normal angehoben	71
angehoben	77
laut	83
sehr laut	89
schreien	95

**Tabelle 2** Auf einen A-bewerteten Gesamtpegel 0 dB normierte lineare Bandpegel von Sprache (nach [1]).

Frequenz (Hz)	125	250	500	1000	2000	4000	8000
$\Delta$ dB (Männer)	2,9	2,9	-0,8	-6,8	-12,8	-18,8	-24,8
$\Delta$ dB (Frauen)	-	5,3	-1,9	-9,1	-15,8	-16,7	-18,0

nehmung von Lauten, Silben und Sätzen führt und das Verstehen der damit übertragenen Information ermöglicht.

Diese Schwankung bzw. Modulation erfolgt in einer zeitlichen Folge, die zur Beurteilung der Übertragungsqualität von Sprache durch die 14 Terzfrequenzen  $f_m$  von 0,63 bis 12,5 Hz beschrieben werden kann.

Mit der Prinzip-Darstellung (**Bild 1**) soll dies verdeutlicht werden. Der vom Sprecher (A) emittierte und zeitlich je nach Modulationstiefe schwankende Sprachschall (B) breitet sich aus und weist durch unterschiedliche Einflüsse beim Hörer eine oft kleinere Schwankung bzw. eine geringere Modulationstiefe auf.

Man kann diesen Effekt beschreiben, indem man diese Verringerung der Modulationstiefe für jedes der sieben Frequenzbänder aus Tabelle 2 getrennt untersucht. Hierzu wird dem Oktavbandrauschen bei der betreffenden Frequenz  $f_k$  - z. B. 1000 Hz - eine Modulation der Frequenz  $f_m$  - z. B. 1 Hz - aufgebracht, wie dies in Zeile C bildlich und in Zeile D mathematisch ausgedrückt ist. Die Modulationstiefe des vom Sprecher emittierten Schalls der Intensität  $I_s$  ist im Formalismus der Zeile D durch die Amplitude bzw. dem Modulationsindex  $m_s$  bestimmt. Beim Hörer kommt dieser Schall mit einer Intensität  $I_R$  und einem in der Regel verminderten Modulationsindex  $m_R$  an. Die Verringerung der Modulationstiefe von der Sprecher- bis zur Hörerposition wird dann für dieses Frequenzband  $f_k$  (im Beispiel 1000

Hz) und bei dieser Modulationsfrequenz  $f_m$  (im Beispiel 1 Hz) durch den Quotienten der beiden Modulationsindizes, den sog. Modulations-Transfer-Faktor MTF, ausgedrückt.

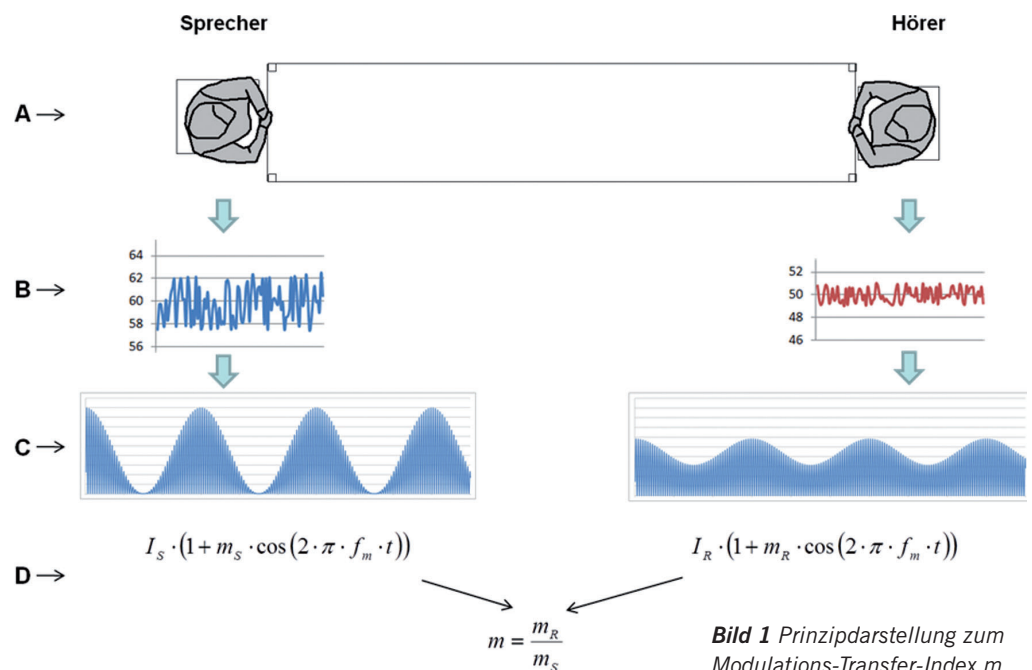
Dieses mathematisch hier nicht weiter ausgeführte Konzept zeigt den wesentlichen Kern der Beurteilung der Sprachübertragung durch die Verringerung der Modulationstiefe. Die zahlreichen oben genannten die Sprachübertragung beeinflussenden Phänomene, wie z. B. die Halligkeit eines Raums oder zusätzlich einwirkender Hintergrundlärm, werden letztlich dahingehend beurteilt, wie sie die 98 Modulations-Transfer-Faktoren (sieben Oktav-Frequenzbänder mit je 14 Modulationsfrequenzen) vermindern.

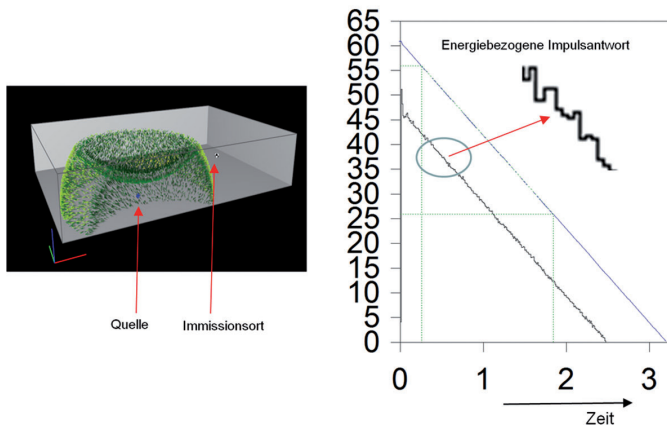
Im Folgenden soll - etwas vereinfacht - dargestellt werden, wie sich die Halligkeit eines Raums über die Verminderung der Modulationstiefe auf die Verschlechterung der Sprachverständlichkeit auswirkt und wie dies mit Simulationsberechnungen prognostiziert werden kann.

Bei der Simulationsrechnung werden, wie links in **Bild 2** dargestellt, von der Quelle Schallteilchen in alle Richtungen ausgesandt und ihre Wege berechnet. Die mit einem Teilchen verbundene Schallenergie wird bei jeder Umlenkung durch Reflexion entsprechend dem Absorptionsgrad dieser reflektierenden Fläche geschwächt. Aus der Zahl der ein kleines Zählvolumen um den Immissionspunkt durchlaufenden Teilchen und über die Summation der von ihnen transportierten Schallenergie wird letztlich der Schalldruckpegel berechnet.

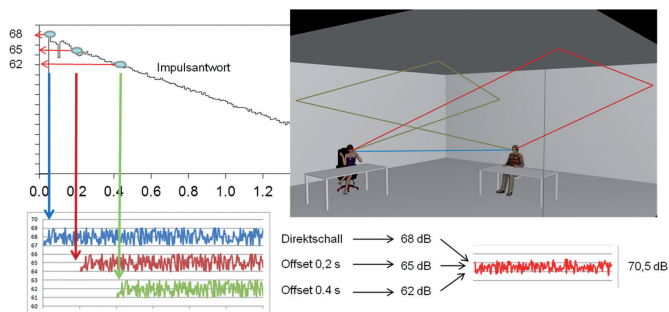
Da die Länge der einzelnen Laufwege bis zum Eintreffen am Zählvolumen bekannt ist, kann so unter Berücksichtigung der Schallgeschwindigkeit auch die seit der Abstrahlung vergangene Zeit bestimmt werden. Damit können die von einer Quelle ausgesandten und am Immissionsort eintreffenden Schallenergien sowie die entsprechenden Schalldruckpegel in Zeitklassen unterteilt und als energiebezogene Impulsantwort entsprechend dem Diagramm rechts in Bild 2 dargestellt werden.

Für das weitere Verständnis ist wesentlich, dass sich der Nullpunkt der Zeitachse rechts in Bild 2 auf den Zeitpunkt der Aussendung der Schallteilchen bezieht, unabhängig davon, in welcher Zeitabfolge die Berechnung der einzelnen Wege erfolgt. Die zur Impulsantwort parallele oberhalb gelegene Abklingkurve





**Bild 2** Ausbreitung von Schallteilchen (links) bei der Simulation und zeitliche Klassierung zur Bestimmung der Impulsantwort (rechts).



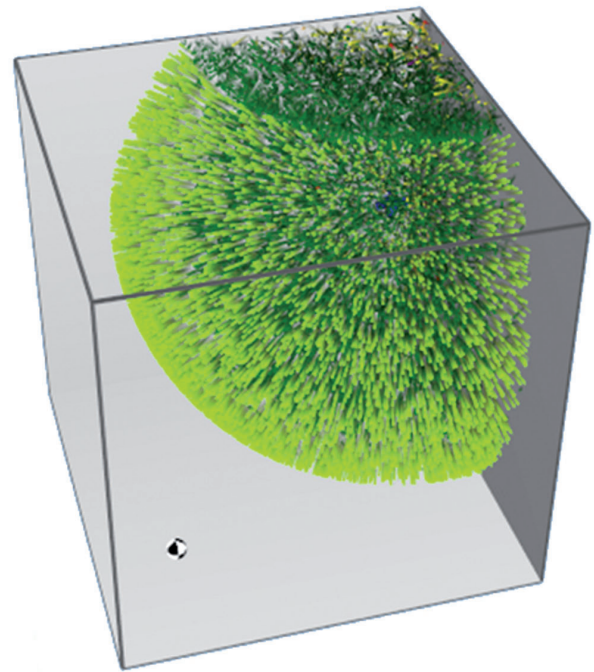
**Bild 3** Ausbreitung von Sprachschall (oben rechts), Impulsantwort (oben links), daraus sich ergebende zeitverschobene Sprachsignale (unten links) und letztlich Summenschall mit geminderter Modulation.

entsteht durch Rückwärtsintegration bzw. durch Summation der Schallenergie vom betreffenden Zeitpunkt bis zum Ende der Zeitskala. Diese Kurve entspricht dem Pegelverlauf nach Abschalten einer stationär abstrahlenden Schallquelle.

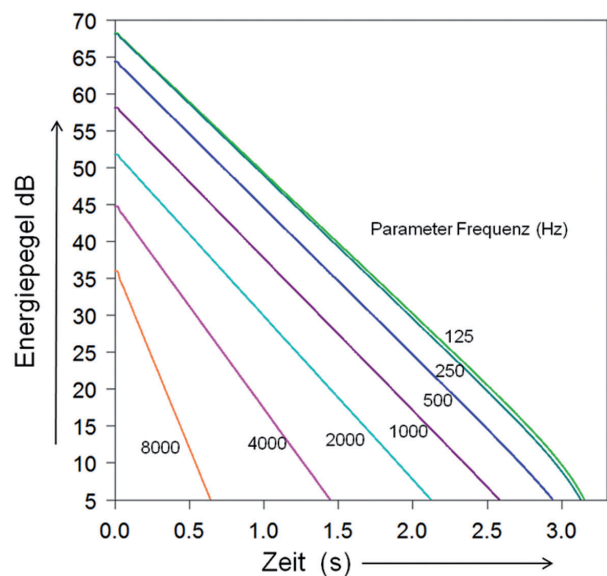
Diese Zusammenhänge gelten auch für Sprachschall. Mit **Bild 3** wurde versucht, dies etwas vereinfachend darzustellen. Oben rechts im Bild sind drei Strahlwege von der telefonierenden Sprecherin zum – hier unfreiwilligen – Hörer dargestellt. Dieselbe zu einem bestimmten Zeitpunkt gesprochene Silbe kommt entsprechend der Impulsantwort mit einem Pegelanteil von 68 dB als Direktschall, aber auch mit 65 dB und einer Verschiebung von 0,2 s sowie mit 62 dB und einer Verschiebung um 0,4 s beim Hörer an (alle Werte sind beispielhaft und dienen nur der Erläuterung).

Die Summe dieser drei Anteile weist mit 70,5 dB zwar einen höheren Pegel auf, die zeitverschobene Überlagerung führt jedoch zu einer Verringerung der Modulationstiefe und damit zu einer schlechteren Verständlichkeit. Dies entspricht auch der Erfahrung – im halligen Raum steigt zwar die empfundene Lautstärke an, die Verständlichkeit sinkt aber wegen der zeitlichen „Verschmierung“ des Sprachsignals.

Bei der Berechnung des oben genannten Modulations-Transfer-Index für das Oktavband  $f_k$  mit der Modulationsfrequenz  $f_m$  wird die das Schallsignal repräsentierende Funktion  $F_s$  (Bild 3 unten links) entsprechend dem zeitlichen Abklingen der über Messung oder Simulationsrechnung bestimmten Impulsantwort schrittweise zeitlich verschoben, abgeschwächt und zum



**Bild 4** Modell-Hallraum mit Teilchenwolke.



**Bild 5** Mit Simulation ermittelte Abklingkurven.

Gesamtssignal summiert (dies entspricht mathematisch der Faltung von Funktion  $F_s$  und Impulsantwort). Mit dem dann verbleibenden Modulationsindex  $m_R$  kann dann der entsprechende Modulations-Transfer-Index  $m$  bestimmt werden.

Aus den 98 Modulations-Transfer-Indizes wird dann in mehreren Schritten der letztlich interessierende Speech Transmission Index STI berechnet.

### Einfluss von Impulsantwort und Nachhallzeit auf den STI

Als Beispiel für den Einfluss von Nachhall auf den STI sei dieser für den Fall berechnet, dass sich Sprecher und Hörer in dem in [2] beschriebenen Modell eines Hallraums befinden. **Bild 4** zeigt dieses Hallraum-Modell mit der sich ausbreitenden Teil-



chenwolke und in **Bild 5** ist das Diagramm der aus der energiebezogenen Impulsantwort – die im Folgenden auch als Echogramm bezeichnet wird – ermittelten Abklingkurven dargestellt.

Im diffusen Schallfeld zeigen die mit logarithmischer Ordinatenskala aufgetragenen Abklingkurven entsprechend Bild 5 einen linearen Verlauf und jede dieser Kurven kann durch eine Nachhallzeit eindeutig beschrieben werden. In diesem Spezialfall führt die oben genannte Berechnungsmethodik zu einer analytisch geschlossenen Lösung

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

mit

$f_m$ : Modulationsfrequenz in Hz

$m$ : Modulations-Transfer-Index

$T$ : Nachhallzeit in s

$m(f_m)$ : Modulations-Transfer-Index bei der Modulationsfrequenz  $f_m$  und für die Mittenfrequenz des Oktavbands, auf das sich die Nachhallzeit  $T$  bezieht.

**Bild 6** zeigt im oberen Teil die mit dem SERT-Verfahren (siehe [2]) durch Simulation ermittelten Nachhallzeiten, die mit Gl. (1) aus diesen berechneten 98 Modulations-Transfer-Indizes und schließlich den daraus berechneten Wert des STI.

Dieses Beispiel eignet sich auch als Testaufgabe. Mit den über die Simulation ermittelten Echogrammen muss sich derselbe Wert ergeben.

Hintergrundlärm liefert einen mit dem Sprachsignal unkorrelierten Beitrag. Er führt ebenso zu einer Minderung der Modula-

$f_k$ (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

$f_m$ Hz	Frequenzband (Hz)						
	125	250	500	1000	2000	4000	8000
0.63	0.74	0.75	0.76	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	0.51	0.54	0.63	0.82
1.6	0.40	0.40	0.41	0.42	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	0.07	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**

**Bild 6** Die ermittelten Nachhallzeiten, Modulations-Transfer-Indizes und der daraus bestimmte STI.

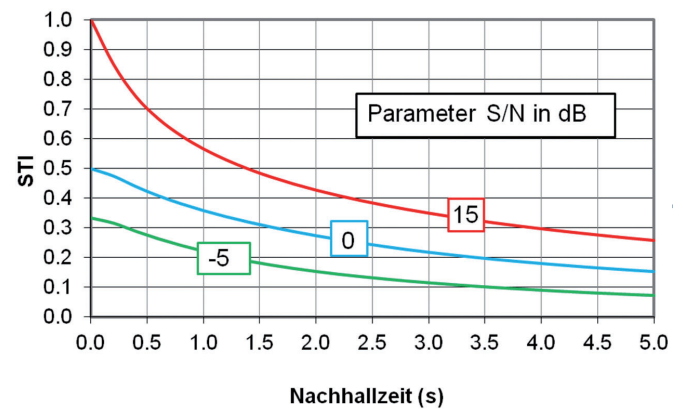
tionstiefe und damit des STI. Dabei ist nicht der absolute auf den Hintergrundlärm bezogene Schalldruckpegel, sondern der S/N-Wert (Signal to Noise Ratio angegeben in dB) bzw. die Differenz der Pegel von Sprachsignal und von Hintergrundgeräusch ausschlaggebend. Die Berechnung bei Variation des S/N und der Nachhallzeit  $T_{60}$  führt auf die in den **Bildern 7** und **8** dargestellte Abhängigkeit.

Die Kurve mit einem S/N von 15 dB – also praktisch kein Fremdgeräuscheinfluss – zeigt, dass selbst in diesem „idealen“ Fall Nachhallzeiten unter ca. 1,5 s erforderlich sind, damit eine noch einigermaßen akzeptable Sprachverständlichkeit gegeben ist.

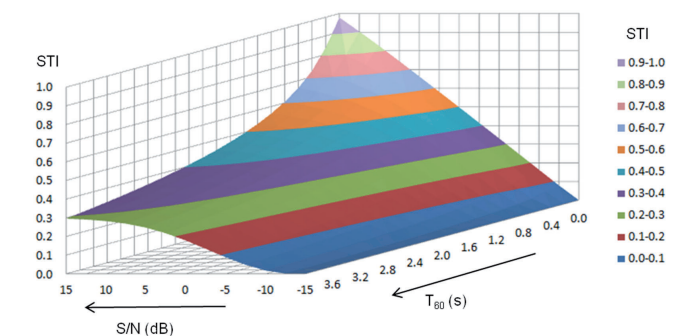
Bei der längsten in Bild 8 einbezogenen Nachhallzeit von 3,6 s kann, wie die entsprechende vordere Randkurve der STI-Fläche zeigt, auch bei fehlendem Hintergrundlärm mit  $S/N \geq 15$  dB kein höherer STI als 0,3 erreicht werden.

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 in **Bild 9**, mit dem der STI bei den drei Werten des Signalabstands in Abhängigkeit von dem beim Hörer vorhandenen Signalpegel in dB(A) dargestellt ist.

Allerdings sollte man die beschriebene Abhängigkeit der Sprachverständlichkeit von der Nachhallzeit nicht überbewerten – sie gilt nur dann, wenn sich der Hörer in einem Bereich befindet, in dem eine durch die Nachhallzeit beschriebene lineare Abklingkurve vorliegt.



**Bild 7** Der STI in Abhängigkeit von der Nachhallzeit für drei S/N-Werte in dB.



**Bild 8** Der STI in Abhängigkeit vom S/N-Abstand in dB und von der Nachhallzeit (nur bei diffusem Schallfeld).

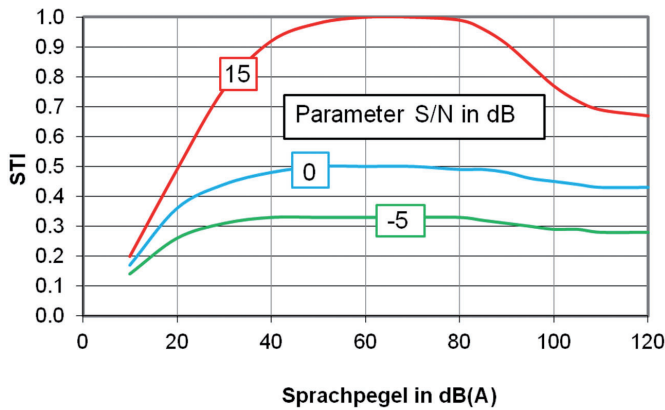


Bild 9 Abhängigkeit des STI vom A-bewerteten Signalpegel.

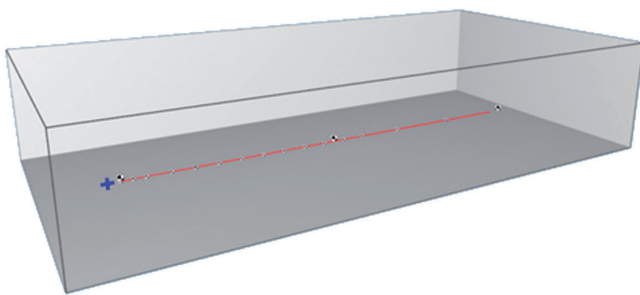


Bild 10 Leere Halle (50 m x 25 m x 10 m) mit Quelle und Messpfad.

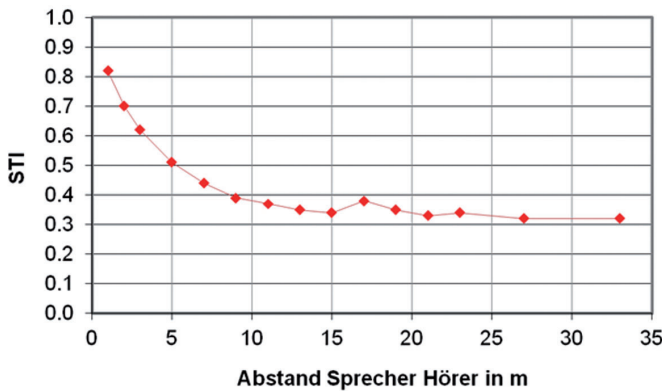


Bild 11 Der STI in Abhängigkeit vom Abstand des Hörers vom Sprecher in der großen Halle.

Dies zeigt schon eine einfache Berechnung für die in Bild 10 dargestellten leeren Halle mit den Abmessungen 50 m x 25 m x 10 m: Sie ist mit einem mittleren Absorptionsgrad von 0,1 als sehr schallhart angenommen. Aufgrund einer Nachhallzeit von 4 bis 5 s wäre nach Bild 7 ohne Fremdlärm ein STI von unter 0,3 zu erwarten.

Tatsächlich ergibt sich aus den Echogrammen an den auf einem geraden Pfad verteilten Immissionspunkten der im Diagramm in Bild 11 dargestellte Verlauf des STI (berechnet mit Simulation).

Der STI von 0,5 bis 0,8 in Sprechernähe zeigt trotz einer hohen Nachhallzeit des Raums von über 4 s eine gute Sprachverständlichkeit an. Dies deckt sich mit der Erfahrung – auch in großen und schallharten Räumen, wie Kirchen oder den großen Aulen in manchen Hotels oder Geschäftsbauten, gibt es bei geringem

Hintergrundpegel kein Problem mit der Sprachverständlichkeit bei üblichem Sprecherabstand. Den Grund zeigt das Echogramm in 1 m Abstand vom Sprecher (Bild 12) – der Pegel fällt erst um fast 30 dB steil ab, und erst dann setzt der reflexionsbedingte Abklingvorgang an.

Dies ist der große Vorteil der Verwendung des aus dem Echogramm berechneten STI. Weit über die Beurteilung von Nachhallzeiten hinausgehend umfasst er viele die Sprachverständlichkeit beeinflussenden Phänomene und liefert so eine wesentlich zuverlässigere Prognose von sprachbedingten Wirkungen.

### STI in Umgebungen mit abschirmenden Einrichtungen

Beugungseffekte können bei der Berechnung der Schallausbreitung in Räumen oft vernachlässigt werden, weil Schall über zahlreiche Ausbreitungswege auch die vom Direktschall abgeschirmten Bereiche erreicht und dort den letztlich sich einstellenden Pegel bestimmt. Mit dem hier zugrundeliegenden Schallteilchenverfahren SERT [2] werden die eine Kante passierenden Schallteilchen so abgelenkt, dass sich hinter einem Schirm das nach Maekawa [3] experimentell gefundene und auch der ISO 9613-2 [4] zugrunde liegende Beugungsfeld ergibt.

Dabei wird der von einer Schallquelle S an einem Immissionsort R bei Ausbreitung im Freifeld verursachte Schalldruckpegel durch Einfügung eines breiten Schirms mit einer für die Beugung bestimmenden Kante um ein frequenzabhängiges Abschirmmaß  $D_z$  vermindert. Ist die Länge des Wegs über die Schirmkante um den Umweg  $z$  größer als die Länge des direkten Wegs – ohne Schirm – zwischen Quelle S und Immissionsort R, so ergibt sich ein Abschirmmaß von

$$D_z = 10 \cdot \log(3 + 20 \cdot N) \text{ dB} \quad (2)$$

$$\text{mit } N = \frac{z}{\lambda} \cdot z$$

$\lambda$  ist die Schall-Wellenlänge.

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

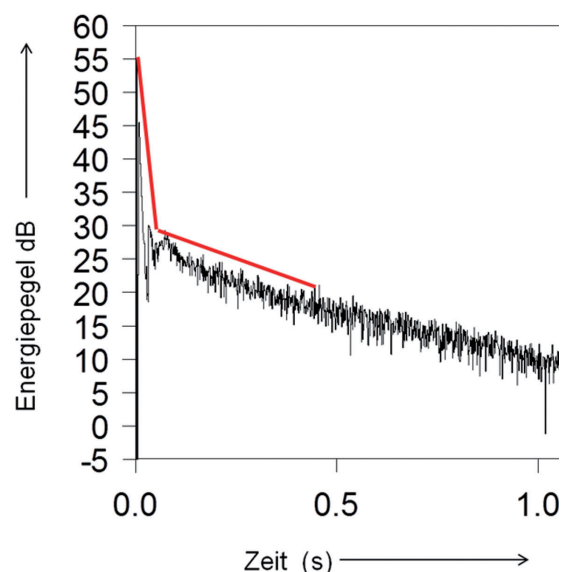


Bild 12 Echogramm in 1 m Abstand vom Sprecher (500 Hz-Band).

Schalldruckpegel entsprechend Gl. (2) ergeben. **Bild 14** zeigt die Ablenkung der Teilchenbahnen in den abgeschirmten Bereich.

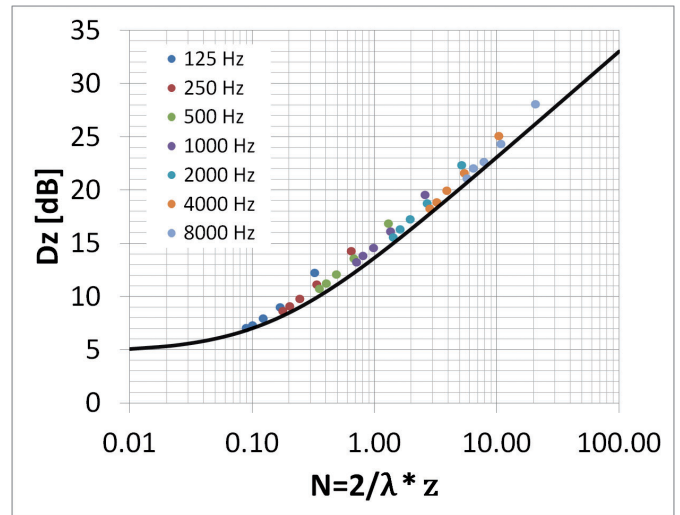
Aus den für die Frequenzbänder 125 bis 8 000 Hz ermittelten Schalldruckpegeln mit und ohne Schirm wurden die Abschirmmaße  $D_z$  ermittelt und mit der Kurve der Gl. (2) in das Diagramm in **Bild 15** eingetragen.

Damit können auch in den seltenen Fällen, in denen keine anderen Quellen einwirken und der Schallpegel hinter einem Schirm nur vom gebeugten Schall bestimmt wird, diese Schallpegel und die STI-Werte korrekt bestimmt werden. Dies ist insbesondere dann von Interesse, wenn Blenden bzw. absorbierende Stellwände zwischen Arbeitsplätzen vorgesehen werden, um die Störung durch Telefonate oder andere sprachliche Kommunikation zu verhindern.

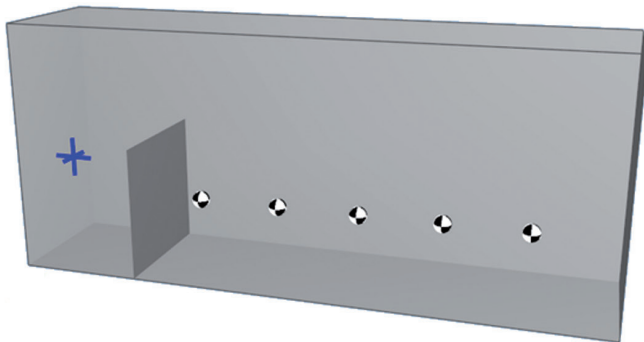
Die Beurteilung der Wirkung von Abschirmungen auf die Sprachverständlichkeit sei mit dem Modell in **Bild 16** gezeigt. Der Raum mit den Abmessungen 20 m x 10 m x 6 m ist durch eine 2,5 m hohe Trennwand in einen 6 m breiten Teilraum mit reflektierenden Begrenzungsflächen und einen 14 m breiten Teilraum mit hochabsorbierenden Begrenzungsflächen unterteilt. Im ersten Schritt werden mit einem Quellspektrum gemäß normal angehobener Stimme nach Tabellen 1 und 2 die Schalldruckpegel und die STI-Werte ohne Schirm berechnet – wie die Ergebnisse in **Bild 17** zeigen, wird der Sprecher im gesamten Raum gut verstanden.

Die Ergebnisse der Berechnung mit Schirm sind in **Bild 18** dargestellt.

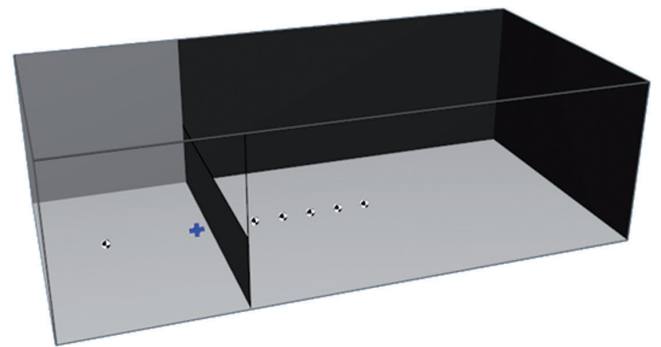
Der hinter dem Schirm deutlich geringere STI ergibt sich, weil die Pegel bereits so niedrig sind, dass die Hörgrenze zur Beschränkung der verbleibenden Modulationstiefe beiträgt. Dies kann gezeigt werden, indem dieselbe Berechnung mit einem Schallleistungspegel der Sprache von 100 dB(A) – statt wie oben mit 70 dB(A) – wiederholt wird. Das Ergebnis in **Bild 19** zeigt, dass der STI trotz sonst gleichem Berechnungsablauf ansteigt, weil die Hörgrenze keine Rolle mehr spielt.



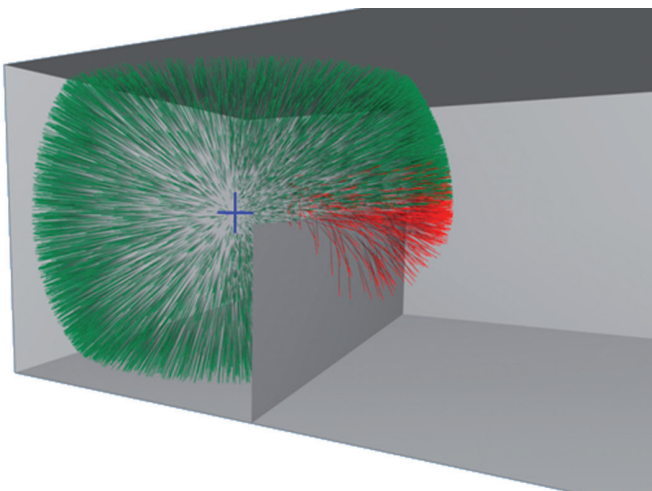
**Bild 15** Die für sieben Frequenzbänder und fünf Immissionsorte ermittelten Abschirmmaße im Vergleich zur Sollkurve.



**Bild 13** Schirm mit Quelle und Berechnungspunkten.



**Bild 16** Raum mit Schirm, Quelle und Empfängerpositionen.



**Bild 14** Ablenkung der Schallteilchen zur Erzeugung des Beugungsfelds.

0.94	Source	0.97	0.95	0.93	0.91	0.90 ← STI
⊗	+	⊗	⊗	⊗	⊗	⊗
52.23		54.52	51.54	49.39	47.73	46.31 ← dB(A)

**Bild 17** Schalldruckpegel und STI-Werte ohne Schirm.

0.94	Source	screen				
⊗	+	0.60	0.63	0.61	0.60	0.58 ← STI
52.23		⊗	⊗	⊗	⊗	⊗
		29.89	33.81	32.80	31.91	31.39 ← dB(A)

**Bild 18** Mit Teilchenbeugung berechnete Schalldruckpegel und STI-Werte mit Schirm.



Schließlich wird die Berechnung noch mit einem – nur mit Lautsprecher als Schallquelle möglichen – Schalleistungspegel von 120 dB(A) wiederholt.

Wie der Vergleich der Ergebnisse in den **Bildern 19** und **20** zeigt, bleiben die STI – Werte im abgeschirmten Bereich auch bei extrem erhöhter Quellen-Schalleistung mit  $L_{WA} = 120$  dB im Wesentlichen gleich – für die Änderung der Modulationstiefe von der Quelle bis zum Empfänger spielt es keine Rolle, ob das emittierte und damit auch das empfangene Sprachsignal um denselben Pegelwert erhöht oder erniedrigt wird. Allerdings tritt – ähnlich wie aufgrund des „Eintauchens“ in die Hörschwelle nach unten – auch bei sehr hohen Pegeln nach oben ein den STI verringernder Effekt auf, wie der Vergleich der STI-Werte an dem von der Quelle nicht abgeschirmten Empfängerposition ganz links zeigt. Bei dem Quellen-Schalleistungspegel von 120 dB(A) führt der Maskierungseffekt jedes Frequenzbands für das benachbarte höhere Frequenzband zur Verschlechterung der empfundenen Modulationstiefe und damit zur entsprechenden Verringerung des STI. Dies ist auch plausibel. Eine Sprachdarbietung durch eine auf die Schalleistung von 120 dB(A) hochgefahrenen Verstärker-Lautsprecher-Kombination in 3 m Abstand führt durch die genannte Verdeckung zur schlechteren Verständlichkeit.

Schließlich sei noch der Fall betrachtet, dass im Beispiel von Bild 18 – also normale erhobene Sprache – ein in Arbeitsbereichen kaum zu unterschreitender Hintergrundpegel von 30 dB(A) vorhanden sei. Dies führt dann zu den Ergebnissen nach **Bild 21** und damit zu einem STI von 0,3 bis 0,4.

Dies ist auch die wesentliche „Botschaft“ des STI-Konzepts: Durch Abschirmungen wird der Sprachschall zwar vermindert, aber dies verringert nur den Absolutpegel, nicht jedoch die Dynamik bzw. Modulationstiefe. Wird ein Sprecher in absolut leiser Umgebung vom Hörer abgeschirmt, wird er nur durch „Anschlag“ an der individuellen Hörschwelle schlechter verstanden. In der Praxis bildet aber nicht die Hörschwelle, sondern das stets vorhandene Hintergrundgeräusch den unteren Anschlag.

0.92 ⊗ 82.23	Source +	screen 0.87 0.78 0.86 0.81 0.86 ← STI ⊗ ⊗ ⊗ ⊗ ⊗ 60.39 63.81 62.93 61.79 61.37 ← dB(A)
--------------------	-------------	--

**Bild 19** Mit Schirm und mit Teilchenbeugung, Schalleistungspegel der Quelle erhöht auf 100 dB(A).

0.72 ⊗ 102.24	Source +	screen 0.81 0.79 0.79 0.81 0.80 ← STI ⊗ ⊗ ⊗ ⊗ ⊗ 79.97 83.88 82.77 81.97 81.50 ← dB(A)
---------------------	-------------	--

**Bild 20** Mit Schirm und mit Teilchenbeugung, Schalleistungspegel der Quelle erhöht auf 120 dB(A).

0.93 ⊗ 52.22	Source +	screen 0.34 0.38 0.39 0.37 0.33 ← STI ⊗ ⊗ ⊗ ⊗ ⊗ 29.78 33.83 32.77 31.81 31.38 ← dB(A)
--------------------	-------------	--

**Bild 21** Mit Schirm und mit Teilchenbeugung, Hintergrundpegel durch Sprachgeräusche 30 dB(A).

Dies wäre auch die Zielrichtung bei einer verbesserten Strategie. Zur Beurteilung der Sprachverständlichkeit zwischen einzelnen Arbeitsplätzen oder Arbeitsplatzgruppen ist es erforderlich, den typischen Hintergrundlärm durch Klassifizierung von typischen Nutzungen oder durch individuelle Berechnung einzubeziehen. Dies würde die akustische Optimierung von Arbeitsstätten und anderen Aufenthaltsräumen gegenüber der derzeit starren Einordnung aufgrund von Ober- und Untergrenzen der Nachhallzeit erheblich verbessern.

### Planungsprinzipien und -beispiele

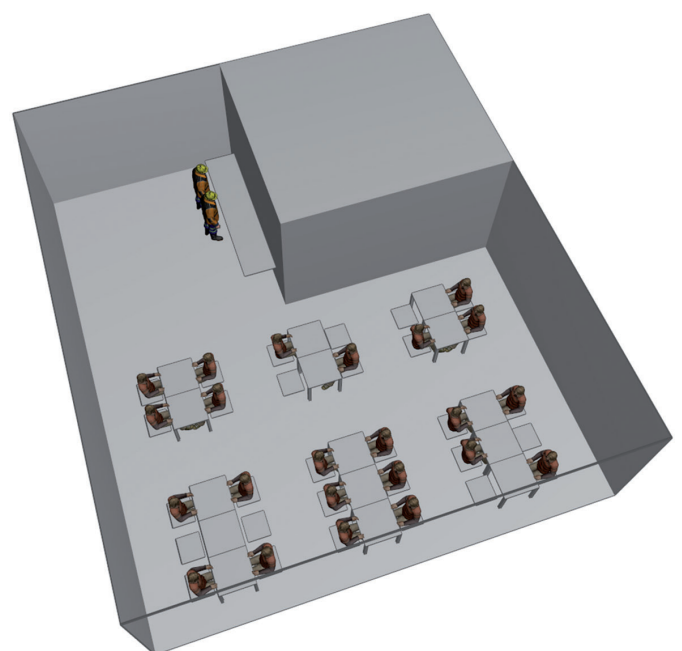
Die planerische Vorausberechnung von Schalldruckpegel und STI ermöglicht eine zielgerichtete Beurteilung und Optimierung von Räumen, in denen die Kommunikation durch Sprache eine wesentliche Rolle spielt.

Jeder kennt dieses Problem von Gaststätten, die man nur deswegen meidet, weil es dort bei guter Belegung erfahrungsgemäß sehr laut wird und eine entspannte Unterhaltung nicht mehr möglich ist. Dies gilt für alle Raumgrößen vom kleinen Bistro mit aus Platzgründen meist enger Bestuhlung bis zum großflächigen dank Glaswänden lichtdurchfluteten Restaurant mit dem bekannten akustischen Charme von Bahnhofshallen.

Ursache der „akustischen Verschmutzung“ sind die Gespräche der übrigen Gäste – jede sprechende Person stellt eine Schallquelle dar, die bei der Planungsberechnung durch eine Emission gemäß Tabelle 1 zu berücksichtigen ist.

Wie laut eine Person spricht, hängt vom Pegel des zu über-tönenden Hintergrundlärms und vom maximalen Abstand der Personen ab, mit denen man kommuniziert. In diesem Sinne bilden die Bier- und Weinkeller mit langen Tischen in historischen Gewölben das negative Ende der Bewertungsskala.

Das Prinzip eines Planungsvorgangs sei mit dem Modell eines kleinen Restaurants neben einer Essenausgabe in **Bild 22** erläutert. Ausgangspunkt ist die Erstellung eines derartigen Modells, die aber unter Nutzung moderner Softwaretools zeitsparend zu erledigen ist. Zielsetzung ist dabei, dass auch bei voller Belegung die Gesprächspartner am eigenen Tisch verstanden werden kön-



**Bild 22** Kleines Restaurant mit Essenausgabe.

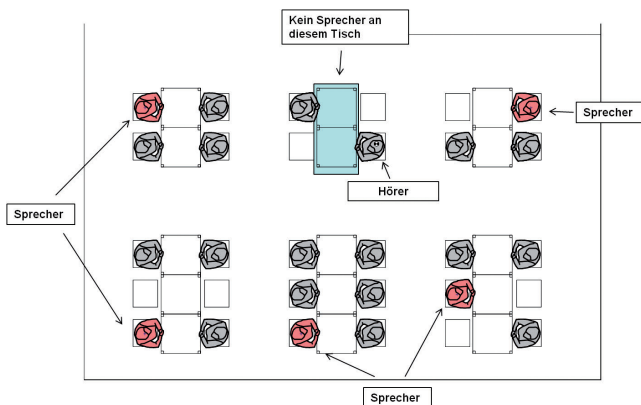
nen. Hierbei ist der STI wie nachfolgend beschrieben eine entscheidende Kenngröße.

Zunächst wird an jedem Tisch eine Sprecher- und eine Hörerposition in typischer Position festgelegt. Ihr Abstand ist umso größer, je länger oder großflächiger die Tische sind.

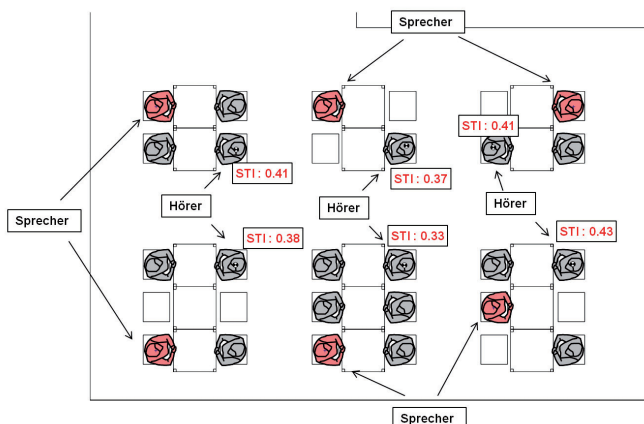
Anschließend erfolgt für einen ersten Tisch die Berechnung des an der Hörerposition zu erwartenden Schalldruckpegels mit den Sprecherpositionen an allen anderen Tischen als Schallquelle, wie es in **Bild 23** dargestellt ist. Der Schalldruckpegel an der Hörerposition des blau markierten Tisches ergibt sich durch die in rot markierten Sprecher an allen anderen Tischen. Dieser Schalldruckpegel bildet den Hintergrundpegel, der für die Berechnung des STI für das Gespräch an diesem Tisch herangezogen wird. Weitere ständig vorhandene Fremdgeräusche durch raumlufttechnische Anlagen oder Außenlärm werden ggfs. durch energetische Addition berücksichtigt.

Unter Berücksichtigung dieses Hintergrundpegels wird nun mit einer Simulationsrechnung der diesen Tisch kennzeichnende STI bestimmt, wobei der Sprecher die Schallquelle und der entfernteste Hörer an diesem Tisch die Empfangsposition darstellt.

Dieser Ablauf wird – möglichst automatisiert – für alle Tische wiederholt. Exemplarische Ergebnisse für die kennzeichnenden STI-Werte für die einzelnen Tische sind für dieses Beispiel in **Bild 24** dargestellt. Die STI-Werte sind deutlich kleiner als 0,5 und zeigen eine schlechte Sprachverständlichkeit an.



**Bild 23** Bestimmung des Hintergrundpegels an der Hörerposition am blau hinterlegten Tisch.

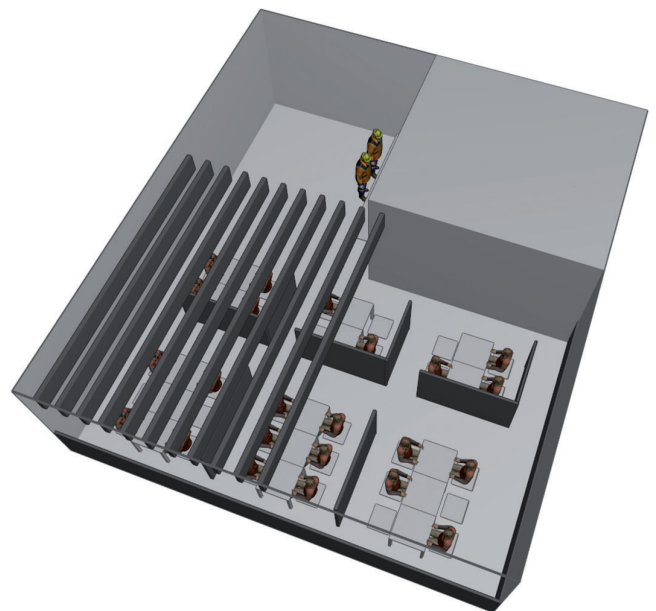


**Bild 24** Kennzeichnende STI-Werte für einzelnen Tische.

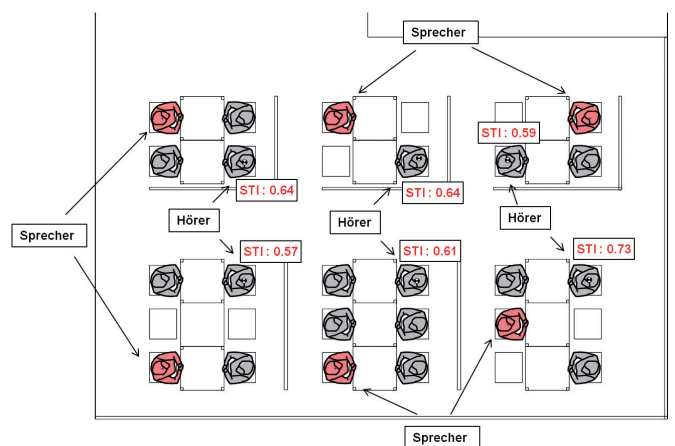
Hier setzt der bekannte Effekt ein, dass an allen Tischen lauter gesprochen wird, um trotz des Hintergrundlärms verstanden zu werden – eine Neuberechnung mit diesem erhöhten Sprachschallpegel würde aber mit dem entsprechend erhöhten Hintergrundpegel zum selben Ergebnis für die STI-Werte führen.

Die richtige Lösung ist eine akustische Planung mit dem hier nicht weiter behandelten Instrumentarium von Absorption und Abschirmung sowie ggfs. einer Änderung der Anordnung. In dem einfach gehaltenen Beispiel erfolgt eine gewisse akustische Trennung durch absorbierende Stellwände in Verbindung mit absorbierender Verkleidung zweier zueinander senkrechter Wände sowie der Abhängung eines schallabsorbierenden Kulissensystems (**Bild 25**).

Die erneute Simulationsrechnung entsprechend dem beschriebenen Ablauf führt auf die in **Bild 26** dargestellten Werte des STI und zeigt damit eine wesentliche Verbesserung der von den Gästen wahrgenommenen akustischen Atmosphäre an.



**Bild 25** Akustische Verbesserung durch Kulissendecke und absorbierende Stellwände.



**Bild 26** Durch Simulation ermittelte STI-Werte unter Berücksichtigung der geplanten Maßnahmen.



## Fazit

Die Verstehbarkeit von Sprache spielt bei der Beurteilung der akustischen Qualität von Aufenthaltsräumen eine erhebliche Rolle. Von der Störung der Konzentrationsfähigkeit durch unfreiwilliges Mithören der Kommunikation anderer im Mehrpersonenbüro bis zur Erschwernis einer entspannten Unterhaltung mit dem eigenen Tischnachbarn im Restaurant gibt es zahlreiche Aspekte, die durch eine vorausschauende akustische Planung über die Quantifizierung der Sprachverständlichkeit durch den STI „entschärft“ werden können. Aufgrund seiner Konstruktion berücksichtigt diese genormte Kenngröße die wichtigsten Parameter, die über die akustische Planung im Sinne der gewünschten Optimierung beeinflusst werden können. Der STI berücksichtigt die über eine Simulationsrechnung bestimmte energetische Impulsantwort und gilt somit – im Unterschied zur Nachhallzeit – auch für beliebige Raumformen und -ausstattungen. Durch die Einbeziehung der unvermeidlichen Hintergrundgeräusche sowie der Maskierungseffekte bei sehr hohen Pegeln (Lautsprecherdurchsagen) oder der frequenzabhängigen Hörschwelle bei sehr niedrigen Pegeln sind die wichtigsten Abhängigkeiten abgedeckt. Aufgrund dieser Leistungsfähigkeit des STI und der Tatsache, dass die Bestimmung über die akustische Simulationsberechnung auf derselben Datengrundlage erfolgt wie die übliche Berechnung von Schalldruckpegeln oder Nachhallzeiten sollte er in Normen und Richtlinien mit Anforderungen für Arbeitsstätten und andere Aufenthaltsräume in zunehmendem Maße Berücksichtigung finden.



Dipl.-Phys. Dr. **Wolfgang Probst**,  
Geschäftsführer und Senior Consultant,  
DataKustik GmbH, **Dipl.-Phys. Michael  
Böhm**, Senior Consultant, DataKustik  
GmbH.

## Literatur

- [1] IEC 60268-16: Sound system equipment – Part 16: Objective rating of speech intelligibility by speech transmission index. Genf: International Electrotechnical Commission 2011.
- [2] *Probst, W.*: Validierung eines energiebasierten Schallteilchen-Verfahrens bei der Berechnung der Schallausbreitung in Arbeitsräumen. *Lärmbekämpfung* 11 (2016) Nr. 2, S. 56-60.
- [3] *Maekawa, Z.*: Noise reduction by screens. *Appl. Acoustics* 1 (1968) Nr. 3, S. 157-173.
- [4] ISO 9613-2: Acoustics – Attenuation of sound during propagation outdoors. Part 2: General method of calculation. Genf: International Organization for Standardization 1996.