

Berücksichtigung binauraler Effekte bei der Vorhersage der Leistungsfähigkeit in Büroumgebungen bei Hintergrundsprache

Tobias Renz¹, Philip Leistner², Andreas Liebl²

¹ *Institut für Akustik und Bauphysik, Universität Stuttgart, 70569 Stuttgart, Email: tobias.renz@iabp.uni-stuttgart.de*

² *Fraunhofer-Institut für Bauphysik, 70569 Stuttgart*

Einleitung

Geringe akustische Privatheit und Störungen durch Telefonate und Gespräche sind in Mehrpersonenbüros häufig berichtete Ursachen für Unzufriedenheit. Die akustische Gestaltung von Mehrpersonenbüros mit schallabsorbierenden und -schirmenden Maßnahmen zielt zwar auf eine Reduktion störender Hintergrundsprache und -geräusche, kann die Unzufriedenheit aber nicht beseitigen. Das gezielte Einbringen eines Geräusches über elektronische Sound Masking Anlagen ist an sich eine effiziente Maßnahme, um bereits pegelreduzierte aber immer noch störende Sprachschalle zu verdecken und deren Verständlichkeit zu reduzieren. Das eingebrachte Maskiergeräusch hat selbst keine leistungsmindernde Wirkung auf die kognitive Leistung, dennoch werden derartige Systeme in Büroumgebungen hierzulande von den Mitarbeitern häufig als unangenehm empfunden. Normen zur akustischen Gestaltung von Mehrpersonenbüros, wie die Richtlinie VDI 2569 [1], nehmen bislang kaum Bezug auf Sound Masking.

Die kognitionspsychologischen Wirkungen bürotypischer Schallumgebungen lassen sich nur schwer mittels physikalischer Messgrößen beschreiben. In Mehrpersonenbüros liegen häufig komplexe räumliche Schallumgebungen mit mehreren aus unterschiedlichen Richtungen eintreffenden Sprachschallen vor, die durch ungleich verteilte reflektierende und absorbierende Oberflächen verstärkt werden können. In solchen Konstellationen stellt sich die Frage, ob die Berücksichtigung binauraler Effekte die Vorhersagegenauigkeit erhöhen kann. Die aktuellen Normen in diesem Bereich (z.B. DIN EN ISO 3382-3 [2]) verweisen auf ein Modell von Hongisto [3], das den Sprachübertragungsindex (Speech Transmission Index, STI) als Prädiktor für die Leistungsfähigkeit in Büroumgebungen vorschlägt. STI ist eine Messgröße zur Vorhersage der Sprachverständlichkeit in Räumen [4]. Diesem Modell entsprechend steigen die Leistungsverluste (decrease of performance, DP) mit zunehmenden STI-Werten, die eine gestiegene Sprachverständlichkeit widerspiegeln

$$DP(STI) = \frac{-7}{1 + \exp((STI - 0.4)/0.06)} + 7 \quad [\%]. \quad (1)$$

Dieses von Hongisto eingeführte Modell wird im Folgenden als DP-Modell bezeichnet. Aus dem Verlauf der Kurve wurden die Messgrößen Ablenkungsabstand r_D und Vertraulichkeitsabstand r_P abgeleitet, die den Abstand bezeichnen, bei dem der STI 0,5 bzw. 0,2 erreicht. Die Werte beruhen darauf, dass gemäß dem Modell die

Störwirkung von Sprache auf das Arbeitsgedächtnis für STI-Werte unterhalb 0,5 schnell beginnt zu sinken und für STI-Werte unterhalb 0,2 verschwindet. Der Entwurf der neuen VDI-Richtlinie zur akustischen Gestaltung von Büros, VDI 2569 [1], nimmt zwar Bezug auf den Ablenkungsabstand r_D , spricht aber keine Empfehlung für Richtwerte aus. Dies dürfte u.a. der Tatsache geschuldet sein, dass die Erreichung kurzer Ablenkungsabstände (z.B. $r_D \leq 5$ m, was in der DIN EN ISO 3382-3 [2] für gute akustische Bedingungen exemplarisch genannt wird) sich in der Praxis als schwierig erweist (z.B. [5, 6, 7]).

Nach Normen, wie der DIN EN ISO 3382-3 [2], wird jeweils eine Sprachquelle zu einem Zeitpunkt mittels eines Lautsprechers unter der Annahme simuliert, dass ein einzelner Sprecher am störendsten ist. Aktuelle Untersuchungen haben gezeigt, dass Umgebungen mit zwei und vier Sprechern an unterschiedlichen Arbeitsplätzen störender wirken als mit einem Sprecher [8]. Dieser Effekt kann durch das räumliche Hörvermögen erklärt werden (für einen Überblick des sog. "Cocktail Party" Phänomens siehe z.B. [9]). Da binaurale Schallumgebungen mit mehreren räumlich verteilten Sprechern störender wirken können als gleichwertige monaurale Schallumgebungen, erscheint eine Erweiterung vorhandener Vorhersagemodelle zur Abschätzung der Störwirkung von Hintergrundsprache um binaurale Effekte sinnvoll.

Die Norm DIN EN 60268-16 [10] empfiehlt bei binauralen Messungen mittels Kunstkopf die STI-Werte des besseren Ohres zu verwenden. Das bessere Ohr bezeichnet hierbei das Ohr, an dem höhere STI-Werte anliegen. Van Wijngaarden und Drullman konnten zeigen, dass der STI-Wert des besseren Ohres mit der Sprachverständlichkeit korreliert [11]. Aus diesem Grund liegt die Annahme nahe, dass das DP-Modell binaurale Effekte berücksichtigen kann, wenn der STI-Wert des besseren Ohres als Eingangsgröße verwendet wird. Im Rahmen dieser Studie wurde auf zwei Laborexperimente mit je elf binauralen Schallbedingungen und einer Ruhebedingung zurückgegriffen, um die zuvor genannte Hypothese zu überprüfen.

Methode

In beiden Experimenten wurde je ein Sprecher mittels trockener Aufnahmen des HSM Satztestes [12] simuliert. In Experiment 1 wurden die Aufnahmen im Freifeldraum am Fraunhofer-Institut für Bauphysik und in einem unbesetzten Mehrpersonenbüro mithilfe eines Yamaha MSP3 Studiolautsprechers (Yamaha Corporation,

Hamamatsu, Shizuoka, Japan) wiedergegeben und über einen Kunstkopf HMS III (HEAD acoustic GmbH, Herzogenrath, Deutschland) aufgenommen. Diese Aufnahmen wurden anschließend Probanden im HiPIE-Labor am Fraunhofer-Institut für Bauphysik über Sennheiser HD 280 PRO Kopfhörer (Sennheiser electronics GmbH & Co. KG, Wedemark, Deutschland) wiedergegeben. In Experiment 2 wurden im gleichen Mehrpersonenbüro die Schalle Probanden über Adam A7X Studiolausprecher (ADAM Audio GmbH, Berlin, Deutschland) präsentiert. Die Raumakustik-Parameter wurden nach DIN EN ISO 3382-3 [2] bestimmt ($r_{D,1} = 3$ m, $r_{D,2} = 5$ m, $r_{P,1} = 7$ m, $r_{P,2} = 11$ m, $D_{2,S,1} = 14$ dB, $D_{2,S,2} = 16$ dB, $L_{p,A,S,4m,1} = 28$ dB(A), $L_{p,A,S,4m,2} = 38$ dB(A)). Diese Werte erfüllen die Empfehlungen für gute akustische Bedingungen entsprechend Anhang A der Norm DIN EN ISO 3382-3 [2].

Am Experiment 1 nahmen 24 Probanden (20-65 Jahre, Median = 24) und am Experiment 2 25 Probanden (21-64 Jahre, Median = 38) teil. Alle Teilnehmer waren deutsche Muttersprachler. Experiment 1 wurde im HiPIE-Labor am Fraunhofer-Institut für Bauphysik durchgeführt und Experiment 2 fand in Räumen der Renz Solutions GmbH statt.

Die Teilnehmer bearbeiteten eine Aufgabe zur Messung der Arbeitsgedächtnisleistung (serial recall): Die Ziffern 1 bis 9 wurden in zufälliger Reihenfolge nacheinander dargeboten (700 ms Anzeige mit einer darauffolgenden Pause von 300 ms). Nach einem kurzen Retentionsintervall von 8 s waren die Ziffern in der präsentierten Reihenfolge auf dem Bildschirm auf einem 3x3 Feld auszuwählen. Diese Aufgabe war zwölfmal je Schallbedingung zu bearbeiten. Jede Ziffer, die nicht exakt an der präsentierten Position wiedergegeben wurde, wurde als Fehler gewertet.

In beiden Experimenten wurden Kontrollbedingungen getestet, die kein Signal bzw. das ungestörte Sprachsignal beinhalteten. In Experiment 1 wurde Aufnahmen des Sprachsignals in Freifeldbedingungen und im Mehrpersonenbüro berücksichtigt, während das Sprachsignal in Experiment 2 nur in Bürobewegungen untersucht wurde. In den übrigen neun bzw. zehn Bedingungen wurde dem Sprachsignal ein Maskiersignal überlagert. Hierzu wurde in Experiment 1 ein stationäres Rauschsignal mit einem an das Langzeitspektrum des Sprechers angepassten Frequenzspektrum verwendet, wobei in Experiment 2 auf ein Signal mit einem um 5 dB pro Oktave fallenden Frequenzspektrum zurückgegriffen wurde. Das angepasste Rauschsignal kann bei höherem Signal-Rausch-Abstand (signal-to-noise ratio, SNR) die Arbeitsgedächtnisleistung unter Hintergrundsprache verbessern [13]. Alle Maskiersignale wurden auf 45 dB(A) an der Position der Probanden kalibriert, was einem typischen Schalldruckpegel für Sound Masking Anlagen entspricht (z.B. [14, 15]). Das Sprachsignal wurde mit 33 bis 39 dB(A) wiedergegeben. Das unmaskierte Sprachsignal wurde mit 36 dB(A) dargeboten. Die genannten Schalldruckpegel beziehen sich auf den A-bewerteten äquivalenten Dauerschallpegel L_{pAeq} .

Die Schallbedingungen umfassten SNR-Werte zwischen -12 und -6 dB. Außerdem wurden in beiden Experimenten Richtungsabweichungen zwischen Sprach- und Rauschquelle getestet. Experiment 1 verfolgte das Ziel, den Effekt der räumlichen Trennung einer Maskier- und Sprachquelle im Freifeld und in Büroumgebungen auf die Arbeitsgedächtnisleistung zu testen. Experiment 2 untersuchte drei unterschiedliche Sound Masking Anlagen, ein konventionelles System in einer abgehängten Decke, einen Lautsprecher in einem Sideboard, sowie einen Lautsprecher in einem Schallschirm.

In das in Gleichung 1 beschriebene Modell wurde der maximale STI-Wert am linken und rechten Ohr STI_{max} eingesetzt. Zusätzlich wurde der mittlere STI-Wert am linken und rechten Ohr STI_{mean} ermittelt und dem zuvor genannten Vorgehen gegenübergestellt. Die in Experiment 1 und Experiment 2 gemessenen mittleren DP-Werte je Schallbedingung wurden normalisiert, indem die Werte durch den maximalen DP-Wert innerhalb des Versuches geteilt wurden (mittlere Fehlerrate von 13 bzw. 12 %). Der Spearmans Rangkorrelationskoeffizient r_S und die mittlere quadratische Abweichung (mean squared error, MSE) zu Hongisto's Modell [3] wurden berechnet, um die beiden Modelle zu vergleichen. Außerdem wurde eine nichtlineare Regression mit einer Boltzmann-Sigmoidfunktion durchgeführt analog zu dem von Hongisto vorgeschlagenen Modell (vgl. [3]).

Ergebnisse

Tabelle 1 zeigt die ermittelten MSE-Werte und Spearmans Rangkorrelationskoeffizienten r_S der beiden Modelle mit den Prädiktoren STI_{max} und STI_{mean} . Wie erwartet, eignet sich STI_{max} besser als Prädiktor für die Arbeitsgedächtnisleistung bei Hintergrundsprache.

Tabelle 1: Übersicht über die MSE-Werte und den Spearmans Rangkorrelationskoeffizienten r_S der untersuchten DP-Modelle

Modell	MSE	r_S
STI_{max}	0,10	0,85
STI_{mean}	0,13	0,82

Abbildung 1 zeigt das DP-Modell von Hongisto [3] (angepasst auf DPnorm-Werte von 0-1) sowie die gemessenen Datenpunkte. Die nichtlineare Regression mit einer Boltzmann-Sigmoidfunktion, die auf Basis des vorliegenden Datensatzes durchgeführt wurde, resultierte in den in der Gleichung 2 beschriebenen Werten und ist ebenfalls in Abbildung 1 dargestellt

$$DPnorm(STI_{max}) = \frac{1,4}{1 + \exp(-(STI_{mean} - 0,16)/0,18)} - 0,4 \quad [\%]. \quad (2)$$

Ein Vergleich zu Hongisto's Modell [3] zeigt deutliche Abweichungen zu niedrigen STI-Werten. Entsprechend dem

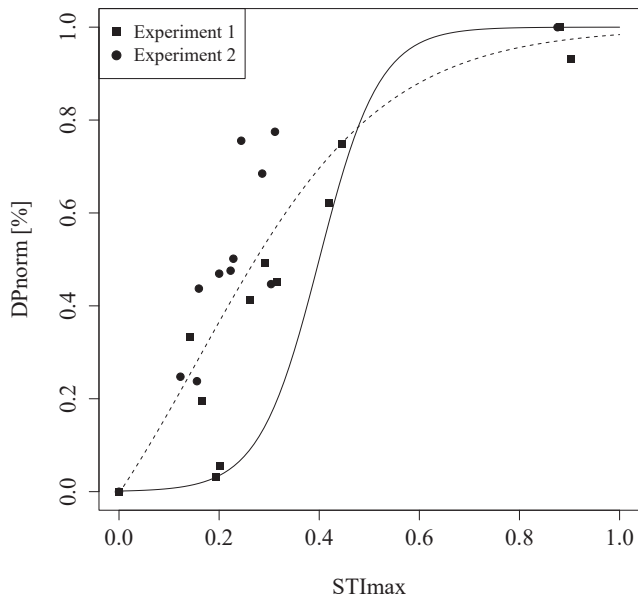


Abbildung 1: Normalisierte DP-Werte über STI_{max} mit den untersuchten Datenpunkten; die durchgezogene Linie zeigt das Modell von Hongisto [3], während die gestrichelte Linie das angepasste Modell $DPnorm(STI_{max})$ darstellt.

angepassten Modell würde die Leistung bereits ab STI -Werten oberhalb 0 anstatt 0,2 beginnen einzubrechen. Beide Modelle nähern sich dem maximalen DP für Werte oberhalb von 0,5 an.

Der Verlauf des in Gleichung 3 und Abbildung 2 dargestellten Modells, das auf dem STI_{mean} basiert, ist ähnlich zu dem in Gleichung 2 beschriebenen Modell. Die Kurve ist jedoch zu etwas niedrigeren STI -Werten verschoben, da der STI_{mean} gleich oder kleiner dem STI_{max} ist

$$DPnorm(STI_{mean}) = \frac{4,3}{1 + \exp(-(STI_{max} + 0,34)/0,30)} - 3,2 \quad [\%]. \quad (3)$$

Diskussion

Im Rahmen der vorliegenden Studie konnte gezeigt werden, dass Hongisto's DP-Modell [3] binaurale Effekte berücksichtigen kann, wenn der STI -Wert des besseren Ohres als Eingangsgröße verwendet wird. Das ermittelte Modell kann als erster Ansatzpunkt dienen, um die Wirkung binauraler Effekte auf die Arbeitsgedächtnisleistung abzuschätzen. In der Praxis würde dies bedeuten, dass monaural gemessene STI -Werte bei der Abschätzung der Störwirkung irrelevanter Hintergrundsprache die Störwirkung in Situationen mit starken Signalunterschieden am linken und rechten Ohr unterschätzen.

Bei der Anwendung von Hongisto's Modell [3] muss berücksichtigt werden, dass aufbauende Studien gezeigt

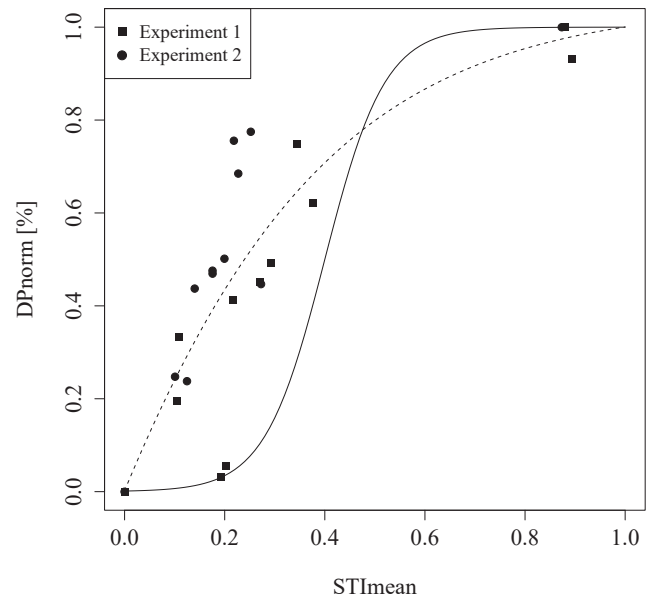


Abbildung 2: Normalisierte DP-Werte über STI_{mean} mit den untersuchten Datenpunkten; die durchgezogene Linie zeigt das Modell von Hongisto [3], während die gestrichelte Linie das angepasste Modell $DPnorm(STI_{mean})$ darstellt.

haben, dass das Modell gewissen Ungenauigkeiten unterliegt (z.B. [16, 17, 18]). Es sollte beachtet werden, dass der STI entwickelt wurde, um die Sprachverständlichkeit abzuschätzen. Bei niedrigen STI -Werten besitzt der STI größere Unsicherheiten (z.B. [19]). Daneben ist bekannt, dass die Sprachverständlichkeit nur in Teilen mit der Arbeitsgedächtnisleistung korreliert und Hongisto's Modell [3] die temporal-spektrale Variabilität von Hintergrundschall unberücksichtigt lässt (z.B. [18]). Ferner kann das Modell nicht berücksichtigen, ob es sich um unverständliche Sprache (Fremdsprache, synthetische Sprachschalle, etc.) oder um verständliche Sprache und damit ggf. störendere Sprache handelt.

Das Modell kann genutzt werden, um die Arbeitsgedächtnisleistung abzuschätzen, wenn Sprache und Hintergrundrauschen binaural und aus unterschiedlichen Richtungen auftreten. In einer vereinfachten Form kann die Störwirkung eines Sprechers unter Hintergrundgeräusch abgeschätzt werden. Situationen, in denen mehrere gleichzeitig redende Sprecher aus unterschiedlichen Richtungen verständlich sind, können nicht abgebildet werden. Ein erweitertes Modell könnte dazu dienen, in frühen Phasen der akustische Planung und Dimensionierung von Mehrpersonenbüros, beispielsweise durch Computersimulationen, die Störwirkungen zwischen Arbeitsplätzen für definierte Sprach- und Hintergrundschalle abzuschätzen.

Literatur

- [1] E VDI 2569:2016-02. Schallschutz und akustische Gestaltung im Büro. Verein Deutscher Ingenieure e.V., Düsseldorf, Deutschland, 2016.

- [2] DIN EN ISO 3382-3:2012-05. Akustik – Messung von Parametern der Raumakustik – Teil 3: Großraumbüros. DIN Deutsches Institut für Normung e.V., Berlin, Deutschland, 2012.
- [3] Hongisto, V.: A model predicting the effect of speech of varying intelligibility on work performance. *Indoor Air* 15(6) (2005), 458-468.
- [4] Steeneken, H. J. M. & Houtgast, T.: A physical method for measuring speech-transmission quality. *The Journal of the Acoustical Society of America* 67(1) (1980), 318-326.
- [5] Haapakangas, A., Hongisto, V., Eerola, M., Kuusisto, T.: Distraction distance and perceived disturbance by noise—An analysis of 21 open-plan offices. *The Journal of the Acoustical Society of America* 141(1) (2017), 127-136.
- [6] Virjonen, P., Keränen, J., Hongisto, V.: Determination of acoustical conditions in open-plan offices: Proposal for new measurement method and target values. *Acta Acustica united with Acustica* 95(2) (2009), 279-290.
- [7] Keränen, J., Virjonen, P., Hongisto, V.: Characterization of acoustics in open offices - four case studies. *Acoustics*, Paris, Frankreich, 2008.
- [8] Yadav, M., Kim, J., Cabrera, D., de Dear, R.: Auditory distraction in open-plan office environments: The effect of multi-talker acoustics. *Applied Acoustics* 126 (2017), 68-80.
- [9] Bronkhorst, A. W.: The cocktail party phenomenon: A review of reserach on speech intelligibility in multiple-talker conditions. *Acta Acustica united with Acustica* 86(1) (2000), 117-128.
- [10] DIN EN 60268-16:2012-05. Elektroakustische Geräte – Teil 16: Objektive Bewertung der Sprachverständlichkeit durch den Sprachübertragungsindex. DIN Deutsches Institut für Normung e.V., Berlin, Deutschland, 2012.
- [11] van Wijngaarden, S. J. & Drullman, R.: Binaural intelligibility prediction based on the speech transmission index. *The Journal of the Acoustical Society of America* 123(6) (2008), 4514-4523.
- [12] Hochmair-Desoyer, I., Schulz, E., Moser, L., Schmidt, M.: The HSM sentence test as a tool for evaluating the speech understanding in noise of cochlear implant users. *The American Journal of Otology* 18(6) (1997), 83.
- [13] Renz, T., Leistner, P., Liebl, A.: Auditory distraction by speech: Sound masking with speech-shaped stationary noise outperforms -5 dB per octave shaped noise. *The Journal of the Acoustical Society of America* 143(3) (2018), EL212-EL217.
- [14] Veitch, J. A., Bradley, J. S., Legault, L. M., Norcross, S. G., Svec, J. M.: Masking speech in open-plan offices with simulated ventilation noise: Noise level and spectral composition effects on acoustic satisfaction. *Institute for Research in Construction* (2002), internal report IRC-IR-846.
- [15] Bradley, J. S.: The Acoustical design of conventional open plan offices. *Canadian Acoustics* 31(2) (2003), 23-31.
- [16] Jahncke, H., Hongisto, V., Virjonen, P.: Cognitive performance during irrelevant speech: Effects of speech intelligibility and office-task characteristics. *Applied Acoustics* 74 (2013), 307-316.
- [17] Ebissou, A., Parizet, E., Chevret, P.: Use of the speech transmission index for the assessment of sound annoyance in open-plan offices. *Applied Acoustics* 88 (2015), 90-95.
- [18] Liebl, A., Assfalg, A., Schlittmeier, S. J.: The effects of speech intelligibility and temporal-spectral variability on performance and annoyance ratings. *Applied Acoustics* 110 (2016), 170-175.
- [19] Mapp, P.: New techniques for measuring speech privacy & efficiency of sound masking systems. 122nd Convention of the Audio Engineering Society, Wien, Österreich, 2007.