

**Universität Für Musik Und Darstellende Kunst Graz**

**Institut für Elektronische Musik und Akustik**

O.Univ.Prof. Mag. Dipl.-Ing. Dr. Robert Höldrich

---

durchgeführt am

---

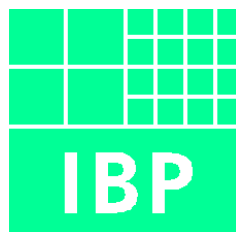
**Fraunhofer-Institut für Bauphysik Stuttgart**

**Abteilung Raumakustik / Technische Akustik**

Prof. Dr.-Ing. Helmut Fuchs

**Diplomarbeit**

**Einfluss tieffrequenter Störgeräusche  
auf die Sprachverständlichkeit**



**Constanze Rau**

01. August 2003 bis 31. Januar 2004

## Zusammenfassung

Die Gestaltung von Räumen gewinnt im breiten Feld der Akustik immer größere Bedeutung. Dabei ist es jedoch sehr umstritten, in welchem Frequenzbereich diese Planungen durchgeführt werden müssen, um eine optimale Sprachverständlichkeit zu erzielen. Während häufig die Meinung vertreten wird, Frequenzen unterhalb 200Hz spielen in geschlossenen Räumen für die Verständlichkeit keine Rolle, gibt es auch Stimmen, dass genau diese vernachlässigten tiefen Frequenzen extrem stören und entsprechend bedämpft werden müssen.

Diese Arbeit sollte deshalb klären, welchen Einfluss tiefe Frequenzen tatsächlich auf die Sprachverständlichkeit in Büros, in Konferenzräumen usw. haben und in welchem Maß dies auch mit einfachen objektiven Messungen vorausgesagt werden kann. Dafür wurden zunächst verschiedene Mithörschwellen ermittelt. Im Anschluss daran wurden Logatomtests durchgeführt, um den Einfluss von unterschiedlichen Nachhallzeiten und Störgeräuschen auf die Sprachverständlichkeit zu bestimmen. Diese Ergebnisse wurden mit den, aus der objektiv gemessenen Impulsantwort berechneten, raumakustischen Gütemaße verglichen.

Mit den gewonnenen Erkenntnissen sollen in Zukunft optimierte raumakustische Planungskonzepte erstellt werden können.

## Abstract

The improvement of speech intelligibility has become an important part in the wide field of acoustics, especially in acoustical designing of rooms. In opposition to conventional ideas some experts hold the view that attenuation of low frequencies below 200Hz improves speech intelligibility significantly.

In this diploma theses the real influence of low frequencies to speech intelligibility is investigated. Therefore, various masked thresholds are determined to measure the effect of simultaneous masking quantitatively. Furthermore, tests with logatoms are conducted to show the influence of reverberation and background noise on speech intelligibility. In a second step it should be analysed if its possible to predict the degree of speech intelligibility by standard measurements (calculated from the impulse response). In future this results should support the optimal acoustical room designing.

---

## **Danksagung**

An dieser Stelle möchte ich mich bei allen bedanken, die mir bei der Erstellung dieser Arbeit behilflich waren!

Dies gilt an erster Stelle Herrn Prof. Helmut Fuchs und Herrn Prof. Dr. Robert Höldrich, die diese Arbeit in Stuttgart ermöglicht und mich bei der Durchführung sehr gut betreut haben.

Weiters möchte ich mich besonders bei Frank Zickmantel und bei Karl-Heinz Bay bedanken, die mir jederzeit tatkräftig zur Seite standen und die mich immer bei der Lösung zahlreich auftretender Probleme unterstützt haben. Danke auch an Robert Hanisch und an Alexander Jauer, die mir bei den Umbauten in den Messräumen und beim Schleppen der schweren Absorberplatten geholfen haben.

Ein weiterer Dank gilt allen Kollegen, die bei den Hörversuchen teilgenommen haben und dadurch teilweise fürchterliche Geräusche über sich ergehen lassen mussten. Dem ganzen Team am IBP danke ich für die angenehme Atmosphäre und die ständige Hilfsbereitschaft!

---

# Inhaltsverzeichnis

<b>1. EINLEITUNG .....</b>	<b>1</b>
<b>1.1. ALLGEMEINES .....</b>	<b>1</b>
<b>1.2. AUFGABENSTELLUNG .....</b>	<b>1</b>
<b>2. PHYSIOLOGISCHE GRUNDLAGEN .....</b>	<b>3</b>
<b>2.1. AUFBAU DES PERIPHEREN GEHÖRS .....</b>	<b>3</b>
2.1.1. DAS AUBENOHR .....	3
2.1.2. DAS MITTELOHR .....	4
2.1.3. DAS INNENOHR .....	4
<b>2.2. ZENTRALE SCHALLVERARBEITUNG .....</b>	<b>6</b>
<b>2.3. LAUTSTÄRKEPEGEL <math>L_S</math>, ANPASSUNG UND COCKTAILPARTY-EFFEKT .....</b>	<b>7</b>
2.3.1. DAS HÖRFELD .....	7
2.3.2. KURVEN GLEICHER LAUTSTÄRKEPEGEL .....	7
2.3.3. AUFTRETEN UND FOLGEN DER ANPASSUNG .....	8
2.3.4. DER COCKTAIL-PARTY-EFFEKT .....	9
<b>3. THEORETISCHE GRUNDLAGEN DER VERDECKUNG .....</b>	<b>10</b>
<b>3.1. VERDECKUNG IM FREQUENZBEREICH – SIMULTANVERDECKUNG .....</b>	<b>11</b>
<b>3.2. VERDECKUNG IM ZEITBEREICH– VOR- BZW. NACHVERDECKUNG .....</b>	<b>13</b>
3.2.1. DIE NACHVERDECKUNG .....	13
3.2.2. DIE VORVERDECKUNG .....	14
<b>3.3. MITHÖRSCHWELLEN-PERIODENMUSTER .....</b>	<b>16</b>
<b>3.4. MITHÖRSCHWELLEN-ZEITMUSTER .....</b>	<b>17</b>
<b>3.5. FREQUENZGRUPPEN UND TONHEIT <math>Z</math> .....</b>	<b>17</b>
<b>3.6. VERDECKUNG DURCH ZUFALLSRAUSCHEN .....</b>	<b>19</b>
3.6.1. SCHALLDRUCKPEGEL UND DICHTPEGEL .....	19
3.6.2. BREITBANDRAUSCHEN .....	21
3.6.3. SCHMALBANDRAUSCHEN .....	22
3.6.4. TIEFPASS- UND HOHPASSRAUSCHEN .....	23
<b>3.7. VERDECKUNG DURCH SINUSTÖNE .....</b>	<b>24</b>
<b>3.8. VERDECKUNG DURCH KLÄNGE .....</b>	<b>25</b>
<b>3.9. VERDECKUNG DURCH LOW-NOISE NOISE .....</b>	<b>26</b>
3.9.1. DEFINITION VON LOW-NOISE NOISE .....	26
3.9.2. ERZEUGUNG VON LOW-NOISE NOISE .....	27
3.9.3. VERDECKUNGSEFFEKTE BEI VERWENDUNG VON LNN ALS MASKIERER .....	31

<b>4. THEORETISCHE GRUNDLAGEN DER SPRACHVERSTÄNDLICHKEIT .....</b>	<b>32</b>
<b>4.1. AKUSTISCHE EIGENSCHAFTEN VON SPRACHE.....</b>	<b>32</b>
4.1.1 ALLGEMEINES .....	32
4.1.2. DAS SPRACHSPEKTRUM.....	33
4.1.3. SPRACHSCHALLPEGEL .....	34
4.1.4. DYNAMIKVERLAUF.....	35
<b>4.2. MESSUNG DER SPRACHVERSTÄNDLICHKEIT .....</b>	<b>36</b>
4.2.1. HÖRSCHWELLE VON VERBALEN REIZEN .....	36
4.2.2. DEFINITIONEN .....	37
4.2.3. SPRACHTESTS MIT SILBEN, WÖRTERN UND SÄTZEN .....	38
<b>4.3. BERECHNUNG DER SPRACHVERSTÄNDLICHKEIT DURCH PHYSIKALISCHE PARAMETER .....</b>	<b>42</b>
4.3.1. DEUTLICHKEITSGRAD NACH THIELE $D_{50}$ .....	42
4.3.2. DEUTLICHKEITSMAB $C_{50}$ .....	42
4.3.3. SCHWERPUNKTZEIT $TS$ (CENTER TIME $CT$ ) .....	43
4.3.4. OBJEKTIVE BESTIMMUNG DER SILBENVERSTÄNDLICHKEIT $V$ .....	44
4.3.5. ARTIKULATIONSVERLUST BZW. KONSONANTENVERLUST $AL_{CONS}$ .....	45
4.3.6. DER ARTIKULATIONSINDEX $AI$ .....	47
4.3.7. SPRACHÜBERTRAGUNGSINDEX $STI$ UND $RASTI$ .....	49
<b>5. VERSUCHSPLANUNG .....</b>	<b>51</b>
<b>5.1. AUSGANGSLAGE.....</b>	<b>51</b>
<b>5.2. AUFBAU FÜR DIE HÖRSCHWELLENMESSUNGEN .....</b>	<b>52</b>
5.2.1. ANFORDERUNGEN AN DIE MESSUMGEBUNG .....	52
5.2.2. MESSUMGEBUNG AUDIOKABINE .....	53
5.2.3. VERSTÄRKER UND LAUTSPRECHER .....	54
5.2.4. COMPUTER ZUR SOFTWARE-STEUERUNG .....	54
5.2.5. MESSGERÄTE.....	55
<b>5.3. MESS-SIGNALE FÜR DIE HÖRSCHWELLENTESTS.....</b>	<b>56</b>
5.3.1. NUTZSIGNALE.....	56
5.3.2. STÖRSIGNALE .....	57
<b>5.4. MESSAUFBAU FÜR DIE VERSTÄNDLICHKEITSTESTS .....</b>	<b>62</b>
5.4.1. MESSUMGEBUNG TIEFFREQUENZ-LABOR.....	62
5.4.2. SIGNALDARBIETUNG.....	63
5.4.3. GEPLANTE RAUMAUSKLEIDUNGEN .....	64
<b>5.5. TEST-SIGNALE FÜR VERSTÄNDLICHKEITSUNTERSUCHUNGEN .....</b>	<b>65</b>
5.3.1. LOGATOMLISTEN .....	65
5.3.2. STÖRSIGNALE .....	65
<b>6. MESS-SOFTWARE EARTECT2003 .....</b>	<b>66</b>
<b>6.1. EIGENSCHAFTEN DES VORHANDENEN PROGRAMMS .....</b>	<b>66</b>
6.1.1. ALLGEMEINE DATEN .....	66
6.1.2. ZEITLICHER ABLAUF DES MESSVERFAHRENS .....	66
<b>6.2. NOTWENDIGE ÄNDERUNGEN AN DER SOFTWARE.....</b>	<b>69</b>
<b>6.3. EIGENSCHAFTEN DES ADAPTIERTEN PROGRAMMS .....</b>	<b>70</b>

---

<b>7. VERSUCHSDURCHFÜHRUNG HÖRSCHWELLENMESSUNG</b> .....	<b>72</b>
<b>7.1. VORBEREITUNGEN FÜR HÖRSCHWELLENVERSUCHE</b> .....	<b>72</b>
7.1.1. MESSUMGEBUNG AUDIOKABINE .....	72
7.1.2. TESTMESSUNGEN IM HALBFREIFELDDRAUM.....	76
<b>7.2. ERGEBNISSE DER MESSUNGEN MIT KONVENTIONELLEN STÖRSIGNALEN</b> .....	<b>81</b>
7.2.1. STÖRSIGNAL SINUSTON .....	81
7.2.2. STÖRSIGNAL TERZRAUSCHEN.....	82
7.2.3. STÖRSIGNAL TIEFPASSRAUSCHEN .....	84
7.2.4. STÖRSIGNAL TONGEMISCH .....	87
7.2.5. ALLGEMEINE ERKENNTNISSE AUS DEN HÖRVERSUCHEN MIT ZUFALLSRAUSCHEN.....	88
<b>7.3. MESSUNGEN MIT LOW-NOISE NOISE</b> .....	<b>90</b>
7.3.1. STÖRSIGNAL TERZRAUSCHEN.....	90
7.3.2. STÖRSIGNAL TIEFPASSRAUSCHEN .....	91
7.3.3. ZUSAMMENFASSUNG DER ERGEBNISSE UND FAZIT.....	93
<b>8. VERSUCHE ZUR ERMITTLUNG DER SPRACHVERSTÄNDLICHKEIT</b> .....	<b>94</b>
<b>8.1. MESSUMGEBUNG TIEFFREQUENZ-LABOR</b> .....	<b>94</b>
8.1.1. DIE NACHHALLZEIT.....	94
8.1.2. DIE TRANSFERFUNKTION.....	95
8.1.3. DIE RAUMIMPULSANTWORT .....	96
<b>8.2. DURCHFÜHRUNG DES LOGATOMTESTS</b> .....	<b>99</b>
8.2.1. AUSWIRKUNGEN DER STÖRGERÄUSCHE .....	100
8.2.2. AUSWIRKUNGEN DER RAUMZUSTÄNDE .....	103
<b>8.3. VERGLEICH DER SUBJEKTIVEN ERGEBNISSE MIT BERECHNETEN WERTEN</b> .....	<b>106</b>
8.3.1. SCHWERPUNKTZEIT $T_S$ .....	107
8.3.2. DEUTLICHKEITSMAB $C_{50}$ .....	107
8.3.3. ARTIKULATIONSINDEX AI UND SILBENVERSTÄNDLICHKEITSAKTOR $V_{S/N}$ .....	108
8.3.4. OBJEKTIVE GÜTEMAßE: $v$ , $D_{50}$ , $1-AL_{CONS}$ , STI UND RASTI.....	109
<b>8.4. ZUSAMMENFASSUNG DER ERGEBNISSE</b> .....	<b>111</b>
<b>9. SCHLUSSFOLGERUNG</b> .....	<b>113</b>
<b>9.1. HÖRSCHWELLENVERSUCHE</b> .....	<b>113</b>
<b>9.2. UNTERSUCHUNGEN DER SPRACHVERSTÄNDLICHKEIT MIT LOGATOMTESTS</b> .....	<b>114</b>
<b>ANHANG</b> .....	<b>118</b>
<b>ZAHLENWERTE DER BERECHNETEN RAUMAKUSTISCHEN GÜTEMAßE</b> .....	<b>118</b>
<b>HÖRSCHWELLEN DER TESTPERSONEN</b> .....	<b>120</b>
<b>LOGATOMLISTEN</b> .....	<b>121</b>
<b>LITERATURVERZEICHNIS</b> .....	<b>123</b>

---

# 1. Einleitung

## 1.1. Allgemeines

Im großen Feld der Akustik gewinnt der Bereich der Raumakustik immer mehr an Bedeutung. Firmen bemühen sich verstärkt um angenehmere Arbeitsbedingungen für ihre Mitarbeiter, öffentliche Bauten werden akustisch verbessert, gutes Hören und Verstehen von Sprache wird immer häufiger ein wichtiger Bestandteil bei der Bautenplanung.

Um diese Ziele zu erreichen, werden bei den raumakustischen Planungskonzepten unterschiedlichste Arten von Absorbern eingesetzt. Je nach Anforderung (Frequenzbereich, Preis, Lichtdurchlässigkeit, Aussehen usw.) können optimierte Materialien verwendet werden.

Problematisch ist dabei jedoch immer noch, welche exakten Anforderungen nun tatsächlich an die Absorber gestellt werden sollen und müssen.

Während im Allgemeinen die Meinung vertreten wird, Frequenzen unterhalb 200Hz spielen für die Sprachverständlichkeit in geschlossenen Räumen praktisch keine Rolle, ist man am Fraunhofer-Institut für Bauphysik (IBP) in Stuttgart anderer Meinung: gerade tiefe Frequenzen seien besonders zu beachten, da sie die Sprachverständlichkeit stören [1, 2].

Um in diesem strittigen Punkt Klarheit zu erlangen, sollte in dieser Diplomarbeit der Einfluss tiefer Frequenzen auf die Sprachverständlichkeit untersucht werden. Mit den daraus gewonnenen Erkenntnissen können in Zukunft optimierte raumakustische Planungskonzepte erstellt werden.

## 1.2. Aufgabenstellung

In dieser Diplomarbeit sollte untersucht werden, wie stark sich tieffrequente Störungen (Frequenzen unterhalb 500Hz) auf die menschliche Kommunikation auswirken.

Dazu wurden zunächst von mehreren Probanden die Ruhehörschwellen gemessen. Dadurch konnte sichergestellt werden, dass nur Personen mit normalem Gesundheitszustand, sogenannte otologisch normale Personen [3], an den Hörversuchen teilnahmen. Mit diesen Testpersonen wurden dann Mithörschwellen bei unterschiedlichen, tieffrequenten Störgeräuschen ermittelt. Die daraus erhaltenen Ergebnisse wurden mit bekannten Resultaten aus der Literatur verglichen.

Zusätzlich wurden Sprachverständlichkeits-Untersuchungen mit Logatomtests in einer definierten Umgebung durchgeführt. Diese Resultate konnten Berechnungen aus verschiedenen

raumakustischen Gütemaßen gegenübergestellt und evtl. auftretende Unterschiede diskutiert werden.

Dazu wurde ein vorhandenes Audiometrie-Messsystem – bestehend aus einer reflexionsarmen Kabine, einer Mess-Software und den dazugehörigen elektroakustischen Komponenten – aufgebaut. Zusätzlich wurden Verbesserungsmöglichkeiten des Audiometrie-Systems miteinbezogen (Verwendung der 24-Bit Technologie, Einbau eines Koaxial-Lautsprechers, Überprüfung alternativer Sitzpositionen in der Audiokabine).

Als Testsignal wurden Sinustöne, Terzrauschen, tiefpassgefiltertes Weißes Rauschen und Low-noise noise (LNN) verwendet. Diese Signale wurden mit entsprechender Software (Samplitude und Cool Edit Pro) vollständig digital generiert.



## 2. Physiologische Grundlagen

Das menschliche Ohr lässt sich in zwei Bereiche gliedern. Den ersten Teil bildet das periphere Gehör, das in Äußeres Ohr, Mittelohr und Innenohr unterteilt werden kann. Im zweiten Teil des Hörorgans findet die neuronale Verarbeitung von Aktionspotentialen statt, die in höheren Ebenen schließlich zu Empfindungen führen.

### 2.1. Aufbau des peripheren Gehörs

Das periphere Gehör lässt sich in drei Bereiche unterteilen: Außenohr, Mittelohr und Innenohr. Diese drei Stufen lassen sich sowohl funktionell als auch anatomisch voneinander unterscheiden.

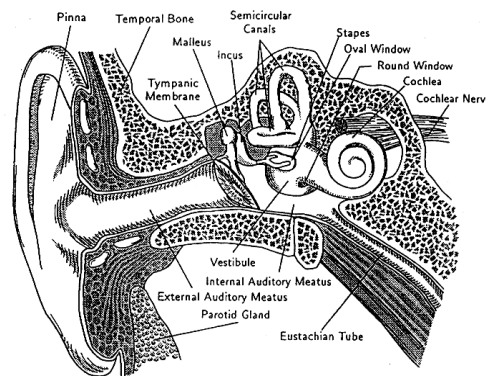


Abbildung 1: Aufbau des peripheren Gehörs [4]

#### 2.1.1. Das Außenohr

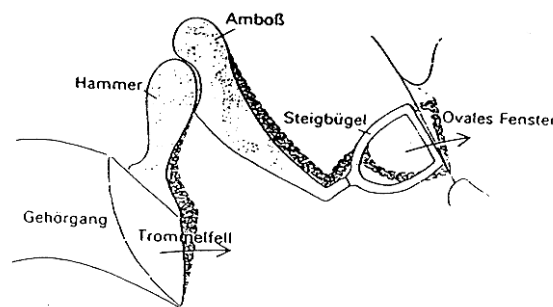
Das Außenohr besteht aus der Ohrmuschel und dem Gehörgang und schließt mit dem Trommelfell ab.

Die Ohrmuschel fängt den Schall wie ein Schalltrichter aus verschiedenen Richtungen auf und leitet ihn über den Gehörgang weiter zum Trommelfell. Sie verzerrt die ankommenden Schallsignale in Abhängigkeit von Schallquellenrichtung und –entfernung unterschiedlich und nimmt eine Transformation räumlicher Schallfeldmerkmale in zeitliche vor. Hierin besteht ihre Bedeutung für das räumliche Hören.

Die akustische Wirksamkeit der Ohrmuschel beruht auf unterschiedlichen physikalischen Phänomenen wie Reflexion, Abschattung, Streuung, Beugung, Interferenz und Resonanz. So ist neben der Lokalisation in der Vertikalebene auch eine Unterscheidung zwischen Vorne und Hinten möglich. Damit stellt sie in ihrer akustischen Funktion ein lineares Filter dar, dessen Übertragungsfunktion von der Richtung und der Entfernung der Schallquelle abhängt [4].

### 2.1.2. Das Mittelohr

Das Mittelohr liegt in einer kleinen Höhle des Schädelknochens. Es erstreckt sich vom Trommelfell bis zum ovalen Fenster der Schnecke und dient als Bindeglied zwischen Außen- und Innenohr. Seine wesentlichen funktionalen Bestandteile sind drei Mittelohrknöchelchen, die wegen ihres Aussehens als Hammer, Amboss und Steigbügel bezeichnet werden.



**Abbildung 2: Aufbau des Mittelohrs [5]**

Im Mittelohr werden zwei wesentliche Mechanismen durchgeführt: Hebelübersetzung und Flächentransformation.

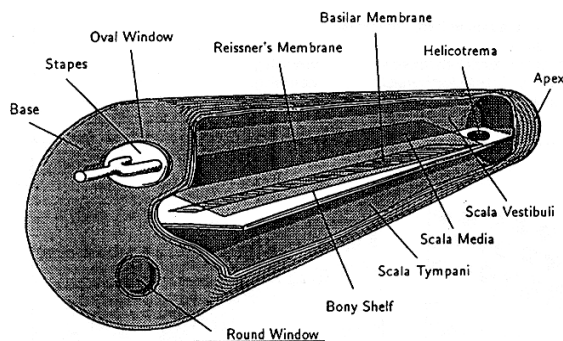
Die Hebelübersetzung findet zwischen dem Trommelfell und dem ovalen Fenster statt. Sie ist für die Impedanzanpassung zwischen den unterschiedlichen Schallwellenimpedanzen von Luft (Gehörgang) und Lymphflüssigkeit (Innenohr) notwendig. Am Trommelfell finden, ausgelöst durch Luftdruckänderungen, also Schallereignissen, bei kleinem Druck relativ große Auslenkungen statt. Diese werden in kleine Amplituden bei sehr hohem Druck am ovalen Fenster und in Flüssigkeitsdruck in der Schnecke umgewandelt. Bei der Flächentransformation wird die große Fläche des Trommelfells auf die kleine Fläche des ovalen Fensters „umgewandelt“. Dadurch werden Schalldrücke, die am Trommelfell wirken, am ovalen Fenster ca. 22-fach verstärkt [4].

### 2.1.3. Das Innenohr

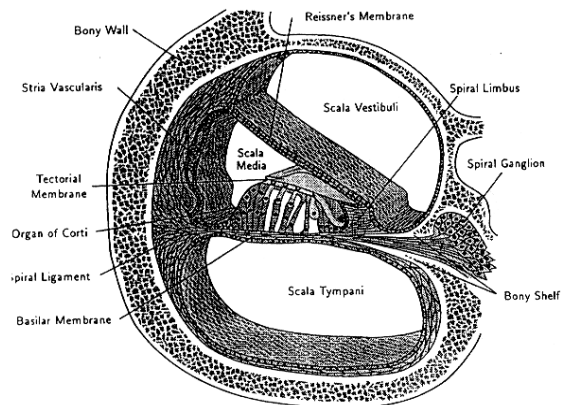
Das Innenohr ist in das Felsenbein, einen sehr harten Knochen, eingebettet. Es besteht aus zwei Teilen: dem Gleichgewichtsorgan und der Schnecke (Cochlea). Die Schneckenform dient dabei primär zur Platzersparnis.

Die Funktion des Innenohres ist von der Funktionsweise des Mittelohres im Wesentlichen unabhängig: auf der Membran des ovalen Fensters sitzt die Fußplatte des Steigbügels auf; sie überträgt die Schallschwingungen auf die Innenohrlymphe. Diese Schwingungen des ovalen

Fensters führen zu einer Volumenverschiebung der inkompressiblen Perilymphe in den Scala tympani und vestibuli.

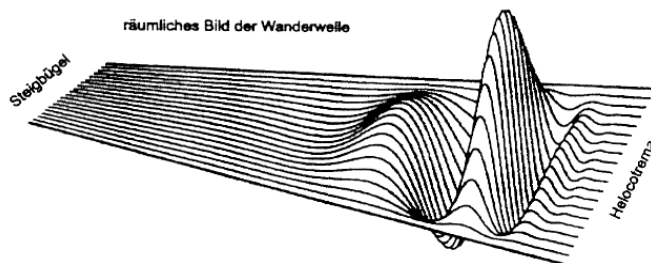


**Abbildung 3:**  
Schnecke in abgewickelter Form [4]

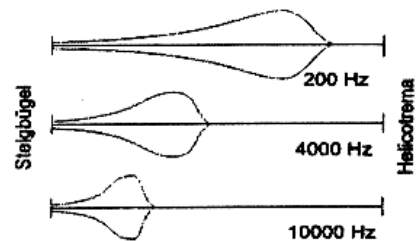


**Abbildung 4:**  
Querschnitt durch den Schnecken gang [4]

Da die Wände elastisch sind, geben sie der wellenförmigen Volumenverschiebung nach. Es bilden sich Wanderwellen aus. Die angeregte Druckverschiebung erreicht das runde Fenster und lenkt dieses aus. Die Hüllkurve dieser Wanderwellen nimmt bei einer sinusförmigen Erregung bis zu einem gewissen Maximum zu und verebbt dann sehr schnell.



**Abbildung 5:** Räumliche Wanderwelle [6]

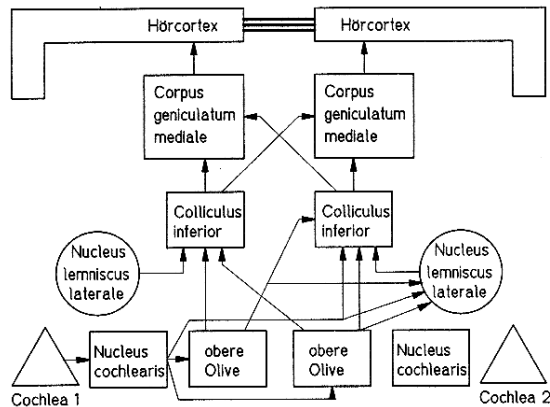


**Abbildung 6:** Maximum der Wanderwelle in Abhängigkeit der Frequenz [6]

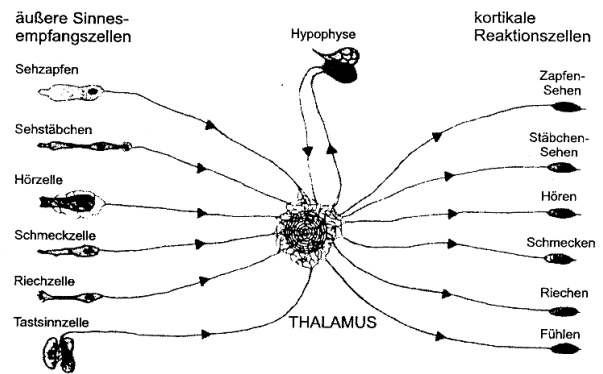
Je höher die angeregte Frequenz ist, desto näher liegt die maximale Auslenkung beim Steigbügel (ovales Fenster), d.h. tiefe Frequenzen erzeugen Wanderwellen, die weiter in die Cochlea vordringen [4].

## 2.2. Zentrale Schallverarbeitung

Eine schematische Abbildung des auditiven Systems, die sogenannte Hörbahn (Abb. 7) belegt die Komplexität der zentralen Schallverarbeitung (der Übersichtlichkeit halber sind nur die aufsteigenden Bahnen eines Ohres (Cochlea 1) eingezeichnet).



**Abbildung 7: Zentrale Verarbeitungsstationen [6]**



**Abbildung 8: Nervenverbindungen von Sinnesorganen, Thalamus und Großhirnrinde [6]**

Die Hörbahn besteht unterhalb der Großhirnrinde (Cortex) aus fünf anatomisch und physiologisch unterscheidbaren Verarbeitungsebenen, die Kerne genannt werden. Diese Kerne, welche jeweils aus einer großen Zahl von Zellen bestehen, sind paarweise vorhanden. Zwischen ihnen bestehen die in der Abbildung angedeuteten Nervenverbindungen.

Über die Nervenverbindungen werden "Feuerimpulse" bis zur Hörrinde, dem eigentlichen Hörzentrum des Gehirns (Thalamus, siehe Abb. 8) weitergeleitet. Dort treten sie als Hörereignis ins Bewusstsein.

Die Cochlea ist damit ein perfekter Schall-Analysator, der auch einzelne Frequenzen in einem komplexen Klang aussortieren kann. Vergleicht man die Cochlea mit einem elektrischen System, so zeigt es sich, dass ihre Eigenschaften bis heute nicht erreicht werden: Das Ohr braucht nur ein paar Perioden einer Schwingung, bis es die Tonhöhe festgestellt hat. Außerdem kann es rund 1500 verschiedene Tonhöhen innerhalb des Hörbereichs erkennen und unterscheiden [4].

## 2.3. Lautstärkepegel $L_S$ , Anpassung und Cocktailparty-Effekt

### 2.3.1. Das Hörfeld

Den Bereich, in dem ein Schallereignis ein Hörereignis auslöst, nennt man Hörfeld, Hörfläche bzw. Hörbereich. Das Hörfeld ist bei kleinen Schalldruckpegeln durch die Hörschwelle, bei großen Schalldruckpegeln durch die Schmerzschwelle begrenzt. Nur Schallereignisse mit Frequenzen zwischen 16Hz und maximal 20kHz rufen Hörereignisse hervor. Mit dem Alter sinkt die obere Hörgrenze im Mittel um 1 kHz pro Lebensjahrzehnt [7].

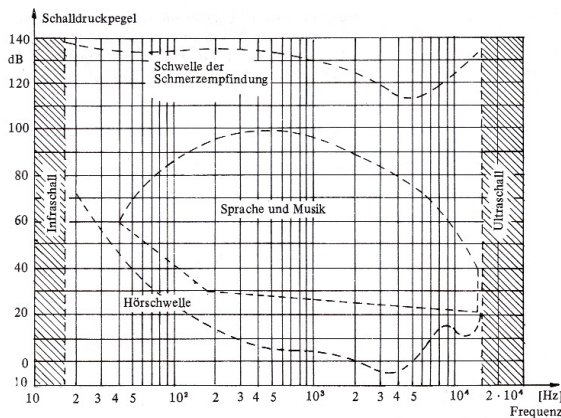


Abbildung 9: Menschliches Hörfeld mit Sprach- und Musikbereich [7]

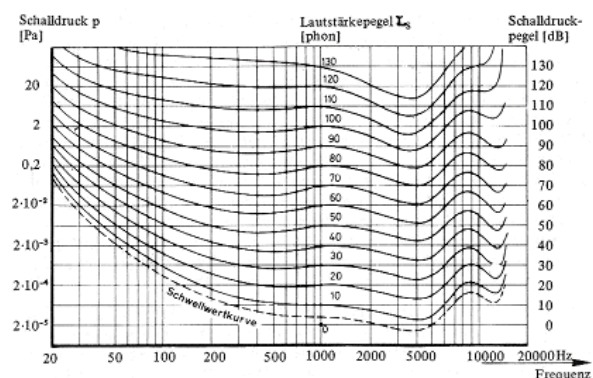


Abbildung 10: Hörschwelle und Kurven gleicher Lautstärkepegel [7]

### 2.3.2. Kurven gleicher Lautstärkepegel

Durchläuft ein Sinuston mit einem konstanten Schalldruckpegel den gesamten hörbaren Frequenzbereich, so bleibt der Ton keineswegs gleich laut.

Er wird mit steigender Frequenz zunächst lauter, ab ca. 4000 Hz schließlich wieder leiser. Um diese Feststellung für verschiedene Schalldruckpegel genau zu erfassen, hat man die Kurven gleicher Lautstärkepegel ermittelt. Sie geben in Abhängigkeit von der Frequenz den Schalldruckpegel  $L_p$  an, der die jeweils gleiche Lautstärkeempfindung hervorruft. Man ordnet jeder Kurve einen bestimmten Lautstärkepegel zu, der in Phon angegeben wird. Für 1 kHz stimmen Schalldruckpegel in dB und Lautstärkepegel in Phon zahlenmäßig überein.

Abbildung 10 zeigt die nach DIN 1318 und 45 630 genormten Kurven, die mit den internationalen ISO-Empfehlungen (R 226 und R 454) übereinstimmen. Die DIN-Kurven beziehen sich auf Personen mit normalem Gehör zwischen 18 und 25 Jahren bei beidohrigem Hören im freien Schallfeld [7].

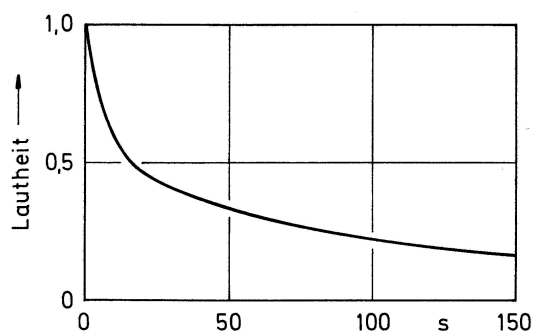
Der Grund dafür, dass sich die Hörschwelle – in Abbildung 10 gestrichelt dargestellt – bei 4 Phon befindet und nicht bei 0 Phon, liegt darin, dass als Bezugsschalldruck der runde Wert 20 Pa international vereinbart wurde. Für andere Schallsignale, z.B. Bandrauschen, oder für andere Abhörbedingungen wie z.B. Diffusfeld, ergeben sich etwas abweichende Normalkurven.

### 2.3.3. Auftreten und Folgen der Anpassung

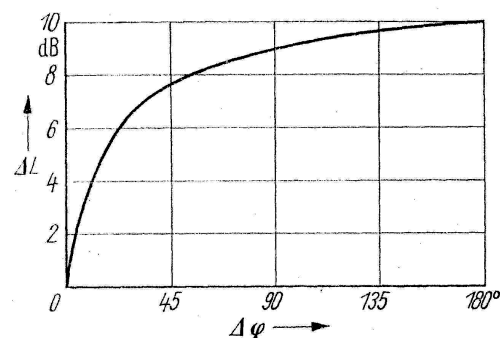
Eine wichtige Eigenschaft des menschlichen Gehörs ist die Fähigkeit, seine Empfindlichkeit einem bestimmten, gerade herrschenden, mittleren Schalldruckpegel  $L_{eq}$  anzupassen. Dadurch werden z.B. gleichmäßige Hintergrundgeräusche im Höreindruck stark zurückgedrängt.

Das Gehör bildet sich aus den einwirkenden Reizen ein Bezugssystem (Anpassungsniveau), an dem sich die einzelnen Urteile wie laut – leise, hell – dunkel oder hoch – tief als Mittelwert orientieren [7].

Der Betrag, um den die Hörempfindlichkeit absinkt, ist abhängig von der Frequenzzusammensetzung und des Pegels des Signals, sowie von der Darbietungsdauer:



**Abbildung 11: Zeitliches Hörempfinden eines Dauerschalls [8]**



**Abbildung 12: Mögliche Störpegelerhöhung ohne Verständlichkeitsverlust als Funktion der Seitenwindeldifferenz zwischen Sprecher und Störung [9]**

Wie in Graphik 11 erkennbar, setzt der Adaptionvorgang rasch nach Beginn des Dauersignals ein. In dieser Anfangsphase der Signaldarbietung zeigt sich eine sehr starke Lautheitsabnahme, dessen Maximum nach etwa drei bis fünf Minuten erreicht wird. Bei noch längerer Signaleinwirkung verändert sich das Lautheitsempfinden nicht mehr weiter, ein konstanter Wert wird beibehalten. Die Redaptation (Rückbildung der ursprünglichen Empfindlichkeit des Ohres) erfolgt in etwa ein bis zwei Minuten [8].

Die Anpassung hat damit zur Folge, dass sich das Ohr an ein Schallereignis gewöhnt und dadurch ein gleichmäßiger Dauerton mit der Zeit immer leiser erscheint: das Gehör ermüdet und ordnet diesen Dauerton als unwichtiges Hintergrundgeräusch ein [7] (siehe auch 5.3.2 *Stör-signale*).

#### 2.3.4. Der Cocktail-Party-Effekt

In diesem Zusammenhang muss noch ein interessantes Hörphänomen erwähnt werden.

Das menschliche Gehör ist in der Lage, aus einer Vielzahl von überlagerten Schallsignalen eine einzelne, willkürlich ausgewählte Quelle gezielt herauszufiltern.

Dieses Phänomen spielt im täglichen Leben eine wichtige Rolle, z.B. wenn sich in einem Raum zahlreiche Personen befinden und sich miteinander unterhalten. Ein Zuhörer ist nun trotz des Stimmengewirrs in der Lage, bei binauralem Hören einen einzelnen Sprecher aus einer Sprechergruppe herauszuhören, ohne sich dem Redner zuwenden zu müssen [8].

Bei einohrigem (monauralem) Hören geht diese Fähigkeit erheblich zurück.

Ursache für dieses Phänomen ist, dass das Gehör Hörereignisse, die aus unterschiedlichen Richtungen kommen, besser trennen und verfolgen kann, als wenn sie aus einer gemeinsamen Richtung eintreffen. Ein Nutzsignal, das aus einer bestimmten Richtung ankommt, wird deshalb von einem Störsignal mit anderem Einfallswinkel bei zweiohrigem Hören weniger stark verdeckt als bei einohrigem Hören [7] (siehe 5.4.1. *Messumgebung Tieffrequenz-Labor*).

Abbildung 12 zeigt, um wieviel der Störschallpegel einer zweiten Quelle beim binauralen Hören höher sein darf, wenn Nutzschall und Störung aus unterschiedlichen Richtungen einfallen und die gleiche Verständlichkeit aufrecht erhalten werden soll wie bei gleichem Schalleinfall [9].

So kann z.B. auf einer Cocktailparty Sprache aus einer bestimmten Richtung (Gespräch mit dem Nachbarn) immer noch verstanden werden, obwohl die Umgebungslautstärke höher liegt als die der Sprache. Daher leitet sich neben der Bezeichnung „Intelligentes Hören“ auch der Name „Cocktail-Party-Effekt“ ab, der von E. C. Cherry 1953 erstmals beschrieben wurde.

### 3. Theoretische Grundlagen der Verdeckung

Unter Verdeckung versteht man den Effekt, dass ein Testschall durch einen Störschall nicht mehr oder nur mit verminderter Lautstärke wahrgenommen werden kann.

Dieser Effekt ist aus dem Alltag gut bekannt: Bei einer Unterhaltung mit einem Partner in ruhiger Umgebung ist keine große Lautstärke notwendig. Tritt jedoch plötzlich ein lauter Störschall auf (z.B. ein vorbeifahrender Zug), so kann der Gesprächspartner nicht mehr verstanden werden. Die Person muss jetzt wesentlich lauter als zuvor sprechen um erneut hörbar zu werden, wobei diese Erhöhung des Sprachpegels meist völlig unbewusst durchgeführt wird. Dieses Verhalten wird auch als Lombard-Effekt bezeichnet [9].

Um die wichtigen Effekte der Verdeckung, die oft auch als Maskierung bezeichnet wird, quantitativ zu beschreiben, wird die sogenannte Mithörschwelle gemessen. Diese gibt denjenigen Schalldruckpegel eines Testschalls an, den dieser besitzen muss, um neben einem Störschall (Maskierer) gerade noch wahrgenommen zu werden.

In der Psychoakustik werden als Testschalle meistens Sinustöne verwendet, als Störgeräusche kommen Breitbandrauschen, Schmalbandrauschen, Sinustöne und Tongemische in Frage [10] (siehe auch 5.3.2. *Störsignale*).

Ist der Störschall leise, so ist seine Wirkung auf den Testschall sehr gering und das Lautstärkeempfinden des Testschalls bleibt erhalten. Ist das Störsignal (Maskierer) laut, so wird die Lautstärkeempfindung des Testschalls verringert (unvollständige bzw. teilweise Verdeckung). Wird die Lautstärke des Störschalls jedoch nochmals erhöht, so wird das Nutzsignal schließlich vollständig verdeckt und es ist nicht mehr hörbar (vollständige Verdeckung).

Dieser Übergang von ungestörter Hörbarkeit des Testsignals zur vollständigen Verdeckung verläuft dabei nicht sprunghaft sondern stetig [11].

Der Effekt der Verdeckung, muss dabei als dreidimensionales Phänomen betrachtet werden.

Maskierung tritt sowohl im Frequenzbereich als auch im Zeitbereich parallel zueinander auf.

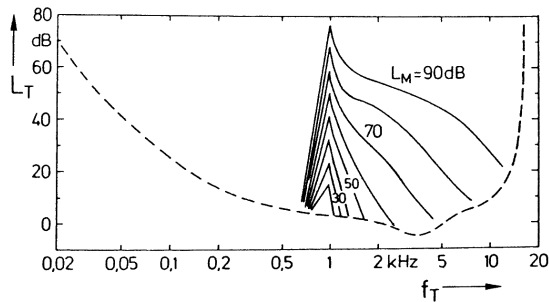
Die jeweils auftretenden Effekte wiederum sind abhängig von der Art der Signale.

Bei lang andauernden Schallereignissen ist die spektrale Verdeckung entscheidend. Bei kurzen Signalen (unterhalb 500ms) gewinnen zusätzlich Effekte der zeitlichen Verdeckung an Bedeutung [10].

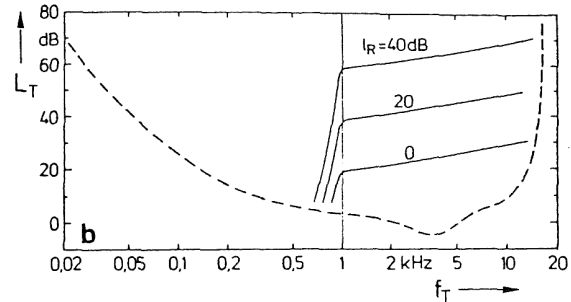


### 3.1. Verdeckung im Frequenzbereich – Simultanverdeckung

Bei der Verdeckung im Frequenzbereich werden diejenigen Effekte betrachtet, die vorkommen, wenn ein Störsignal zeitgleich mit dem Nutzsignal auftritt. Dazu wird auf der x-Achse die Frequenz, auf der y-Achse die Amplitude aufgetragen.



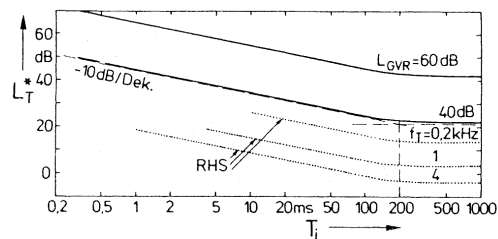
**Abbildung 13:**  
Sinusstörungen bei 1kHz [11]



**Abbildung 14:** Maskierung durch  
Hochpassrauschen [11]

Die Ruheschwelle und auch die Mithörschwelle von Testschallimpulsen hängen dabei stark von der Impulsdauer der Signale ab. Um jedoch zeitliche Effekte der Verdeckung mit kurzen Impulsen messen zu können, muss zunächst der Einfluss der Impulsdauer der Testschalle bestimmt werden.

Dazu werden die Mithörschwellen von Testimpulsen als Funktion der Impulsdauer  $T_i$  ermittelt:



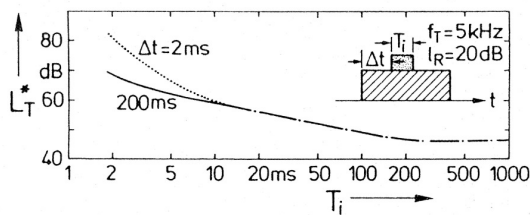
**Abbildung 15:** Tonimpulse [11]

Aufgetragen ist hier Pegel  $L_T^*$  des Dauerschalls, aus dem der Testschall ausgeschnitten ist, als Funktion der Impulsdauer  $T_i$ .

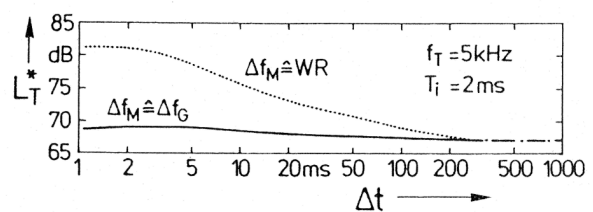
Wie in Abbildung 15 zu sehen, sind die Ruhe- und Mithörschwellen von Tonimpulsen – unabhängig von der Frequenz – stark von ihrer Länge abhängig. Während Impulse mit Dauern von 200ms und länger einen konstanten Wert annehmen, steigt dieser zu kürzeren Impulsen mit ca. 10dB/Dekade an. Dadurch ergibt sich für Testtonimpulse  $\leq 200$ ms ein konstantes Produkt aus Schallintensität und Impulsdauer [11]. Sowohl Ruheschwelle als auch Mithörschwelle zeigen denselben Verlauf.

Beachtet werden muss bei dieser Messung, dass mit einer Verkürzung der Impulsdauer auch gleichzeitig das Spektrum verbreitert wird. Dies begrenzt die Messbarkeit des Verlaufs für kurze Impulsdauern.

Im Bereich der Simultanverdeckung werden noch weitere Effekte beobachtet: Die Mithörschwelle ist abhängig davon, wie lange nach Beginn des Störsignals der Testimpuls dargeboten wird:



**Abbildung 16: Tonimpuls der Dauer  $T_i$   
 $\Delta t$  nach Beginn des Maskierers [11]**



**Abbildung 17: Maskierer  
mit unterschiedlicher Bandbreite [11]**

Wird der Testimpuls ca. 200ms nach dem Einschalten des Maskierers abgespielt, so erhält man den erwarteten Verlauf. Bei einer Verzögerungszeit von lediglich 2ms erhöht sich die Mithörschwelle bei kleinen Dauern jedoch deutlich: ein größerer Pegel des Testschalls ist notwendig, damit er neben dem Störrauschen noch wahrgenommen werden kann [11] (siehe Abb. 16).

Dieser Effekt ist zusätzlich von der spektralen Zusammensetzung der Signale abhängig (siehe Abbildung 17). Gepunktet dargestellt wird die Simultanhörschwelle bei Weißem Rauschen in Abhängigkeit von der Einschaltpause zwischen Stör- und Testsignal. Der durchgezogene dargestellte Verlauf beschreibt den Effekt, wenn die Bandbreite des Maskierers  $\Delta f_M$  auf die Frequenzgruppenbreite  $\Delta f_G$  (siehe Punkt 3.5. *Frequenzgruppen und Tonheit z*) des Testsignals begrenzt ist. Wie gut sichtbar, ist der Unterschied bei kleinen zeitlichen Verschiebungen  $\Delta t$  zwischen dem Beginn des Maskierers und dem Beginn des Testimpulses am deutlichsten. Je später der Testimpuls nach Beginn des Störsignals ertönt, umso geringer wird der Effekt.

Insgesamt kann damit gesagt werden, dass die spektrale Zusammensetzung von Maskierer und Testsignal, sowie deren Übereinstimmung untereinander mit ausschlaggebend ist, so dass es zu individuellen Unterschieden im Bereich von bis zu 10dB kommen kann [11].

### 3.2. Verdeckung im Zeitbereich – Vor- bzw. Nachverdeckung

Neben diesen Effekten, bei denen Testschall und Störsignal gleichzeitig zu hören sind, gibt es auch nichtsimultane Effekte. Um diese jedoch messen zu können, müssen kurze Störgeräusche ( $t \leq 500\text{ms}$ ) und sehr kurze Testtonimpulse ( $t \leq 10\text{ms}$ ) verwendet werden.

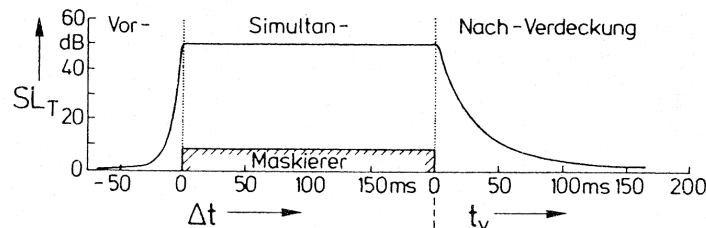


Abbildung 18: Zeitliches Verdeckungsprinzip [11]

Wird der Testschall zeitlich nach dem Störsignal abgespielt, so wird dies als Nachverdeckung bezeichnet. Wird dagegen der Testschall zeitlich vor dem Störschall dargeboten, so wird von Vorverdeckung und der Vorhörschwelle gesprochen.

Die unvollständige Vorverdeckung wird als Folgedrosselung bezeichnet [11].

#### 3.2.1. Die Nachverdeckung

Die Nachverdeckung (engl. Forward Masking → auf der Zeitachse vorwärts) lässt Signale auch nach Abschalten des verdeckenden Signals eine kürzere Zeit lang unhörbar bleiben. Die Zeitspanne der Nachverdeckung hängt sehr von den Signalarten und -dauern ab und liegt im Bereich bis zu 100ms. Erklärt wird die Nachverdeckung durch eine herabgesetzte Empfindlichkeit gerade stimulierter Hörzellen.

Das Gehör weist bei diesem Effekt auch ein spezielles Verhalten auf, das deutlich von dem in der Raumakustik bekannten exponentiellen Verhalten abweicht.

Abbildung 19 zeigt den Abklingvorgang des Gehörs für Pegel zwischen 40dB und 80dB. Die durchgezogene Linie stellt dabei die Messergebnisse dar [11]. Es ist gut sichtbar, dass die Nachhörschwelle ca. 5ms nach dem Abschalten des Maskierers immer noch die Höhe der Simultanhörschwelle besitzt. Danach beginnt die Hörschwelle abzufallen, bis das Gehör nach etwa 200ms auf seinen Ausgangswert (Ruhehörschwelle) zurückgekehrt ist. Bei niedrigeren Schallpegeln ergeben sich dabei flachere Abklingkurven als bei hohen Schallpegeln.

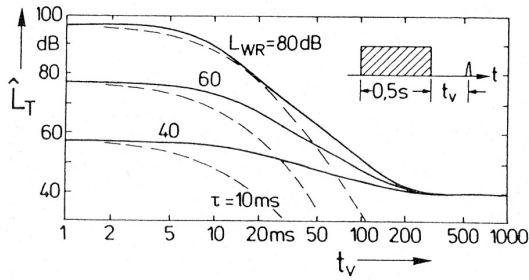


Abbildung 19: Abklingvorgang des Gehörs [11]

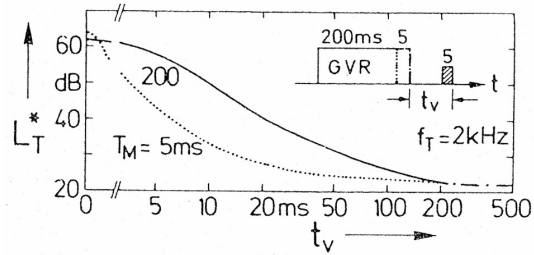


Abbildung 20: Störgeräusch mit 200ms bzw. 5ms Dauer [11]

Der Abklingvorgang des Gehörs kann jedoch nicht mit einem exponentiellen Abklingen (gestrichelte Linie) beschrieben werden. Der Vergleich der beiden Kurven (Messergebnisse und exponentielle Abfallkurve) macht deutlich, dass der Abklingvorgang im Gehör sehr komplex ist [10].

Die Dauer der Nachverdeckung wird außerdem nicht nur vom Schallpegel beeinflusst, sondern auch von der Dauer der Störgeräusche. In Abbildung 20 ist das Abklingverhalten für Störgeräusche von 200ms bzw. von 5ms Dauer dargestellt. Der zeitliche Abklingverlauf ist dabei für die zwei Fälle sehr verschieden: Nach sehr kurzen Störschallimpulsen (gepunktete Linie) erfolgt das Abklingen der psychoakustischen Erregung wesentlich schneller als nach längeren Störungen (durchgezogene Linie). Bereits nach 50ms ist wieder die Ruheshwelle erreicht (ca. 200ms bei langen Störsignalen). Die Nachverdeckung hängt also von der Maskierdauer ab. Dies muss als nichtlinearer Effekt angesehen werden [11].

### 3.2.2. Die Vorverdeckung

Auch eine Vorverdeckung für voreilende Signale (engl. Backward Masking → auf der Zeitachse rückwärts) kann beobachtet werden. Dabei beginnt der Maskierer erst bis zu 20 ms später als das Testsignal, das trotzdem nicht wahrgenommen werden kann. Es wird angenommen, dass die Vorverdeckung erst in höheren Verarbeitungszentren des Gehörs bewirkt wird. Dabei wird bei der neuronalen Verarbeitung der Signale das intensitätsstärkere Signal als wichtiger bewertet und vorrangig bearbeitet [4].

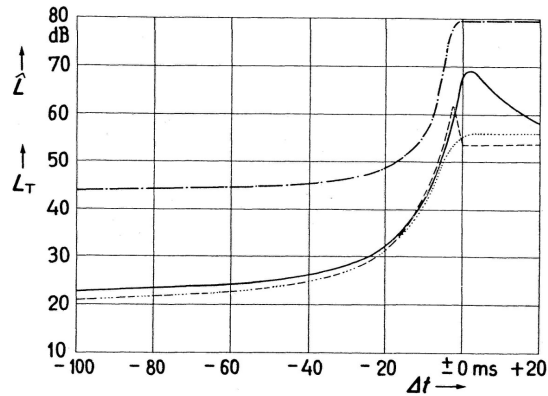


Abbildung 21: Vorhörschwellen für verschiedene Störsignale [12]

In Abbildung 21 sind die Vorhörschwellen und die anschließenden Mithörschwellen für verschiedene Kombinationen von Stör- und Testschall als Funktion der Verzögerungszeit aufgetragen.

Wenn beide Signale ähnliche Spektren haben (punktierte bzw. strichpunktierte Linie) geht die Vorhörschwelle monoton in die Mithörschwelle über.

Geht einem Ton ein kohärenter Testtonimpuls mit gleicher Frequenz und gleicher Phase voraus, so zeigt sich dieselbe Unstetigkeit in der Vorhörschwelle, die bereits bei der Nachhörschwelle aufgetreten ist (gestrichelte Linie). Sie liegt ebenfalls wieder dort, wo die Verzögerungszeit gleich der Dauer des Testimpulses ist. Dadurch verlängert der Testimpuls den Störton lückenlos nach vorne. Dieser Fall ist ein Spezialfall, der auf die Versuchsanordnung zurückzuführen ist.

Wenn jedoch die Spektren der beiden Signale extrem verschieden sind (Tonimpuls wird durch Weißes Rauschen verdeckt) kann ein neuer Effekt beobachtet werden (durchgezogene Linie): Die Vorhörschwelle steigt zunächst weit über die Mithörschwelle an, erreicht dann kurz nach Beginn des Störrauschens ein Maximum und fällt dann innerhalb der ersten 25ms des Störrauschens auf die Mithörschwelle ab [12].

Im Mittel liegt der Übergang von der Ruhehörschwelle zur simultanen Mithörschwelle bei ca. 10ms bis 20ms. Bei der Nachverdeckung liegt die mittlere Dauer, bis wieder die Ruhehörschwelle erreicht wird, im Bereich von 30ms bis 40ms. Dadurch wird die Hörbarkeit eines Testimpulses durch den Beginn eines Störschalls weniger gestört als durch sein Ende [12]. Die Vorverdeckung spielt also im Vergleich zur Nachverdeckung eine untergeordnete Rolle und wird deshalb häufig vernachlässigt [11].

### 3.3. Mithörschwellen-Periodenmuster

Bei Schallen mit zeitlichen Lücken wirken Vor-, Simultan- und Nachverdeckung zusammen. Es ergeben sich dabei die sogenannten Mithörschwellen-Periodenmuster (MSPM) [11].

In Abbildung 22 ist das Mithörschwellen-Periodenmuster dargestellt, das von einem verdeckenden 1kHz-Ton, der rechteckförmig moduliert ist, hervorgerufen wird.

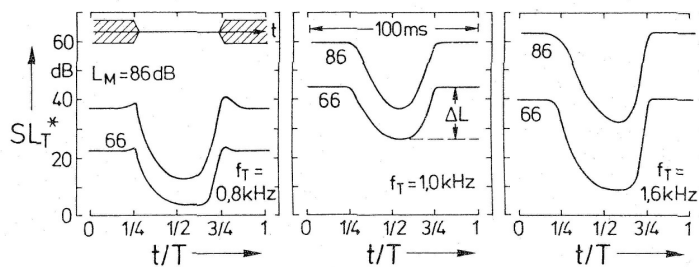


Abbildung 22: Mithörschwellen-Periodenmuster durch 1kHz-Ton [11]

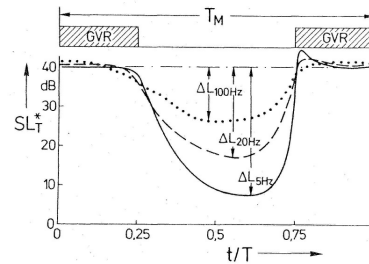


Abbildung 23: MSPM, durch Gleichmäßig Anregendes Rauschen [11]

Der Zeitzusammenhang wurde in der Darstellung so gewählt, dass der Zeitpunkt Null mit der Mitte des Maskierers zusammenfällt. Die Pause des Maskierers liegt also in der Mitte. Zwischen den einzelnen Bildern wird jeweils die Testtonfrequenz verändert.

Während für die Testtonfrequenzen  $f_T$  von 800Hz und 1kHz die Kurven bis auf die unterschiedlichen Höhenlagen sehr ähnlich verlaufen, hebt sich die 1,6kHz-Kurve deutlich ab. Sowohl der Abstand der beiden Kurven voneinander als auch die Differenz zwischen dem Minimum und dem Maximum der Mithörschwellen-Periodenmuster ist größer. Dies lässt sich mit der nichtlinearen Auffächerung der oberen Mithörschwellen-Flanke (siehe Punkt 3.7 *Verdeckung durch Sinustöne*) erklären [11]. In allen Kurven von Abbildung 22 ist jedoch die deutlich geringere Auswirkung der Vorverdeckung gegenüber der Nachverdeckung zu erkennen.

Zu Auswirkungen kommt es ebenfalls, wenn die Modulationsfrequenz des Maskierers verändert wird. Dadurch ändert sich die Form der Mithörschwellen-Periodenmuster: bei hohen Modulationsfrequenzen nähert sich die Steilheit des Abfalls der Nachverdeckung der Steilheit des Anstiegs der Vorverdeckung an (Abbildung 23).

Zu beachten ist bei dieser Graphik, dass sich die absolute Zeitdauer der Periode für verschiedene Modulationsfrequenzen verändert:  $T = 200\text{ms}$  für  $f_{\text{mod}} = 5\text{Hz}$  (durchgezogene Linie),  $T = 50\text{ms}$  für  $f_{\text{mod}} = 20\text{Hz}$  (gestrichelte Linie), und  $T = 10\text{ms}$  für  $f_{\text{mod}} = 100\text{Hz}$  (gepunktete Linie). Die Dauer der Lücke beträgt dementsprechend 100ms, 25ms und 5ms.

### 3.4. Mithörschwellen-Zeitmuster

Wird die Mithörschwelle mit dem Maskierpegel verglichen, so ist zu beobachten, dass eine Übereinstimmung nicht nur im Mittel erreicht wird, sondern dass die Mithörschwelle der zeitlichen Struktur des maskierenden Schalles sehr genau folgt. Dies kann in Abbildung 24 nachvollzogen werden: Als durchgezogene Linie ist der Logarithmus der Umhüllenden des Schalldruckverlaufs als Funktion der Zeit abgebildet. Die charakteristischen Einbrüche von Rauschen sind im Verlauf dieser Hüllkurve deutlich zu sehen. Punktiert dargestellt sind die Mithörschwellen eines 3kHz-Testtonimpulses mit der Länge 2ms.

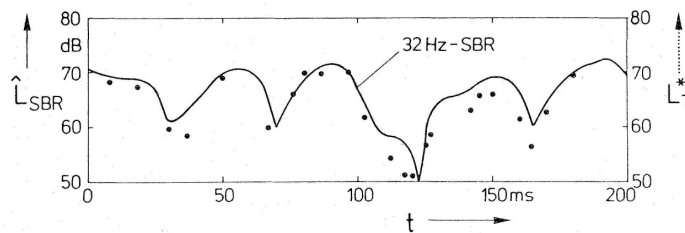


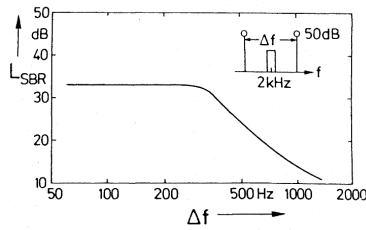
Abbildung 24: Mithörschwellen-Zeitmuster [11]

Je kleiner dabei die Bandbreite des verdeckenden Rauschens ist, umso besser wird die Korrelation. Bei großen Bandbreiten ändert sich der Schalldruckzeitverlauf so rasch, dass die Mithörschwelle diesem Zeitverlauf auf Grund der Nachverdeckung nicht mehr folgen kann [11].

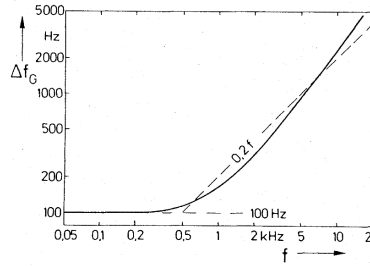
### 3.5. Frequenzgruppen und Tonheit z

Stark vereinfacht kann man sich die Frequenzgruppen als im Gehör realisierte Filter vorstellen. Innerhalb bzw. außerhalb derer arbeitet das Gehör nach unterschiedlichen Gesetzmäßigkeiten: in einer Frequenzgruppe fasst das Gehör alle auftretenden Schallintensitäten zusammen, außerhalb nicht [8] (siehe auch 5.3.2. *Störsignale*). Auch bei der Lautstärke- und der Richtungsempfindung spielt die Frequenzgruppe eine Rolle.

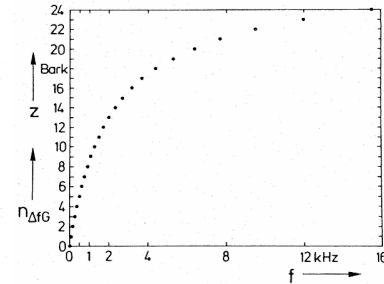
Dies kann sehr gut durch Mithörschwellenmessungen dargestellt werden. Die Mithörschwelle von Schmalbandrauschen, das von zwei Tönen im Frequenzabstand  $\Delta f$  maskiert wird, ist zunächst bei kleinen Bandbreiten unabhängig vom Frequenzabstand der beiden Töne (Abbildung 25). Erst wenn ein bestimmter Frequenzabstand  $\Delta f$ , nämlich die Frequenzgruppenbreite  $\Delta f_G$  überschritten wird, fällt die Mithörschwelle zu kleineren Pegeln hin ab [11].



**Abbildung 25:**  
Mithörschwelle mit Schmalbandrauschen [11]



**Abbildung 26:** Breite der Frequenzgruppe als Funktion der Frequenz [11]



**Abbildung 27:** Zusammenhang zwischen Frequenzgruppen, Tonheit z und der Frequenz [11]

Die Breite der Frequenzgruppen wurde aus Messungen mit vielen Testpersonen gewonnen. In Abbildung 27 ist sie als Mittelwert über der Frequenz aufgetragen. Zur genaueren Verdeutlichung sei hier zusätzlich eine Tabelle mit den wichtigsten Zusammenhängen angegeben:

**Tabelle 1:** Zusammenhang zwischen Tonheit z und Frequenz f, sowie zwischen Frequenzgruppenbreite  $\Delta f_G$  und Mittenfrequenz  $f_m$  [11]

z [Bark]	$f_u$ [Hz]	$f_o$ [Hz]	$f_m$ [Hz]	$\Delta f_G$ [Hz]	z [Bark]	$f_u$ [Hz]	$f_o$ [Hz]	$f_m$ [Hz]	$\Delta f_G$ [Hz]
1	0	100	50	100	13	1720	2000	1850	280
2	100	200	150	100	14	2000	2320	2150	320
3	200	300	250	100	15	2320	2700	2500	380
4	300	400	350	100	16	2700	3150	2900	450
5	400	510	450	110	17	3150	3700	3400	550
6	510	630	570	120	18	3700	4400	4000	700
7	630	770	700	140	19	4400	5300	4800	900
8	770	920	840	150	20	5300	6400	5800	1100
9	920	1080	1000	160	21	6400	7700	7000	1300
10	1080	1270	1170	190	22	7700	9500	8500	1800
11	1270	1480	1370	210	23	9500	12000	10500	2500
12	1480	1720	1600	240	24	12000	15500	13500	3500

Die Frequenzgruppe besitzt bei tiefen Frequenzen einen konstanten Wert von 100 Hz. Ab 500 Hz steigt sie zunächst proportional mit der Frequenz an, später etwas überproportional. Da Terzfilter eine Bandbreite von 23% der Mittenfrequenz besitzen, können sie bei Frequenzen über 500 Hz als brauchbare Näherung an die „Gehörfilter“ betrachtet werden [10].



Insgesamt gibt es 24 Frequenzgruppen im Hörbereich bis 16kHz. Diese spielen bei der Beschreibung von Hörwahrnehmungen eine so wichtige Rolle, dass sie zum Aufbau einer neuen Skale benützt werden: Die Frequenzgruppenskale (englisch: critical band rate) entsteht dadurch, dass die Frequenzgruppen so aneinandergereiht werden, dass die obere Grenze der unteren Frequenzgruppe gleichzeitig die untere Grenze der oberen Frequenzgruppe darstellt (Abbildung 27). Diese Skale wird als Tonheit  $z$  bezeichnet und hat die Einheit Bark erhalten (zu Ehren des Forschers von Barkhausen, der sich um die Schallmessung verdient gemacht hat und die Einheit „phon“ eingeführt hat) [7, 11].

### 3.6. Verdeckung durch Zufallsrauschen

Eine Möglichkeit der Signalform als Störsignal ist Rauschen. Rauschen ist ein Schallsignal von statistischer Natur, bei dem nur ein kontinuierliches Intensitätsspektrum angegeben werden kann. Die beiden wichtigsten Arten von Rauschen sind Weißes Rauschen und Rosa Rauschen.

Weißes Rauschen ist aus sehr vielen, dicht nebeneinander liegenden Sinusschwingungen zusammengesetzt. Diese Teilschwingungen besitzen alle die gleichen Amplituden, die Phasen sind statistisch gleichmäßig unabhängig voneinander verteilt. Die Teiltondichte pro Hz Bandbreite ist dabei konstant. Dies bedeutet, dass die spektrale Intensitätsdichte  $dI/df$  in einem interessierenden Frequenzbereich konstant ist. Damit bleibt die Intensität in einem absoluten Frequenzband (z.B. Band von 100Hz Breite) konstant [7].

Beim Rosa Rauschen nimmt die Amplitude der einzelnen Teilschwingungen pro Frequenzverdopplung um den Faktor 0,7 ab, der Pegel also um 3dB. Dadurch ist die spektrale Intensitätsdichte umgekehrt proportional zur Frequenz. Die Intensität bleibt in einem relativen Frequenzband, also in einem bestimmten Intervall (z.B. in einem Terzbereich) konstant [7].

#### 3.6.1. Schalldruckpegel und Dichtepegel

Da, wie oben beschrieben, Rauschen ein statistisches Signal ist, ist es sinnvoll, anstelle eines Spitzenwertes der Schallintensität (wie bei Sinustönen) die Schallintensität oder den Effektivwert des Schalldruckes zu verwenden.

Häufig wird anstelle der Schallintensität  $I$  auch die spektrale Schallintensitätsdichte  $dI/df$  angegeben. Diese Dichte umfasst jedoch einen so großen Wertebereich, dass es zweckmäßig ist, ein logarithmisches Maß zu benutzen.

Als Bezugsdichte wird

$$I_0 = \frac{10^{-12} \text{ W / m}^2}{1 \text{ Hz}} \quad \text{Gleichung 1}$$

verwendet. Daraus ergibt sich der Schallintensitätsdichtepegel, der zur Unterscheidung vom Schallpegel L mit l angegeben wird:

$$l_{\text{Rauschen}} = 10 \cdot \lg \frac{d(\tilde{p}/20 \mu\text{Pa})^2}{df / \text{Hz}} \text{ dB} = 10 \cdot \lg \frac{dI / (10^{-12} \text{ W / m}^2)}{df / \text{Hz}} \text{ dB} \quad \text{Gleichung 2}$$

Wenn Verwechslungen ausgeschlossen werden können, wird der Schallintensitätsdichtepegel auch als Dichtepegel bezeichnet [11].

Sehr häufig wird nicht breitbandiges Rauschen, sondern bandbegrenztes Rauschen verwendet. Ist dies der Fall, wird mit Hilfe eines Schalldruckmessers der Effektivwert des Schalldruckes als einzige sinnvolle Größe gemessen, die das Rauschen charakterisiert.

Sind neben dem Effektivwert des Schalldruckes noch die Grenzfrequenzen  $f_o$  und  $f_u$ , zwischen denen Weißes Rauschen vorhanden ist, bekannt, kann der Dichtepegel bestimmt werden (siehe auch 5.3.2 *Störsignale*):

$$l_{\text{WR}} = \left( 20 \cdot \lg \frac{\tilde{p}}{20 \mu\text{Pa}} - 10 \cdot \lg \frac{f_o - f_u}{1 \text{ Hz}} \right) \text{ dB} \quad \text{Gleichung 3}$$

Ist der Dichtepegel bekannt und soll der Schalldruckpegel L bestimmt werden, so ist das mit der entsprechend umgeformten Formel möglich:

$$L = \left( l_{\text{WR}} + 10 \cdot \lg \frac{f_o - f_u}{\text{Hz}} \right) \text{ dB} \quad \text{Gleichung 4}$$

Unter der Annahme, dass Weißes Rauschen zwischen 20Hz und 20kHz vorhanden ist, ergibt sich damit

$$L_{\text{WR}} = l_{\text{WR}} + 10 \cdot \lg(20000 - 20) \text{ dB} = l_{\text{WR}} + 43 \text{ dB} \quad \text{Gleichung 5}$$

3.6.2. Breitbandrauschen

Die durch Weißes Rauschen erzeugten Mithörschwellen verlaufen, wie in Abbildung 28 zu sehen, bei tiefen Frequenzen horizontal. Oberhalb von 500Hz steigen sie mit der Frequenz um 10dB/Dekade an. Dieser Frequenzgang kann mit den Frequenzgruppen erklärt werden: unterhalb 500Hz sind die Frequenzgruppen gleich breit, es werden immer gleich viele Frequenzanteile und damit gleiche Intensitäten vom Gehör zusammengefasst. Oberhalb 500Hz nimmt die Breite der Frequenzgruppen proportional mit der Frequenz zu. Dadurch wird – die Leistung des Weißen Rauschens ist in konstant breiten Frequenzabschnitten konstant – vom Ohr mehr Leistung zusammengefasst, es ist also ein höherer Pegel des Nutzsignals notwendig, um überhaupt wahrgenommen zu werden [12].

Auffällig bei diesen Versuchen ist die sehr genaue Übereinstimmung der Ergebnisse. Die starken individuellen Schwankungen, wie sie bei den Ruhehörschwellen vorhanden sind, verschwinden [11]. Änderungen des Pegels des Störschalles bewirken bei den Hörschwellen praktisch eine Parallelverschiebung [12].

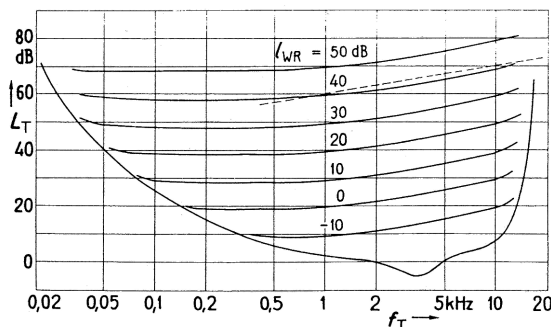


Abbildung 28: Mithörschwellen von WR [12]

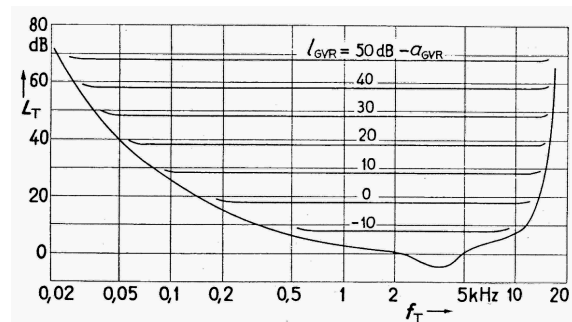


Abbildung 29: Frequenzunabhängige Mithörschwelle durch GVR [12]

Teilweise ist es für Versuche sinnvoll, im gesamten Frequenzbereich horizontal verlaufende, d.h. frequenzunabhängige Mithörschwellen zu erzeugen (Abbildung 29). Dieses Rauschen wird Gleichmäßig Verdeckendes Rauschen (GVR) genannt.

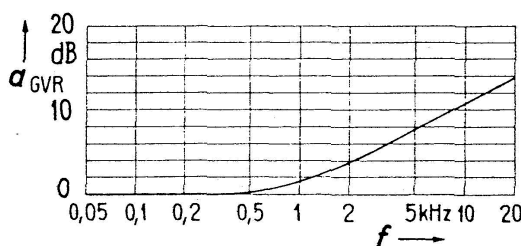


Abbildung 30: Verlauf des Dämpfungsfaktors  $a_{GVR}$  [12]

$$l_{GVR} = l_{WR} - a_{GVR}$$

Gleichung 6

Erreicht wird dies dadurch, dass das Rauschen oberhalb von 500Hz in seinem Dichtepegel so abgesenkt wird, wie die Mithörschwellen von Weißem Rauschen ansteigen. Praktisch wird dafür dem Weißen Rauschen ein Dämpfungsglied mit dem Dämpfungsverlauf  $a_{GVR}$  nachgeschaltet (Abb. 30). Dadurch erhält man aus dem Dichtepegel  $l_{WR}$  von Weißem Rauschen den Dichtepegel von Gleichmäßig Verdeckendem Rauschen  $l_{GVR}$  [12].

### 3.6.3. Schmalbandrauschen

Unter Schmalbandrauschen versteht man Rauschsignale mit Bandbreiten kleiner als die Frequenzgruppe (siehe auch 7.2.2. *Störsignal Terzrauschen*). Mit diesem Signal ist es zweckmäßig, anstelle des Dichtepegels wieder mit dem Pegel des Schmalbandrauschens zu rechnen. Die Umrechnungen sind mit Hilfe von Gleichung 4 leicht möglich [11].

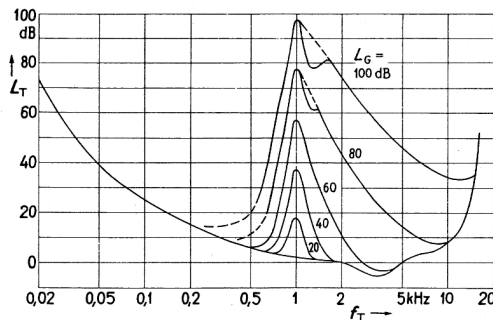


Abbildung 31: Schmalbandrauschen bei 1kHz [12]

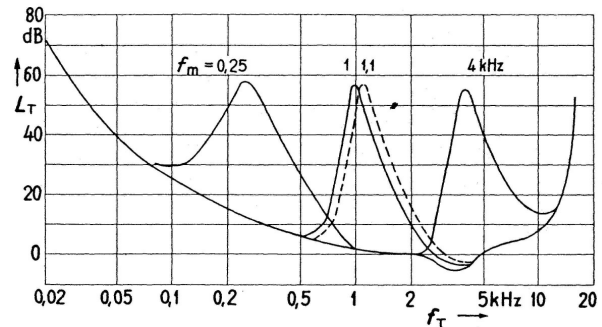


Abbildung 32: Variation der Mittenfrequenzen des Schmalbandrauschens [12]

Abbildung 31 zeigt die Mithörschwellen von Testtönen mit der Mittenfrequenz 1kHz und unterschiedlichen Schallpegeln  $L_G$  des Störrauschens.

Alle Mithörschwellen zeigen einen sehr steilen Anstieg, der vom Pegel des Störrauschens weitgehend unabhängig ist. Ihr Maximum erreichen sie bei der Mittenfrequenz des Schmalbandrauschens mit ca. 4dB unter  $L_G$ . Nach höheren Frequenzen fallen die Mithörschwellen bei kleinen Störpegeln steil ab, bei mittleren und v.a. bei hohen Pegeln wird dieser Abfall deutlich flacher. Diese Abhängigkeit vom Störpegel ist ein nichtlinearer Effekt, der als nichtlineare Auffächerung der oberen Flanke bezeichnet wird [11].

Bei 80dB und 100dB Pegel des Störrauschens ergeben sich oberhalb von 1kHz Einsattelungen in der Mithörschwelle. Diese werden auf ein Hörbarwerden von Differenzrauschen zurückgeführt, was vermutlich mit der Nichtlinearität des Mittelohres zusammenhängt [12]. Die punktierten Kurven (Abb. 31) zeigen den Verlauf der Hörschwelle wenn tatsächlich der Testton hörbar wird.

Wird die Mittenfrequenz des Schmalbandrauschens variiert, während der Pegel in allen Fällen unverändert bleibt, entstehen die Kurven aus Abbildung 32.

Dabei sind zwei Effekte auffällig: Während die Mithörschwellen bei 4kHz und 1kHz ziemlich ähnlich verlaufen, erscheint die Mithörschwelle bei 250Hz deutlich verbreitert. Offenbar setzt zu tiefen Frequenzen hin eine Veränderung der Form der Mithörschwelle ein. Der zweite Effekt besteht darin, dass das Maximum, das von den Mithörschwellen erreicht wird, zu hohen Frequenzen hin abnimmt, obwohl der Pegel des Störrauschens unverändert ist. Die Differenz des Pegels des Hörschwellenmaximums zum Störpegel beträgt bei 250Hz nur 2dB, bei 1kHz bereits 3dB und bei 4kHz sogar 5dB [11].

Insgesamt kann jedoch bei den Mithörschwellen von Schmalbandrauschen immer beobachtet werden, dass die Flanken, die von den tiefen Frequenzen kommen, steiler ansteigen als sie nach hohen Frequenzen hin abfallen. Dieser Anstieg beträgt ca. 100dB/Oktave.

Deshalb muss bei ähnlichen Versuchen darauf geachtet werden, dass die Dämpfung der verwendeten Bandpässe an den Grenzfrequenzen steil genug ansteigt. Werden normale Terz- oder Oktav-Bandfilter verwendet, werden die Flanken der Mithörschwellen durch deren Dämpfung bestimmt [12].

#### 3.6.4. Tiefpass- und Hochpassrauschen

Wenn als Störschall tiefpass- bzw. hochpassgefiltertes Weißes Rauschen verwendet wird (Parameter ist, wie beim breitbandigen Rauschen, der Dichtepegel), erhält man die Mithörschwellen aus Abbildung 33 und 34 (siehe auch 7.2.3. *Störsignal Tiefpassrauschen*).

Dieser Verlauf wird am besten dadurch beschrieben, dass man von Weißem Rauschen ausgeht. An der Stelle der Grenzfrequenz des Rauschens wird dann die untere bzw. obere Flanke der Mithörschwelle von Schmalbandrauschen angehängt [11]. Dadurch wird der Unterschied mit steigendem Pegel immer größer, die obere Flanke verläuft auch hier bei hohen Pegeln sehr flach, während die untere Flanke sehr steil bleibt.

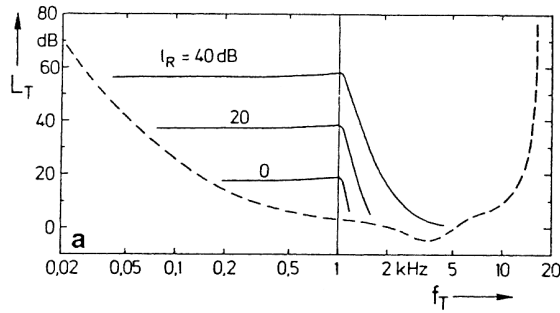


Abbildung 33: Tiefpassrauschen als Maskierer [11]

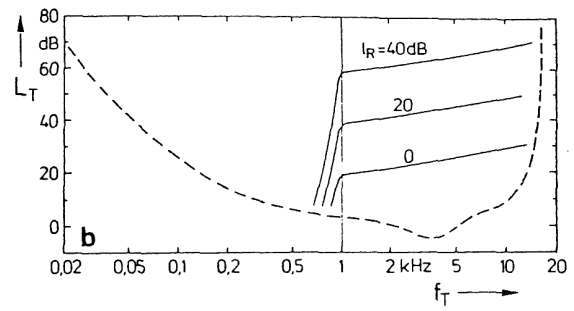


Abbildung 34: Verdeckung mit Hochpassrauschen [11]

Dadurch reicht die Kraft von starkem Tiefpassrauschen zur Verdeckung bis weit über die Grenzfrequenz hinaus [12]. Zusätzlich wird deutlich, wie wichtig die Mithörschwellen von Schmalbandrauschen sind.

### 3.7. Verdeckung durch Sinustöne

Die Bestimmung der Mithörschwelle, die durch einen einzelnen Sinusstörton hervorgerufen wird, bereitet v.a. bei mittleren und hohen Störpegeln einige Schwierigkeiten. Diese sind am Beispiel eines 1kHz-Sinusstörtones mit 80dB Pegel in Abbildung 35 dargestellt.

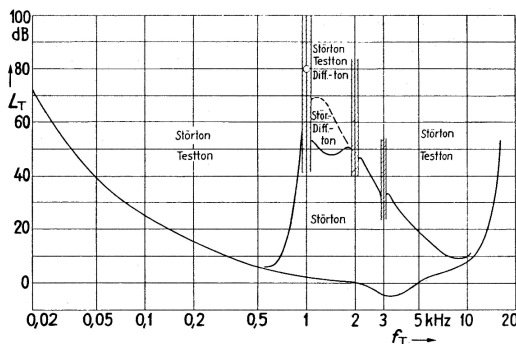


Abbildung 35: Auftretende Schwierigkeiten bei Messungen mit Sinus-Störtönen [12]

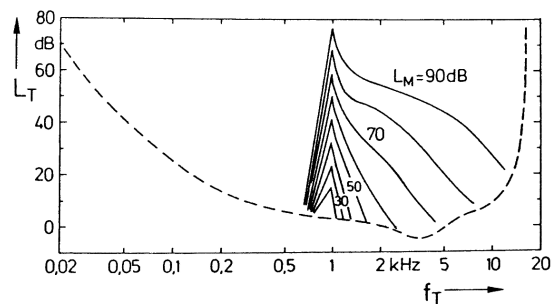


Abbildung 36: Effektive Mithörschwellen mit Sinusstörton [11]

In der Umgebung des 1kHz-Tones wird nicht mehr das Testsignal wahrgenommen, sondern Schwebungen. Diese Schwebungen, die ein völlig anderes Kriterium aufzeichnen (Hörbarkeit von Schwankungen), treten zusätzlich bei 2kHz und in abgeschwächter Form auch noch bei 3kHz auf. Eine weitere Schwierigkeit ist das Hörbarwerden eines Differenztones bei 600Hz, wenn die Hörschwelle bei ca. 1400Hz bestimmt werden soll. Dieser wird durch nichtlineare

Verzerrungen im Gehör erzeugt. Da er jedoch nicht der gesuchten Mithörschwelle für den Testton entspricht, muss hier bei der Messung genau unterschieden werden [11].

Mit geübten Testpersonen können dennoch genaue Ergebnisse gewonnen werden (Abb. 36).

Hier wird erneut die steile untere Flanke der Mithörschwellen deutlich. Im Gegensatz zum Schmalbandrauschen ist jedoch die Tendenz zu beobachten, dass diese Flanke bei kleinen Pegeln flacher ist als bei hohen Pegeln. Die obere Flanke dagegen wird, identisch zum Schmalbandrauschen, mit größer werdenden Pegeln des Störsignals immer flacher. Damit tritt die nichtlineare Auffächerung auch bei der Verdeckung durch Sinustöne auf.

Das deutliche Verdeckungsmaximum tritt in der unmittelbaren Umgebung des Störtones auf [12] (siehe auch 7.2.1. *Störsignal Sinuston*).

Der Einfluss der Maskierer-Frequenz ist identisch zum Einfluss bei Schmalbandrauschen:

Bei tiefen Frequenzen der Maskierer werden die Mithörschwellen als Funktion der Testtonfrequenz etwas breiter, oberhalb von 500Hz können die verschiedenen Kurven nahezu durch horizontale Verschiebungen gewonnen werden [11].

### 3.8. Verdeckung durch Klänge

Verdeckt man den Testton nicht durch einen reinen Sinusstörton, sondern durch einen Klang, der sich aus mehreren oder vielen Teiltönen zusammensetzt, so verläuft die Hörschwelle im gesamten Frequenzgebiet der Teiltöne über der Ruhehörschwelle [12] (siehe auch 7.2.4. *Störsignal Tongemisch*). Vor allem im Bereich der Musik ist dieser Fall sehr wichtig. Bis auf wenige Ausnahmen (z.B. Flöte) sind die Töne der Musikinstrumente sehr obertonhaltig, da sie aus einer großen Zahl von harmonischen Schwingungen zusammengesetzt sind [11].

Je nach Zusammensetzung der Amplituden bzw. Pegeln dieser Teiltöne sind die hervorgerufenen Mithörschwellen sehr verschieden.

Als Beispiel dafür werden in den folgenden Abbildungen die Mithörschwellen eines laut gespielten tiefen Violintones (195Hz) (Abb. 37) und die eines laut gespielten hohen Violintons (Abb. 38) mit der Hörschwelle eines künstlich erzeugten 200Hz-Tones mit 10 harmonischen Obertönen (alle Harmonischen mit gleichen Amplituden und statistisch verteilten Phasen) (Abb. 39) verglichen.

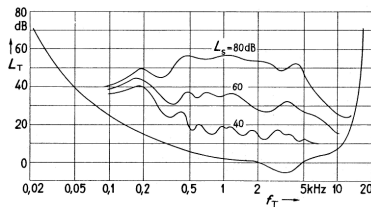


Abbildung 37: Lauter tiefer Violinton [12]

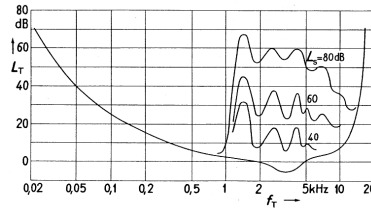


Abbildung 38: Lauter hoher Violinton [12]

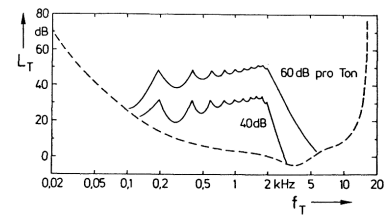


Abbildung 39: 200Hz-Ton mit 10 Harmonischen [11]

Hier kann bei allen Graphiken beobachtet werden, dass bei den einzelnen Spektralanteilen des Klanges Verdeckungserscheinungen wie bei Sinustönen auftreten. Diese überlagern sich gegenseitig und führen zu den dargestellten Mithörschwellen.

Ganz allgemein kann daraus gefolgert werden: Hohe leise Töne werden durch tiefe laute Töne maskiert, aber tiefe leise Töne werden durch hohe laute Töne nicht verdeckt.

Die Auswirkungen dieses Phänomens sind sehr weitreichend. So sind in einem gemischten Chor die Frauenstimmen immer stärker besetzt als die Männerstimmen, in einem Streichorchester spielen nur ein Bruchteil von Celli im Gegensatz zu Violinen. Auch für die Wiedergabe von Musik ist der flache Verlauf der Mithörschwellen zu hohen Frequenzen hin von großer Bedeutung: ein hoher Ton, der bei leisem Abspielen neben einem tiefen Ton gut hörbar ist, kann durch diesen tiefen Ton vollständig verdeckt werden, wenn die Wiedergabelautstärke zu stark erhöht wird. Falsche Wiedergabelautstärke verändert deshalb die Musik nicht nur quantitativ, sondern auch qualitativ [12].

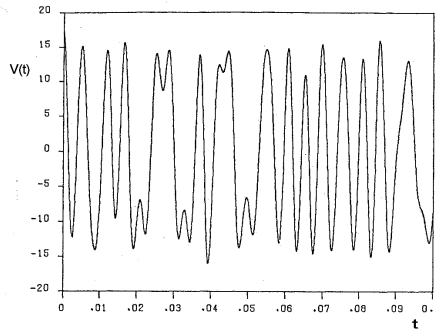
### 3.9. Verdeckung durch Low-noise noise

#### 3.9.1. Definition von Low-noise noise

Anstelle von natürlichen Signalformen können selbstverständlich auch künstliche, nach bestimmten Definitionen erzeugte Signale als Maskierer verwendet werden.

Eine solche Möglichkeit bietet Low-noise noise (LNN). Dieses periodische Signal wurde 1985 von Jon Pumplin definiert [13]. Es besitzt die Eigenschaft, dass es bei einem bestimm- baren Frequenzspektrum die geringstmögliche Amplitudenschwankung aufweist (siehe Ab- bildung 40). Es stellt damit eine Verbesserung zu den bisherigen zufälligen Rausch- Signaltypen dar und ist auch aus psychoakustischer Sicht von großem Interesse [13].





**Abbildung 40: Ausschnitt eines LNN-Signal [13]**

$$x(t) = \sum_{n=N_1}^{N_2} A_n \cdot \cos\left(\frac{2\pi n t}{T} + \phi_n\right) \quad \text{Gleichung 7}$$

Formelmäßig wird Low-noise noise durch Gleichung 7 beschrieben.

Dabei bezeichnet  $T$  die Periodendauer,  $n$  die Anzahl der Teiltöne,  $A_n$  die Amplituden der einzelnen Teiltöne und  $\Phi_n$  die Phasenwinkel im Bereich von  $0$  bis  $2\pi$ . Die Anzahl der Harmonischen beträgt  $N = N_2 - N_1 + 1$ .

Die Amplituden und die Periodendauer können speziell für die gewünschte Anwendung bestimmt werden. Um damit die minimale Amplitudenschwankung zu erhalten, müssen die Phasenwinkel der einzelnen Teiltöne nach einem speziellen Verfahren zugeordnet werden. Dieses Verfahren der Low-noise noise Technik wurde so entwickelt, dass das vierte Moment der Signalform

$$\overline{x^4} = \frac{1}{T} \int_0^T [x(t)]^4 dt \quad \text{Gleichung 8}$$

minimiert wird. Dies ist äquivalent damit, die Abweichung der Momentanleistung  $x^2$  vom Durchschnittswert zu minimieren, da gilt

$$\overline{(x^2 - \overline{x^2})^2} = \overline{x^4} - \overline{x^2}^2 \quad \text{Gleichung 9}$$

und die durchschnittliche Leistung unabhängig von den Phasenlagen ist [13]:

$$\overline{x^2} = \sum \frac{A_n^2}{2} \quad \text{Gleichung 10}$$

### 3.9.2. Erzeugung von Low-noise noise

Das vierte Moment des Signals, das für die Erzeugung von Low-noise noise minimiert wird, wird zunächst mit dem Quadrat der Durchschnittsleistung normiert:

$$W = \overline{x^4} / \left(\overline{x^2}\right)^2 \quad \text{Gleichung 11}$$

Da die Leistung, wie oben erwähnt, unabhängig von den Phasenwinkeln ist, spielt diese Normalisierung bei der Minimierung des vierten Moments keine Rolle [14].

Jetzt müssen die Phasenwinkel der einzelnen Teiltöne gewählt werden. Dabei kann ein Winkel beliebig eingesetzt werden, alle anderen  $N-1$  Phasenlagen werden berechnet. Es wird dafür der Algorithmus von Pumpilin [14] verwendet. Bildlich gesehen geschieht dabei folgendes: Der multidimensionale Raum der Phasenwinkel ist übersät mit Einkerbungen, in denen das vierte Moment klein wird. Allerdings sind dabei einige Vertiefungen tiefer und damit besser als andere. Ziel ist es nun, den tiefsten Krater zu finden. Dazu startet man an verschiedensten, zufällig gewählten Raumpunkten und sucht das jeweilige nächste Minimum. Mit einer genügend großen Zahl an Startpunkten ist dann das kleinste gefundene Minimum auch gleichzeitig das globale Minimum und damit das Endergebnis.

Die Einzelschritte für die exakte Berechnung sehen dazu folgendermaßen aus:

Zunächst wandert man von einem beliebig gewählten Startpunkt in Richtung des nächstgelegenen Minimums. Diese Richtung wird durch den Gradienten  $g$  vorgegeben: Man variiert alle Phasen und hält gleichzeitig

$$\phi'_n = \phi_n + xg_n \quad \text{Gleichung 12}$$

konstant. Dadurch wird  $\overline{V^4}$  als Funktion von  $x$  minimiert.

Der Gradient  $g_n$  am Ausgangspunkt  $\Phi_n$  ist dabei definiert als

$$g_n = \partial(\overline{V^4})/\partial\phi_n \quad \text{Gleichung 13}$$

und gibt die Richtung an, in der  $\overline{V^4}$  sich am stärksten ändert.

Anschließend berechnet man den neuen Gradienten  $g'$  am neuen Ausgangspunkt. Mit diesem minimiert man wiederum  $\overline{V^4}$ , indem man erneut alle Phasen variiert:

$$\phi''_n = \phi'_n + yg_n + zg'_n \quad \text{Gleichung 14}$$

Das gefundene Minimum von  $\overline{V^4}$  ist damit eine Funktion von  $y$  und  $z$ .

Diese Schritte, den neuen Gradienten und damit das neue Minimum zu berechnen, werden solange wiederholt, bis sich  $\overline{V^4}$  nicht mehr ändert [13].

Dabei gelten folgende Zusammenhänge [14]:

1. Die Anzahl der Einkerbungen steigt überproportional mit der Anzahl der harmonischen Teiltöne an:  
 $N = 12 \rightarrow 28 \text{ Krater} \rightarrow 1000 \text{ Starts notwendig}$   
 $N = 24 \rightarrow 12500 \text{ Krater} \rightarrow 1\,000\,000 \text{ Starts notwendig}$
2. Für ein vorgegebenes Spektrum besitzen viele Einbuchtungen sehr ähnliche Tiefen, die sich nur unwesentlich vom globalen Minimum unterscheiden. Dadurch ist es ohne große Folgen, welches der gefundenen Löcher für das vierte Moment verwendet wird.
3. Die Anzahl der lokalen Minima ist begrenzt. Wenn die Anzahl der harmonischen Teiltöne zu groß gewählt wird, wird es unwahrscheinlich, mit dieser Methode ein globales Minimum und damit ein Minimal-Schwankungs-Signal zu finden.
4. Abgesehen von wenigen, sehr einfachen Fällen sind das lokale Minimum mit dem kleinsten vierten Moment und das lokale Minimum mit kleinstem Crest-Faktor unterschiedlich. Allerdings unterscheiden sich die Crest-Faktoren nur um wenige Prozent.

Entscheidend bei dieser Anwendung ist, dass alle partiellen Ableitungen, die in einem multi-dimensionalen Raum vorhanden sind, mit einer einzigen FFT bestimmt werden können [14].

Bei der Generierung von Low-noise noise muss auch unterschieden werden, ob das erzeugte Signal schmalbandig oder breitbandig ist [14]. Dabei gilt folgende Definition:

Ein LNN-Signal ist schmalbandig, wenn

$$f_m = N_1 + \left( \frac{N_2 - N_1}{2} \right) \geq N_2 - N_1 \quad \text{Gleichung 15.}$$

Eine äquivalente Definition von schmalbandig lautet:

$$N_2 \leq 3 \cdot N_1 \quad \text{Gleichung 16}$$

Wenn das Signal **schmalbandig** ist, ist das vierte Moment der Wellenform äquivalent 3/2-mal dem vierten Moment der Einhüllenden [14].

Dadurch besitzen diese Signale besondere Eigenschaften:

- Die Signalkomponenten können entlang der Frequenzachse verschoben werden, ohne dass sich das vierte Moment verändert, d.h. zu allen Phasenwinkeln kann ein konstanter Wert addiert werden.

Beispiel: Es ist für das vierte Moment egal, ob man die Lösungen für die Phasenwinkel für die Harmonischen Teiltöne 3, 4, 5, 6 und 7 oder für die Teiltöne 99, 100, 101, 102 und 103 benützt. Es muss nur beachtet werden, dass tatsächlich beide Signale schmalbandig sind.

Achtung: beim Verschieben über die Frequenzachse ändert sich der Crest-Faktor!

- Bei schmalbandigen Signalen kann nicht nur ein einziger Phasenwinkel beliebig gewählt werden, sondern zwei. Durch diesen gewonnenen Freiheitsgrad findet man – bildlich gesehen – keine Löcher bei der Berechnung der Phasenwinkel, sondern zusammenhängende Gräben, innerhalb denen sich das vierte Moment nicht ändert.

Dieser zusätzliche Freiheitsgrad kann nun dazu benutzt werden, ein möglichst geringes sechstes Moment zu berechnen und stellt damit eine optimierte Lösungsmöglichkeit dar. Es kann kein Signal mit geringerer Amplitudenfluktuation gefunden werden. Wenn allerdings die Ergebnisse dieser Berechnung über die Frequenzachse verschoben werden, ändert sich auch das sechste Moment und bleibt nicht minimal.

Ist das generierte Signal hingegen **breitbandig**, so gelten die Ergebnisse nur für die zur Berechnung benutzten Teiltöne. Die Wellenform dieser Signale ist abhängig von der Anzahl der Teiltöne und der relativen Amplituden der Teiltöne untereinander [14].

Phasenwinkel für ausgewählte Teiltonanzahlen bei schmalbandigen und breitbandigen Signalen wurden von Pumplin und Hartmann 1991 berechnet und können vom American Institut of Physics in New York angefordert werden [15], z.B. für:

- Breitbandige Signale, Teiltonanzahl 1 bis 18, 20, 24, 31, 36, 48, 60
- Signale mit abnehmender Amplitude ( $N = 40$ )
- Schmalbandige Signale, Teiltonanzahl 1 bis 17, 21, 24, 31, 36, 48
- Signal ausschließlich mit Oktav-Komponenten,  $N = 1-7$
- ...

Angegeben sind dabei jeweils die Anzahl der Teiltöne und die dazugehörigen Phasenwinkel, die Summe aller Phasenwinkel zur Überprüfung der richtigen Eingabe, das Vierte Moment, der Crest-Faktor und der Relative Peak-Faktor [15]. Mit diesen Werten ist es schnell und relativ einfach möglich, LNN-Signale mit vorgegebenem Aussehen zu erzeugen (siehe 5.3.2. *Störsignale*).

### 3.9.3. Verdeckungseffekte bei Verwendung von LNN als Maskierer

Um die Art der Verdeckung durch LNN zu untersuchen, wurden bereits in der Vergangenheit Hörschwellentests durchgeführt. Dazu sollte ein 1kHz-Ton im LNN-Rauschen detektiert werden. Als LNN-Signale wurden breitbandige und schmalbandige LNN-Signale erzeugt und mit Zufallsrauschen (Phasenlage der Teiltöne rein zufällig) verglichen [16] (siehe auch 7.3. *Messungen mit LNN*).

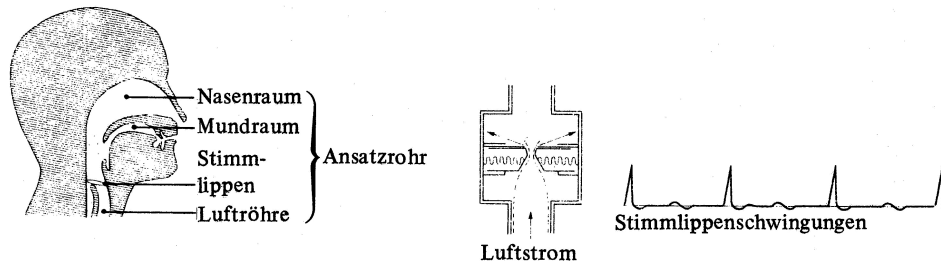
Wird dabei die spektrale Dichte der Signale konstant gehalten, ergeben sich folgende Ergebnisse: Bei schmalbandigen LNN-Signalen liegt die Hörschwelle ca. 5dB unterhalb der Hörschwelle von Zufallsrauschen. Bei breitbandigen Maskier-Signalen hingegen ist kein Unterschied feststellbar [16]. Dies bedeutet, dass der Verdeckungseffekt bei schmalbandigem Low-noise noise weniger effektiv ist als bei zufälligem Rauschen, nicht jedoch bei breitbandigem LNN.

Die Erklärung dafür lautet: Bei der Signaldetektion basiert das Resultat, ob das Testsignal gehört wird oder nicht, nicht nur auf einer einzelnen Signalerkennung, sondern auf der Kombination mehrerer Entdeckungen in aufeinanderfolgenden Subintervallen. Bei Low-noise noise ist die Energie in diesen Subintervallen auf Grund der geringen Amplitudenfluktuation nahezu gleich groß. Bei Zufallsrauschen hingegen schwankt diese Energie beträchtlich. Unterscheidet sich nun wie hier das Testsignal im Spektralbereich vom Maskierer, führen die unterschiedlichen Energiegehalte in den Subintervallen zu höheren Mithörschwellen [16].

Das negative Ergebnis bei breitbandigem LNN-Maskier-Signal kann mit Hilfe der Frequenzgruppen erklärt werden. Die minimale Amplitudenschwankung wird durch das Zusammenwirken der Phasen aller Teilkomponenten bewirkt. Werden nun einige dieser Komponenten durch ein Testsignal am Gehör der Testperson unterdrückt, so erscheint das LNN-Signal nicht mehr als Signal mit geringstmöglicher Amplitudenschwankung. Es verhält damit ähnlich zum Zufallssignal, es ergeben sich gleiche Mithörschwellen des Testsignals [16].

## 4. Theoretische Grundlagen der Sprachverständlichkeit

Die menschliche Stimme ist ein sehr vielseitiger Schallerzeuger. Sie kann Klänge und Geräusche mit unterschiedlichster akustischer Zusammensetzung erzeugen und diese je nach Notwendigkeit rasch ändern oder über längere Zeit beibehalten.



**Abbildung 41: Stimmorgan, mechanisches Modell der Stimmlippen und Schwingungen der Stimmlippen [7]**

So können stimmhafte oder stimmlose Laute, Zischlaute (F, S, SCH, CH), Explosivlaute (P, T, K, B, D, G) und Vokale erzeugt werden. Die Vokale werden dabei durch sogenannte „Formantbereiche“ – Frequenzbereiche, in denen alle Teiltöne unabhängig vom Grundton durch Resonanz verstärkt werden – bestimmt. Emotionen, wie z.B. Angst, Freude oder Wut prägen ebenfalls die Klangfarbe der Stimme durch unterschiedliche Hervorhebungen der Formanten im Frequenzbereich bis 3,7kHz [7].

### 4.1. Akustische Eigenschaften von Sprache

#### 4.1.1 Allgemeines

Die oben charakterisierten, unterschiedlichen Sprachlaute besitzen unterschiedliche akustische Eigenschaften. Diese werden durch ihre unterschiedliche Erzeugung hervorgerufen.

Die Sprechstimmlage bzw. die Sprechmelodie ändert sich dabei ständig. Bei Männern liegt sie vorrangig zwischen 120Hz und 160Hz, bei Frauen und Kindern zwischen 220Hz und 330Hz [7]. Dabei kann der Augenblickswert der Sprechstimmlage diese Frequenzbänder z. T. erheblich verlassen (meist in einen höheren Frequenzbereich), kehrt jedoch immer wieder zur mittleren Lage zurück [8].

Lautart	Akustische Eigenschaften
Vokale (A, E, I, O, U)	Harmonische Klänge, Grundfrequenz ist die Sprechtonhöhe, Charakterisierung durch Formantgebiete, Komponenten bis ca. 5kHz
Stimmhafte Konsonanten (L, M, N, R, W u.a.)	Harmonische Klänge mit der Grundfrequenz der Sprechtonhöhe, zusätzlich kontinuierliche Spektralanteile
Zischlaute (F, S, SCH, Z)	Nur geräuschhafte, kontinuierliche Spektren mit Komponenten bis über die Hörgrenze hinaus, v.a. bei S
H	Nur Strömungsgeräusch an den geöffneten Stimmlippen
Explosivlaute (B, D, G, P, T, K)	Impulsartige Vorgänge mit kontinuierlichen Spektren

Tabelle 2: Akustische Eigenschaften von Sprachlauten

## 4.1.2. Das Sprachspektrum

Der Frequenzbereich der Sprache umfasst ein breites Spektrum von 80Hz (Männer) bzw. 100Hz (Frauen) bis ca. 10kHz (Abbildungen 42 und 43) [17]. Tieferfrequente Komponenten fallen beim üblichen Abstand des Zuhörers unter die Hörschwelle.

Bei tiefen Frequenzen ist dabei der absolute Schalldruckpegel nahezu unabhängig von der Sprechstärke, nur der Abstand des Sprechers hat darauf einen bestimmten Einfluss. Hingegen wird das Sprachspektrum umso obertonreicher, je lauter gesprochen wird [7].

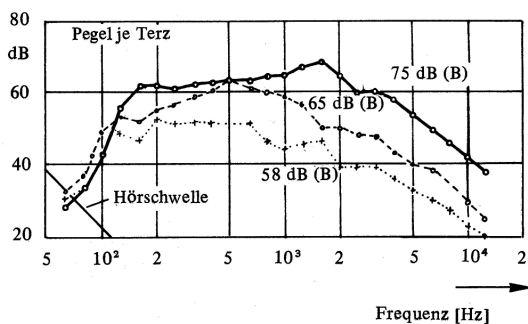


Abbildung 42: Männliches Spektrum [7]

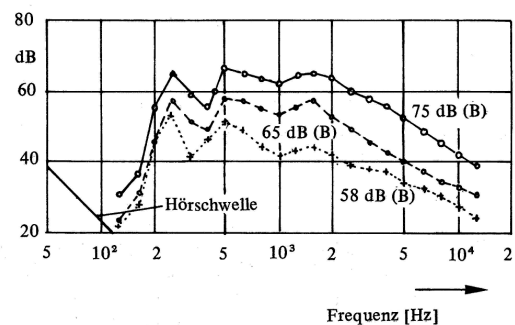


Abbildung 43: Spektrum einer Frauenstimme [7]

Die Unterschiede zwischen den Spektren verschiedener Sprecher sind dabei so gravierend, dass die Werte in den Abbildungen nur als statistischer Durchschnitt verstanden werden dürfen. Die größten Differenzen sind dabei unter 500Hz und über 1,5kHz festzustellen [17].

Wichtig ist auch die Energieverteilung auf die unterschiedlichen Frequenzbereiche. So konzentriert sich der größte Teil der Energie im Gebiet der niedrigen Frequenzen: unterhalb 500Hz liegt mehr als 50% der Gesamtenergie [19]. Die Sprachverständlichkeit wird jedoch nicht durch dieses Intervall, sondern durch die höheren Frequenzen von 600Hz oder 800Hz bis 2500Hz bestimmt. Die für die meisten Vokale und Konsonanten charakteristischen Laute liegen gerade in diesem Bereich. Daraus kann man schließen, dass Lärm, der in diesen wichtigen Frequenzbereich verdeckt, das Sprachverständnis besonders stark erschwert [19].

#### 4.1.3. Sprachschallpegel

Der Sprachschallpegel ist der durch Sprache hervorgerufene Schallpegel in einem gegebenen Abstand vom Sprecher. Da sich die Sprache in ihrem Pegel und Spektrum dauernd ändert, ist die Angabe eines gemittelten Spitzenpegels sinnvoll. Die Angabe in bewerteten Pegeln erleichtert zusätzlich den Vergleich mit Störgeräuschen [7].

In störungsfreier Umgebung spricht der Mensch mit einem Schallpegel, der – gemessen in 1m Entfernung – zwischen 55dB und 65dB liegt. Wird eine längere Unterhaltung oder ein Gespräch in entspannter Atmosphäre geführt, so wird in der Regel etwas leiser gesprochen ( $L_{\text{Sprache}} = 50 - 60\text{dB}$ ). Wirken hingegen Störgeräusche auf den Sprecher oder versucht dieser einer weiter entfernten Person etwas mitzuteilen, so steigt der Pegel um ca. 6dB an [17] (siehe auch 5.5. *Testsignale für Verständlichkeitsuntersuchung*). Bei unnatürlich lautem Sprechen können Werte in 1m Entfernung von 76dB(A) (Männer) bzw. 68dB(A) (Frauen) erreicht werden. Diese Werte liegen bei Entfernungshalbierung nochmals rund 3dB höher, bei Entfernungsverdopplung um 3dB niedriger. Die extremen Formen des Sprechens wie Murmeln oder Schreien erweitern die Dynamik nach unten bzw. oben um weitere 5dB [7].

Wichtig bei diesen Pegelangaben ist dabei auch der Ortsbezug. So ist der Sprechpegel am Ohr des Sprechers um ca. 11dB (!) höher als der Pegel in 1m Entfernung vor dem Mund. Der rückwärtige Schallpegel liegt um 4dB niedriger [17].



## 4.1.4. Dynamikverlauf

Der Pegelwert der gesprochenen Sprache ist, wie oben bereits erwähnt, zeitlich stark schwankend. Die Struktur kann als Abfolge von Impulsen, gebildet durch einzelne Logatome, beschrieben werden [7].

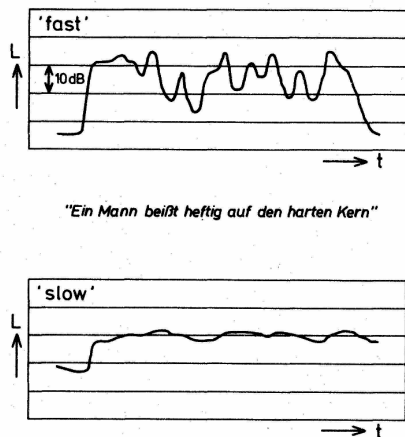


Abbildung 44: Unterschiedliche Dynamik des Sprachpegels [17]

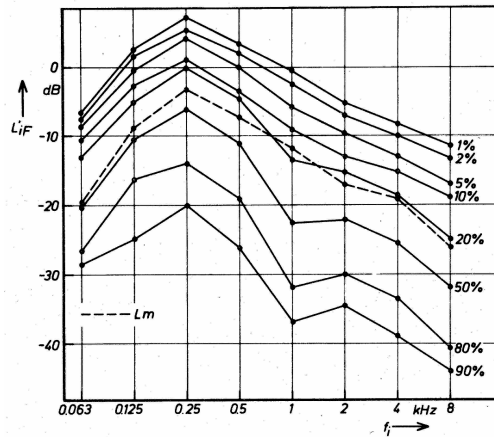


Abbildung 45: Frequenzbereiche mit den höchsten Sprachpegeln [17]

Die Dynamik des Sprechpegels (Abb. 44) wird einerseits durch die Betonung bestimmter Sprachlaute, Silben oder Worte hervorgerufen. Andererseits werden die einzelnen Laute auch innerhalb eines Wortes unterschiedlich laut gesprochen. Dies liegt an der Art der Erzeugung (Kehlkopf, Mund, Rachenraum). Die Vokale werden dabei um 20 – 25dB lauter ausgesprochen als die leisesten Konsonanten (F, P, D, P). Die Konsonanten M, NG, G, K, Z und S liegen in der Sprechintensität dazwischen [17]. Dies führt zu einem Durchschnittspegel, der rund 12dB unterhalb des Spitzenpegels liegt. Dadurch wirkt Sprache leiser als U-Musik mit derselben Aussteuerung, aber lauter als E-Musik [7].

Die Häufigkeiten des Sprachpegels sind im Wesentlichen durch die Pegel der höchsten Intensitäten bestimmt. Diese liegen bei etwa 250Hz (siehe Abbildung 45). Diese Häufigkeiten der bandgefilterten Sprache sind im gesamten Frequenzbereich der Sprache ähnlich. Dadurch können einheitliche linearisierte Häufigkeitsverteilungen für gleiche Pegeldifferenzen im jeweiligen Filter festgelegt werden [17].

Diese Häufigkeitssummenverteilung wird auch bei der Bestimmung des Artikulationsindex (AI) benützt: von Filter zu Filter muss der höchste Pegelwert der Sprache festgelegt werden (siehe 4.3.6. *Der Artikulationsindex AI*).

## 4.2. Messung der Sprachverständlichkeit

### 4.2.1. Hörschwelle von verbalen Reizen

Die Sprachverständlichkeit ist ein grundlegendes Kriterium für die Beurteilung der Hörsamkeit<sup>1</sup> von Räumen für Sprachdarbietungen [7].

Im Unterschied zur Schwellenbestimmung eines reinen Tons ist die Definition der Schwelle bei einem Sprachsignal nicht mehr eindeutig. Zum einen kann angegeben werden, ob ein Sprachreiz gehört wird oder nicht. In diesem Fall wird die Entdeckungsschwelle gemessen [17]. Andererseits kann eine Testperson die Aufgabe bekommen nachzusprechen, was sie verstanden hat. Es handelt sich dabei um die Bestimmung der Sprachverständlichkeitsschwelle  $L_{HSS}$ , die z.T. auch Hörschwelle für Sprache genannt wird. In diesem Fall liegt der Pegel der Sprache wesentlich höher als bei der Entdeckungshörschwelle.

Sprache, deren Schallpegel bei etwa 5 – 10dB liegt, kann nicht mehr verstanden, sondern nur noch erahnt (entdeckt) werden. Die Schwelle der Sprachverständlichkeit liegt bei ungefähr  $L_{HSS} = 20 - 35\text{dB}$ . Die Schmerzschwelle von Sprache entspricht der Schmerzschwelle von Sinustönen. Dabei ergeben sich die gleichen Werte für Sprachreize durch einen Kopfhörer und unter Freifeldbedingungen, wenn der Grundgeräuschpegel des Raumes genügend gering ist [17]. Betrachtet man die Verständlichkeit von Vokalen und Konsonanten in sinnlosen Silben, so kann festgestellt werden, dass Vokale schon verstanden werden, wenn der mittlere Pegel der Silben etwa 25dB beträgt. Konsonanten hingegen werden erst verstanden, wenn der Schallpegel bei 30 – 60dB liegt. Dies liegt daran, dass Vokale im Sprachfluss mit höherer Intensität gesprochen werden als Konsonanten. Allerdings enthalten die Konsonanten erheblich mehr Information als Vokale und sind deshalb für die Sinnverständlichkeit der Sprache enorm wichtig [17].

Die Sprachverständlichkeit wird bei Vorhandensein eines Störschalls dann deutlich beeinträchtigt, wenn das Signal-Rausch-Verhältnis am Hörort einen bestimmten Minimalwert unterschreitet. Dieser Grenzwert ist jedoch sehr stark vom individuellen Hörvermögen, der Art der Hörsituation, der Raumakustik und vom psychischen Zustand des Hörers abhängig, so dass darüber keine allgemeine Aussage gemacht werden kann [18].

Die beeinträchtigende Wirkung des Störschalls ist v.a. auf den psychoakustischen Effekt der Verdeckung zurückzuführen. Bei zeitvarianten Schallen, wie sie in realen Situationen meist

---

<sup>1</sup> Hörsamkeit: Siehe Punkt 4.2.2. *Definitionen*

vorhanden sind, kommen die Effekte der zeitlichen Vor- und Nachverdeckung zusätzlich hinzu. Die stärkste verdeckende Wirkung haben Störsignale, die dem Nutzsignal sehr ähnlich sind, sowohl in Bezug auf ihre spektralen als auch auf ihre zeitlichen Eigenschaften. Die Verständlichkeit von Sprache wird daher durch gleichzeitig vorhandenen Sprachschall am stärksten beeinträchtigt [18].

Bei normalem Hörvermögen ergibt das binaurale Hören gegenüber dem einohrigen Hören unter Störschallbedingungen einen beträchtlichen Gewinn an Sprachverständlichkeit, so dass gehörgeschädigte Personen besondere Schwierigkeiten beim Verstehen bekommen [18].

#### 4.2.2. Definitionen

Um bei der Verwendung von Fachbegriffen Missverständnisse ausschließen zu können, werden hier die wichtigsten Begriffe geklärt.

##### **Sprachverständlichkeit:**

Grundlegendes Kriterium für die Hörsamkeit von Sprachräumen, wobei man im wesentlichen zwischen Silbenverständlichkeit und Satzverständlichkeit unterscheidet [8].

##### **Hörsamkeit**

Oberbegriff, der die Wirkungen der akustischen Eigenschaften eines Raumes für Schalldarbietungen (Sprache oder Musik) am Ort des Hörenden beschreiben soll. Die Hörsamkeit beschreibt dabei nicht die physikalischen Eigenschaften eines Raumes, sondern die hörpsychologischen Wirkungen. Sie ist abhängig von den Eigenschaften des Schalls im Raum und damit auch innerhalb eines Raumes örtlich verschieden. Zusätzlich ist sie individuellen Schwankungen (z.B. Hörvermögen, Erfahrungen der Testperson usw.) unterworfen.

Die Hörsamkeit eines Raumes ist damit keine eindeutig festgelegte Größe, trotzdem ist eine „gute mittlere Hörsamkeit“ in Räumen für sprachliche Kommunikation sehr wichtig [8].

##### **Silbenverständlichkeit**

Anteil der richtig verstandenen Silben im Verhältnis zu allen vorgesprochenen Prüfsilben (Angabe in Prozent). Die Messung erfolgt, indem ein Sprecher mit guter Artikulation in den prüfenden Raum eine bestimmte Anzahl von Testsilben, sogenannte Logatome, vorliest. Diese besitzen keine semantische Funktion, damit der Proband nicht aus dem Zusammenhang Silben kombinieren kann. Dies würde das Messergebnis verfälschen.

Eine möglichst hohe Silbenverständlichkeit ist ein wichtiges Beurteilungskriterium für die Hörsamkeit von Sprachräumen [8].

### Satzverständlichkeit

Die für die verbale Kommunikation entscheidende Satzverständlichkeit ist viel höher als die Silbenverständlichkeit, da einzelne, nicht verstandene Silben aus dem Kontext leicht kombiniert werden können [8].

Zwischen der Silbenverständlichkeit, der Satzverständlichkeit und der Hörsamkeit gelten folgende Zusammenhänge:

**Tabelle 3: Zusammenhang zwischen Silbenverständlichkeit, Satzverständlichkeit und Hörsamkeit**

Silbenverständlichkeit	Satzverständlichkeit	Hörsamkeit von Sprache
50 %	90 %	ungenügend
70 %	95 %	befriedigend
80 %	100 %	gut

### Logatom

Eine Einzelsilbe mit der Lautabfolge Konsonant – Vokal – Konsonant (z.B. gul, peg), die mit einer einzigen sprachlichen Anstrengung hervorgebracht werden kann. Bei der Kombinationsbildung werden die Buchstaben statistisch dem Alphabet entnommen, wobei aber durch Zufall entstehende, sinnvolle Silben (z.B. kam) aussortiert werden. Die Logatome spielen für die Prüfung der Silbenverständlichkeit und in der Sprachaudiometrie eine wichtige Rolle [8].

#### 4.2.3. Sprachtests mit Silben, Wörtern und Sätzen

Beabsichtigt man akustische Bedingungen eines Raumes zu prüfen, so wird im Allgemeinen die Silbenverständlichkeit ermittelt, da auch Einzelsilben (z.B. Zahlen) verständlich sein sollen. Dadurch können zusätzlich nicht-akustische Faktoren wie Satzbau oder Inhalt ausgeschaltet werden (siehe auch 5.3.1. *Logatomlisten*).

Geht es darum, das sprachliche Hörvermögen, d.h. die Fähigkeit, Sprache unter bestimmten akustischen Bedingungen zu verstehen, zu erfassen, so wird eher Material verwendet, das die Struktur der Umgangssprache bezüglich Satzbau und Inhalt widerspiegelt [17].

Bei solchen Testverfahren müssen einige Faktoren berücksichtigt werden.

Die phonetischen Bedingungen sollten der Lautstruktur der Umgangssprache entsprechen. Das Verhältnis von Konsonanten zu Vokalen muss in etwa mit dem der deutschen Sprache, also 61% zu 39% übereinstimmen. Des Weiteren ist die Anzahl der Sprachlaute innerhalb einer Silbe zu bestimmen und die Stellung der Laute als Anlaut und Endlaut auszubalancieren. Werden Wörter im Test verwendet, so ist deren Auftrittswahrscheinlichkeit in der Umgangssprache zu kontrollieren, da dies einen wesentlichen Einfluss auf die Wahrnehmung ausübt. Wenn Untersuchungen mit Sätzen durchgeführt werden, so ist eine Analyse der Satzstruktur vom psycholinguistischen Standpunkt aus unumgänglich [17].

Auch die Art der Antwort kann variiert werden. In den meisten Tests muss die Testperson das, was sie verstanden hat nachsprechen oder aufschreiben. Es ist jedoch auch möglich, den Laut aus einer Anzahl von vorgegebenen Möglichkeiten herausuchen zu lassen.

Es kann deshalb eine Vielzahl von unterschiedlichen Sprachtests genannt werden, die jeweils für unterschiedliche Fragestellungen zugeschnitten sind [17]. Das dabei wichtigste Testverfahren ist dabei die Messung der Silbenverständlichkeit mit Logatomen (siehe auch 8.2. *Durchführung des Logatomtests*). Dieser Einsilber-Test hat den Vorteil einer geringen Redundanz (von einem Wort kann nicht auf das nächste geschlossen werden) und bietet eine hohe analytische Aussagekraft [18]. Deshalb soll dieses Verfahren hier etwas genauer beschrieben werden.

Die **Messung der Silbenverständlichkeit mit Logatomen** liefert gute Messgenauigkeiten, auch bei hohen Verständlichkeiten [19]. Dabei muss beachtet werden, dass isoliert dargebotene Logatome für die Verständlichkeitsmessungen in Räumen nicht geeignet sind. Bei einzeln gesprochenen Silben befindet sich der Raum nicht annähernd in eingeschwungenem Zustand. Zusätzlich dazu spielt die Sprechgeschwindigkeit eine große Rolle: wird sie langsamer, nimmt die Verständlichkeit zu.

Daraus ergeben sich zwei Forderungen für die Logatomlisten:

- Logatome müssen in fortlaufender Sprache eingebettet sein
- Das Testmaterial muss die Artikulationsgeschwindigkeit normaler Sprache besitzen

Typischerweise werden deshalb Prädiktorsätze vor die einzelnen Logatome gestellt. Diese bestehen aus je drei bis vier Silben:

Schreiben Sie ...

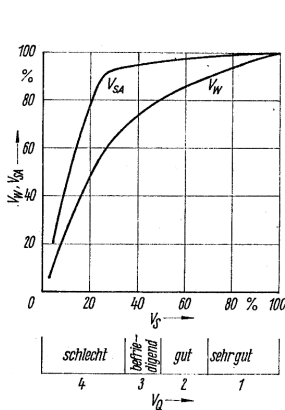
Und nun folgt ...

Jetzt schreiben wir...

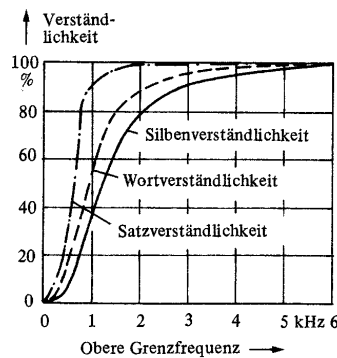
Die Artikulationsgeschwindigkeit<sup>2</sup> wird auf ca. fünf Silben pro Sekunde festgelegt. Dies entspricht in etwa normal gesprochener Sprache. Zwischen den einzelnen Logatomen werden Pausen zum Niederschreiben der verstandenen Silben eingebettet. Diese dürfen nicht zu kurz sein, um ein korrektes Schreiben des Verstandenen zu ermöglichen, andererseits dürfen sie nicht zu lang sein, da sonst die Sprechgeschwindigkeit zu langsam wird [19].

Eine Logatomliste besteht aus mindestens 50 Einzel-Logatomen. Diese wird der Testperson von einer sehr gut artikulierenden Person, im Idealfall ein geschulter Sprecher, vorgelesen.

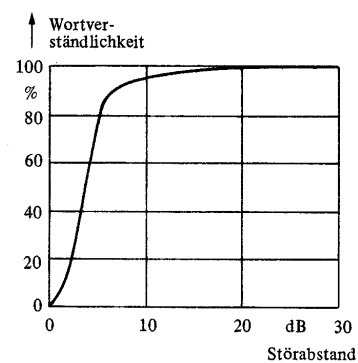
Die mit diesem Verfahren gemessene Silbenverständlichkeit hängt zusätzlich von weiteren Faktoren ab, z.B. von der Nachhallzeit des Raumes, vom Signal-Störabstand (Abb. 48), von der Sprachlautstärke oder der Breite des übertragenen Frequenzbandes des Sprachsignals (Abb.47) [17].



**Abbildung 46: Zusammenhang Verständlichkeit – Raumbewertung [20]**



**Abbildung 47: Verständlichkeit, abhängig von der oberen Grenzfrequenz [7]**



**Abbildung 48: Wortverständlichkeit, abhängig vom Signal-Stör-Abstand [7]**

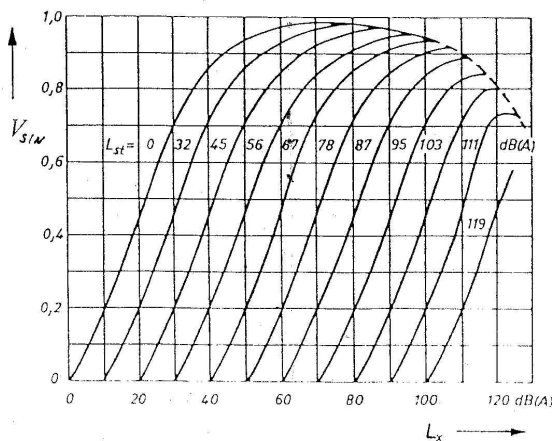
Als Beispiel für die Komplexität der Verständlichkeit ist in Abbildung 50 die Abhängigkeit vom Sprechpegel  $L_S$  und von der Grenzfrequenz  $f_o$  des Übertragungstiefpasses für sinnlose

<sup>2</sup> Artikulationsgeschwindigkeit = Anzahl der Silben bezogen auf die Artikulationszeit

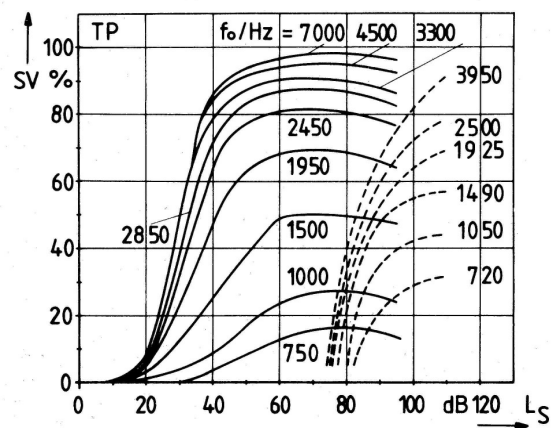
Achtung, darf nicht mit der Sprachgeschwindigkeit verwechselt werden, die sich aus Artikulationsgeschwindigkeit und der Anzahl der Pausen mittlerer Dauer zusammensetzt!

Silben als durchgezogene Linie angegeben. Man kann gut erkennen, dass umso schlechter verstanden wird, je geringer der übertragene Frequenzbereich der Sprache ist.

Mit der Erhöhung des Sprachpegels kann die Verständlichkeit bis zu einem gewissen Wert erhöht werden. Wird das Signal jedoch zu laut ( $L_S > 75\text{dB}$ ), nimmt die Verständlichkeit wieder ab (Abbildung 49). Nicht vernachlässigt werden darf dabei jedoch der Einfluss des Umgebungslärms ( $L_{St}$  in Abb. 49) auf die sprachliche Kommunikation (siehe auch 8.2.1. *Auswirkungen der Störgeräusche*).



**Abbildung 49: Silbenverständlichkeit, abhängig vom Sprachschalldruckpegel  $L_x$  und des Störpegels  $L_{st}$  [20]**



**Abbildung 50: Verständlichkeit, abhängig vom Pegel und der Grenzfrequenz eines Übertragungstiefpasses [17]**

Auch mit steigender Nachhallzeit  $T$  sinkt die Sprachverständlichkeit. Dargestellt ist dies in Abbildung 51 mit Ergebnissen aus verschiedenen Untersuchungen an normalhörenden und schwerhörigen Testpersonen. Auf die genaue Beschreibung aller Kurven soll hier verzichtet werden. Es ist jedoch gut der allgemeine Trend sichtbar, dass alle Kurven mit steigender Nachhallzeit eine deutlich verminderte Verständlichkeit aufweisen (siehe auch 8.2.2. *Auswirkungen der Raumzustände*). Wird zudem noch das Alter der Testpersonen miteinbezogen (Abbildung 52) wird der Trend nochmals verstärkt. Die Verständlichkeit sinkt mit zunehmendem Alter, ohne dass es einer Störung bedarf (Kurve KT). Treten zusätzlich Störfaktoren auf, so nimmt die Verständlichkeit überproportional ab.

Die Sprachverständlichkeit ist damit eine sehr komplexe Größe, bei deren Bestimmung genau auf die Einhaltung der Randbedingungen (geringe Umgebungslautstärke, Eignung der Probanden innerhalb der Testgruppen usw.) geachtet werden muss.

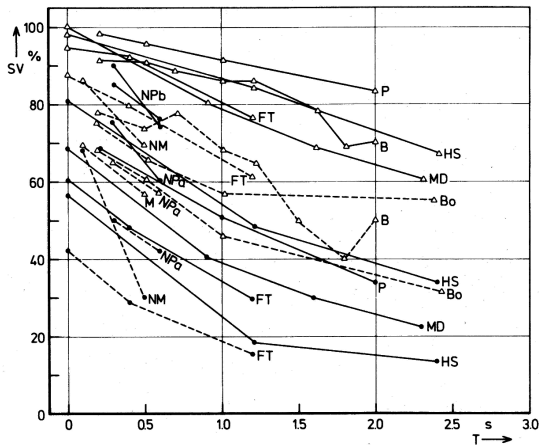


Abbildung 51: Verständlichkeit abhängig von der Nachhallzeit T [17]

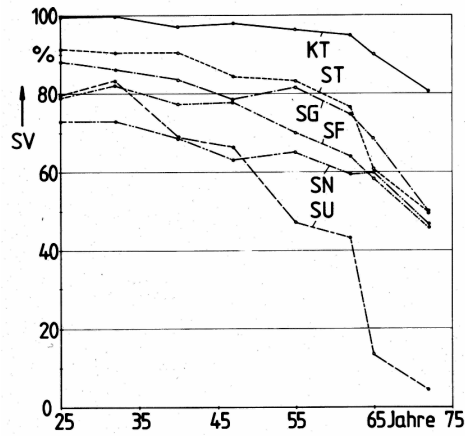


Abbildung 52: Verständlichkeit abhängig vom Alter der Testpersonen [17]

### 4.3. Berechnung der Sprachverständlichkeit durch physikalische Parameter

Die Verständlichkeit von Sprache kann nicht nur subjektiv gemessen werden. Es sind auch objektive Verfahren zur Bestimmung dieser Größe entwickelt worden (siehe auch 8.3. *Vergleich der subjektiven Ergebnisse mit berechneten Werten*). Diese sollen hier kurz vorgestellt werden.

#### 4.3.1. Deutlichkeitsgrad nach Thiele $D_{50}$

Frühe Energie erhöht die Durchsichtigkeit und damit auch die Deutlichkeit. Bei Sprache liegt die Grenze für diesen frühen Energieanteil ca. 50ms nach dem Direktschall [22]. Die messtechnische Beurteilung erfolgt durch den Deutlichkeitsgrad nach Thiele  $D_{50}$  mit

$$D_{50} = \frac{W_{0...50}}{W_{ges}} = \frac{\int_0^{50ms} p^2(t) dt}{\int_0^{\infty} p^2(t) dt} \tag{Gleichung 17}$$

wobei alle auftretenden Frequenzen im Hörbereich in die Energiebetrachtung miteinbezogen werden. Ab Werten von  $D_{50} > 50\%$  ist eine gute Verständlichkeit gegeben.

#### 4.3.2. Deutlichkeitsmaß $C_{50}$

Aus der Definition der Deutlichkeit nach Thiele kann man das Deutlichkeitsmaß  $C_{50}$  ableiten, das eine größere praktische Bedeutung hat. Es gibt die Pegeldifferenz der Schallenergie vor



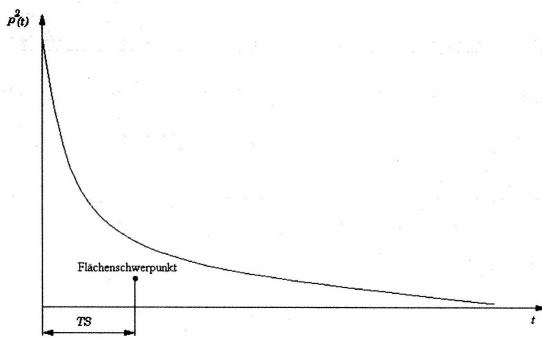
und nach 50ms an. Auch hier werden die alle Frequenzen des gesamten Hörbereichs verwendet.

$$C_{50} = 10 \cdot \log \frac{W_{0...50}}{W_{50...∞}} = 10 \cdot \log \frac{D_{50}}{1 - D_{50}} = 10 \cdot \log \frac{\int_0^{50ms} p^2(t) dt}{\int_{50ms}^{\infty} p^2(t) dt} \quad \text{Gleichung 18}$$

Für  $-8\text{dB} < C_{50} < -3\text{dB}$  ergibt sich eine ausreichende Verständlichkeit. Bei  $-3\text{dB} < C_{50} < 2\text{dB}$  gute und für  $C_{50} > 2\text{dB}$  sehr gute Verständlichkeit [22].

#### 4.3.3. Schwerpunktzeit TS (Center time CT)

Wird das Quadrat des Schalldrucks über die Zeit integriert, liefert das Ergebnis die Gesamtenergie  $W_{\text{ges}}$ . Aus dem Flächenschwerpunktsatz der Mechanik ergibt sich dazu die Schwerpunktzeit TS durch die örtliche Lage des Flächenschwerpunktes über der Zeitachse (Abb.53):



$$TS = \frac{1}{W_{\text{ges}}} \cdot \int_0^{\infty} t \cdot p^2(t) dt \quad \text{Gleichung 19}$$

Abbildung 53: Schwerpunktzeit TS [22]

Für  $TS < 70\text{ms}$  kann von guter Sprachverständlichkeit ausgegangen werden.

Das Deutlichkeitsmaß setzt eine scharfe Grenze zwischen günstigen und ungünstigen Schallenergien. Die Schwerpunktzeit hingegen ist ein Maß für die zeitliche Lage hoher Leistungen und entspricht deshalb der tatsächlichen Schallwirkung besser. Eine kurze Schwerpunktzeit bedeutet also einen steilen Anstieg des Energieverlaufs. Es hat sich jedoch gezeigt, dass sich bei einer Schwerpunktzeit unter 70ms die Silbenverständlichkeit nicht mehr weiter verbessert [22].

4.3.4. Objektive Silbenverständlichkeit  $v$ 

Die Objektive Silbenverständlichkeit  $v$  ist definiert als

$$v = 96 \cdot k_F \cdot k_N \cdot k_C \quad \text{Gleichung 20}$$

und wird in Prozent angegeben.  $90\% < v < 96\%$  ergeben dabei einen ausgezeichneten Wert, für  $v < 30\%$  sind die Testsilben als unverständlich zu bewerten [22].

Dabei berücksichtigt der Faktor 96 die Tatsache, dass auch bei optimalen Versuchsbedingungen maximal 96% aller Logatome richtig erkannt werden.

Der Faktor  $k_F$  berücksichtigt die Bandbreiteneinschränkung der Signalwiedergabe. Dies kann z.B. durch eine Beschallungsanlage geschehen. Dieser Faktor ergibt sich aus der oberen und der unteren Grenzfrequenz (Abb.54) unter dem Zusammenhang von Gleichung 21. Das Beschneiden der tiefen Frequenzen erhöht dabei den Signal-Stör-Abstand, da Störgeräusche in diesem Frequenzbereich in der Regel hohe Pegel besitzen und gleichzeitig Sprache hier kaum Komponenten hat. Liegt keine Bandbreiteneinschränkung vor, so ergibt sich  $k_F = 1$  [22].

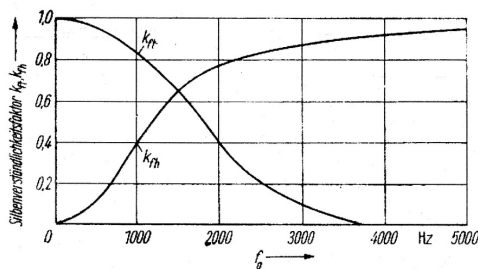
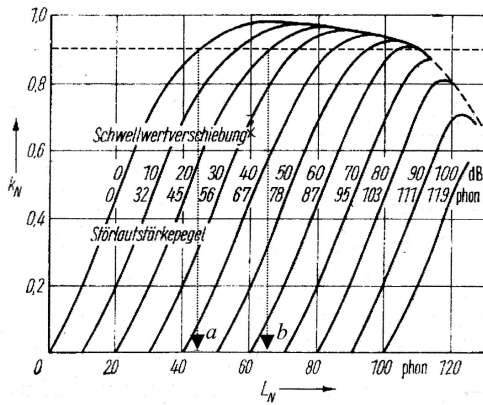
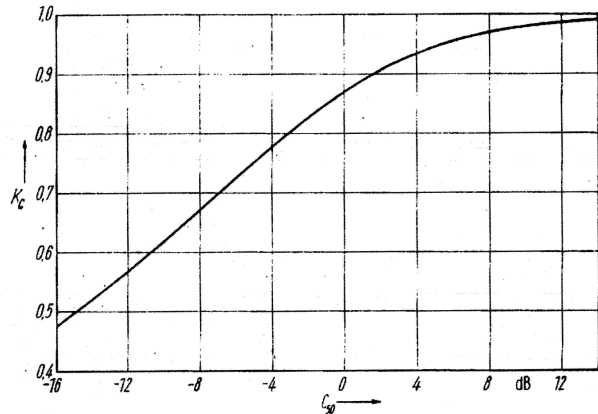


Abbildung 54: Berücksichtigung der Bandbreiteneinschränkung [22]

$$k_F = k_{ft} \cdot k_{fh} \leq 1 \quad \text{Gleichung 21}$$

Der Lautstärkefaktor  $k_N$  berücksichtigt den Signal-Stör-Abstand am Hörort, wenn sich die Spektren von Nutzsignal und dem störenden Grundgeräusch sehr ähnlich sind. Der entsprechende Wert muss aus Abbildung 55 abgelesen werden. Dabei ist auf der x-Achse die Lautstärke  $L_N$  des Nutzsignals aufgetragen. Der Parameter der Kurvenschar ist die Lautstärke des Störgeräuschs (in phon) bzw. die Schwellwertverschiebung<sup>3</sup> in dB.

<sup>3</sup> Die Schwellwertverschiebung gibt an, um wie viel dB der Nutzsignalpegel bei vorhandenem Störsignal erhöht werden muss, um die gleiche Verständlichkeit zu erreichen wie ohne Störsignal.

Abbildung 55: Bestimmung von  $k_N$  [22]Abbildung 56: Zusammenhang zwischen  $C_{50}$  und  $k_C$  [22]

Der Faktor  $k_C$  stellt den eigentlichen Zusammenhang zum Deutlichkeitsmaß  $C_{50}$  dar (siehe Abbildung 56). Die Formel wurde von Kürer empirisch bestimmt:

$$k_C = \sqrt[5]{\frac{1}{1 + 10^{\frac{C_{50}}{10 \text{ dB}}}}} \quad \text{Gleichung 22}$$

Mit dieser objektiven Bestimmung der Verständlichkeit soll der große Aufwand des subjektiven Logatom-Verfahrens erspart bleiben [22].

#### 4.3.5. Artikulationsverlust bzw. Konsonantenverlust $Al_{\text{cons}}$

Der Artikulationsverlust gesprochener Konsonanten (Articulation loss of consonants) ist bei der Beurteilung von Sprachverständlichkeit von großer Bedeutung. Davon ausgehend wurde ein entsprechendes Kriterium für die objektive Bestimmung der Verständlichkeit entwickelt [22].

Für die Voraussetzung eines genügend großen Signal-Stör-Abstandes von  $S/N > 35\text{dB}$  gilt:

$$Al_{\text{cons}} = 0,652 \cdot \left( \frac{r_{LH}}{r_R} \right)^2 \cdot T \quad \text{Gleichung 23}$$

Dabei entspricht  $T$  der Nachhallzeit,  $r_{LH}$  dem Abstand zwischen Schallquelle und Hörer und  $r_R$  der Richtentfernung. Die Richtentfernung wiederum setzt sich zusammen aus

$$r_R = \sqrt{\gamma} \cdot r_H \quad \text{Gleichung 24}$$

wobei  $\gamma$  den Bündelungsgrad der Schallquelle angibt und  $r_H$  den Hallradius (siehe Gleichung 25). Die obige Formel gilt dabei im gesamten Raum [22].

Nach [22] kann diese Formel jedoch für unterschiedliche Quellen-Hörer-Abstände vereinfacht werden. Bis zu einer Entfernung, die dem dreifachen effektiven Hallradius entspricht (siehe Gleichung 25)

$$r_H = 0,057 \cdot \sqrt{\frac{V}{T}} \quad \text{und} \quad r_{H,eff} = 0,057 \cdot \sqrt{\gamma} \cdot \sqrt{\frac{V}{T}} \quad \text{Gleichung 25}$$

gilt

$$Al_{cons} = \frac{r_{LH}^2 \cdot T^2}{\gamma} \cdot 200\%; \quad 0 < r_{LH} < 3 \cdot r_{H,eff} \quad \text{Gleichung 26}$$

In größerer Entfernung hängt  $Al_{cons}$  nur noch von der Nachhallzeit bei 1400Hz ab [22]:

$$Al_{cons} = T \cdot 9\% \quad R \geq 3,5 \cdot r_{H,eff} \quad \text{Gleichung 27}$$

Das Ergebnis wird in Prozent angegeben und weist folgenden Zusammenhang zur Verständlichkeit auf:

$Al_{cons} < 2 \%$	ideale Verständlichkeit
$2 \% < Al_{cons} < 7 \%$	gute Verständlichkeit
$7 \% < Al_{cons} < 20 \%$	befriedigende Verständlichkeit
$20\% < Al_{cons}$	unbrauchbare Verständlichkeit

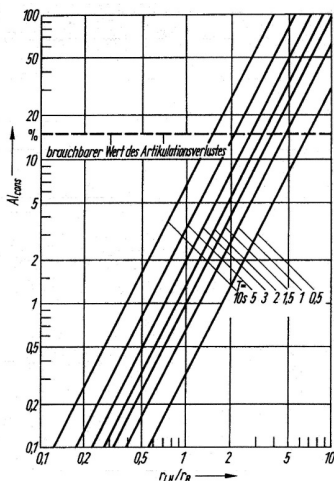


Abbildung 57: Einfaches Ablesen des  $Al_{cons}$  [22]

Der Artikulationsverlust nimmt also mit der Nachhallzeit und dem Abstand zwischen Quelle und Hörer zu.

Wenn der Bündelungsgrad  $\gamma = 1$  gegeben ist, kann der Wert des Endergebnisses auch direkt aus Abbildung 57 abgelesen werden.

Bis jetzt wurde immer davon ausgegangen, dass der Signal-Stör-Abstand mindestens 35dB beträgt. Sinkt dieser Wert auf unter 35dB, verschlechtern (erhöhen) sich die Werte von  $Al_{cons}$  erheblich [22]. Für diesen Fall kann das Ergebnis aus Abbildung 58 herausgelesen werden.

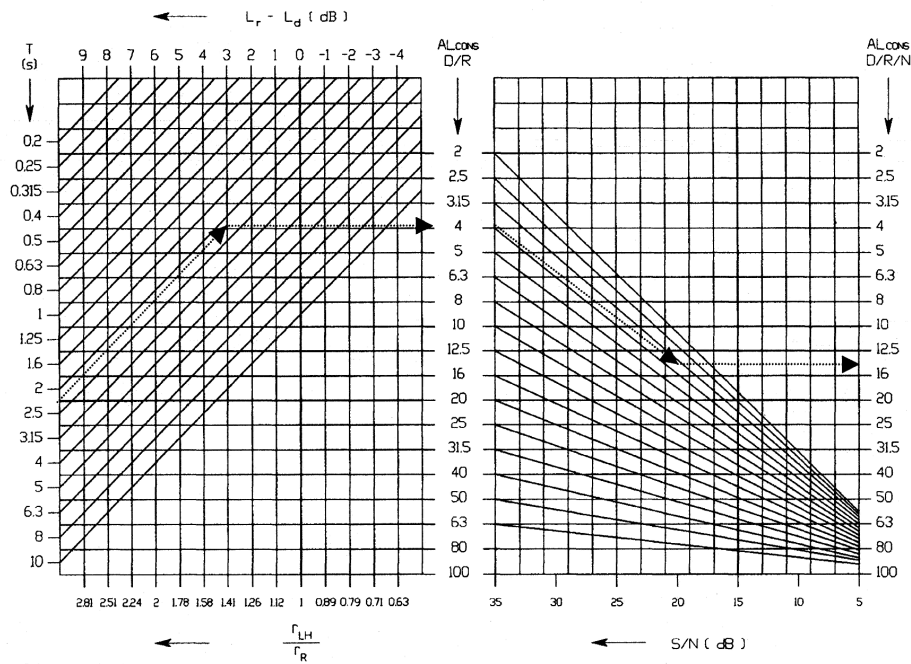


Abbildung 58: Bestimmung des  $AI_{cons}$  für  $S/N < 35dB$

Dabei wurde das Entferungsverhältnis  $r_{LH}/r_R$  zu einer Pegeldifferenz umgerechnet:

$$L_{diffus} - L_{direkt} = 20 \cdot \log \frac{r_{LH}}{r_R} \quad \text{Gleichung 28}$$

Mit Hilfe der Nachhallzeit  $T$  und dem Verhältnis zwischen Direkt- und Diffusschall erhält man aus dem linken Teilbild einen vorläufigen Wert. Dieser muss nun mit dem rechten Teilbild auf den vorhandenen Signal-Stör-Abstand korrigiert werden [22].

#### 4.3.6. Der Artikulationsindex AI

Das Konzept des Artikulationsindex wurde von French und Steinberg mit Hilfe einer Terz- oder Oktavanalyse entwickelt und besitzt Zahlenwerte zwischen 0 und 1. Je größer der AI ist, umso besser ist die Sprachverständlichkeit [23]. Zur Ermittlung des AI wird zunächst in jedem Frequenzband der gemessene RMS-gemittelte Störschallpegel vom Sprachspitzenschallpegel abgezogen. Dies ergibt die Terz- (Oktav-) dynamik. Mit diesen Dynamikwerten wird der relative AI berechnet:

$$AI_{rel} = g_{AI} \cdot \text{Terzdynamik}$$

bzw.

$$AI_{rel} = g_{AI} \cdot \text{Oktavdynamik} \quad \text{Gleichung 29}$$

Es muss jedoch darauf geachtet werden, dass nur Werte zwischen 0dB und 30dB verwendet werden. Werden diese Grenzen unter- bzw. überschritten, werden 0dB bzw. 30dB eingesetzt. Die entsprechenden Gewichtungsfaktoren können aus der folgenden Tabelle entnommen werden [22]:

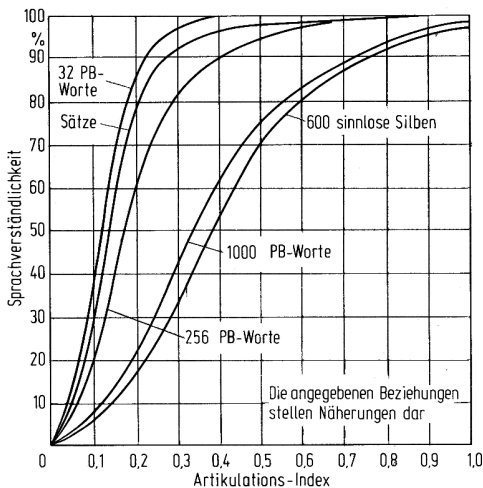
**Tabelle 4: Gewichtungsfaktoren**

Oktav-analyse:	$f_m = 250$	500	1000	2000	4000 [Hz]			
	$g_{AI} = 0,18$	0,5	0,75	1,07	0,83 [%]			
Terz-analyse:	$f_m = 200$	250	315	400	500	630	800 [Hz]	
	$g_{AI} = 0,03$	0,07	0,10	0,16	0,17	0,17	0,27 [%]	
	$f_m = 1,0$	1,25	1,6	2,0	2,5	3,15	4,0	5,0 [kHz]
	$g_{AI} = 0,30$	0,33	0,37	0,36	0,3	0,27	0,26	0,17 [%]

Das Endergebnis für den Artikulationsindex ergibt sich dann aus der Summe der Einzelwerte:

$$AI_{gesamt} = \sum AI_{rel} \quad \text{Gleichung 30}$$

Dabei gilt folgender Zusammenhang zur Sprachverständlichkeit:



$AI_{gesamt} < 35$	unbrauchbar
$35 - 45$	ausreichend
$45 - 60$	befriedigend
$60 - 80$	gut
$> 80\%$	sehr gut

**Abbildung 59: Zusammenhang AI - Verständlichkeit [23]**

Mit dem Artikulationsindex können Verständlichkeitsverluste berechnet werden, die durch nicht stark fluktuierende Störgeräusche oder durch Frequenzbandbegrenzungen hervorgerufen werden [22].

## 4.3.7. Sprachübertragungsindex STI und RASTI

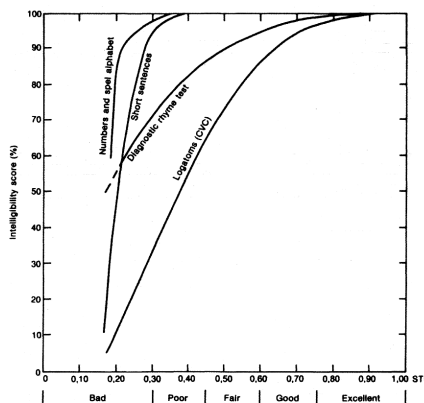
Einer der Gründe, warum die Sprachverständlichkeit reduziert wird, liegt darin, dass die übertragene zeitliche Hüllkurve (die „Modulation“) verfälscht übertragen wird und dass v.a. der Informationsgehalt (die Signaldynamik) schwacher Nutzsingale durch Störgeräusche oder durch Nachhall vorhergehender Signalanteile verringert wird [23].

Während der Artikulationsindex nur Störsignale berücksichtigt, bewerten Klarheitsmaß, Deutlichkeitsmaß und Konsonantenverlust nur Raumreflexionen. Der Sprachübertragungsindex STI (Speech Transmission Index) hingegen berücksichtigt beide Störgrößen und ermöglicht damit eine Klassifizierung der Übertragungsgüte durch *eine* Messgröße [23].

Der STI wird aus der Modulationsübertragungsfunktion MTF abgeleitet. Dabei wird oktavbreit gefiltertes Rosa Rauschen mit dem Amplitudenmodulationsgrad  $m = 1$  mit unterschiedlichen Modulationsfrequenzen sinusförmig amplitudenmoduliert. Der Pegel des gemittelten modulierten Oktavrauschens muss so gewählt werden, dass er dem mit gleichem Oktavfilter gemessenen mittleren Sprachpegel entspricht.

Mit den Oktav-Mittenfrequenzen 125Hz, 250Hz, 500Hz, 1kHz, 2kHz, 4kHz und 8kHz und den Modulationsfrequenzen 0,63Hz, 0,8Hz, 1Hz, 1,25Hz, 1,6Hz, 2Hz, 2,5Hz, 3,15Hz, 4Hz, 5Hz, 6,3Hz, 8Hz, 10Hz und 12,5Hz ergeben sich aus den Kombinationen insgesamt 98 Einzelmessungen. Bei jeder Messung wird ermittelt, um welchen Betrag der Amplitudenmodulationsgrad abgenommen hat. Diese Abnahme – die Modulationsreduktion  $m(f)$  – wird in einen scheinbaren Störabstand umgerechnet [23].

Dieser wird nun für alle 98 Messungen berechnet, gewichtet und aufsummiert. Das Endergebnis ist der Sprachübertragungsindex STI, für den folgende Klassifizierung für die Verständlichkeit gilt der Zusammenhang aus Abbildung 60.



**Abbildung 60: Zusammenhang STI – Verständlichkeit [22]**

$STI < 0,3$	unbrauchbar
$0,3 < STI < 0,45$	mangelhaft
$0,45 < STI < 0,6$	befriedigend
$0,6 < STI < 0,75$	gut
$0,75 < STI < 1$	ausgezeichnet

Für eine genauere Beschreibung des Verfahrens sei hier auf die entsprechende Fachliteratur verwiesen (z.B. 22, 23).

Die Messung des STI ist in der Praxis relativ aufwändig, v.a. wenn man bedenkt, dass oft mehrere Messungen durchgeführt werden, um eine Mittelung durchführen zu können. Deshalb wurde ein vereinfachtes Verfahren entwickelt, das jedoch auf dem gleichen Messprinzip beruht – das RASTI-Verfahren. Anstelle von 98 Messungen werden nun nur noch 9 Messungen in zwei Oktavbändern durchgeführt, die für die Sprachverständlichkeit besonders wichtig sind (500Hz und 2kHz) [22].

$$\begin{array}{ccccc} f_{\text{oktav}} / f_{\text{mod}} = & 2\text{k} / 0,7 & 500 / 1,0 & 2\text{k} / 1,4 & 500 / 2,0 & 2\text{k} / 2,8 \\ [\text{Hz}] / [\text{Hz}] & 500 / 4,0 & 2\text{k} / 5,6 & 500 / 8,0 & 2\text{k} / 11,2 & [23] \end{array}$$

Dem jeweiligen Oktavband werden dabei alle vier bzw. fünf Modulationsfrequenzen gleichzeitig aufmoduliert, wodurch sich kurze Messzeiten von ca. 8 bis 10s ergeben.

Für diese Frequenzen wird nun  $m(f)$  berechnet und daraus die scheinbaren Signal-Störabstände  $X_i$  bestimmt. Mit diesen ergibt sich der RASTI-Wert aus der arithmetischen Mittelung [22].

$$RASTI = \frac{1}{9} \cdot \sum_{i=1}^9 \frac{X_i + 15}{30} \cdot 100 \quad [\%] \quad \text{Gleichung 31}$$

Dabei muss beachtet werden, dass der scheinbare Signal-Stör-Abstand maximal  $\pm 15\text{dB}$  betragen darf.

Ein Problem bei diesem vereinfachten Messverfahren liegt darin, dass eben nur bei 500Hz und 2kHz gemessen wird, die Sprachübertragung aber in einem Bereich von 100Hz bis 8kHz von Bedeutung ist. Deshalb müssen RASTI-Werte immer mit Vorsicht interpretiert werden [22].

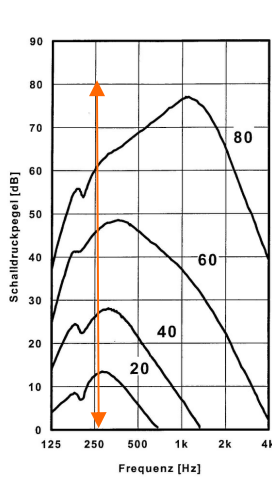


## 5. Versuchsplanung

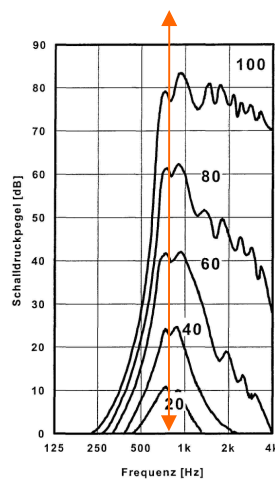
### 5.1. Ausgangslage

Bevor der Aufbau des Messsystems erläutert wird, soll nochmals kurz auf die Ausgangslage und die zu lösende Problematik eingegangen werden:

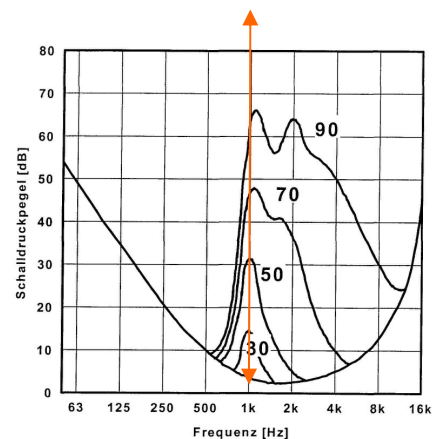
"Die stärkste verdeckende Wirkung zeigen intensive Töne (mit einem Schalldruckpegel über 80dB) niedriger Frequenz. Dabei übertönt der verdeckende Ton alle Töne mit höherer Frequenz viel stärker als die mit niedrigerer Frequenz" [19]. Zu dieser Aussage veröffentlichte Slawin folgende gemessene Mithörschwellen (Abb. 61 und 62), in die zusätzlich das Störsignal mit einem Pfeil eingetragen wurden:



**Abbildung 61:**  
Störton 200Hz [19]



**Abbildung 62:**  
Störton 800Hz [19]



**Abbildung 63:**  
Störton 1kHz [11]

Zwicker veröffentlichte ebenfalls gemessene Mithörschwellen bei Sinusstörtönen (Abb. 63). Deutlich erkennbar ist der Unterschied der ermittelten Kurven, besonders zwischen Abbildung 61 und 63. Während bei Slawin die Mithörschwelle oberhalb des Störsignals weiter stark ansteigt und erst im kHz-Bereich sein Maximum erreicht, besitzt die Kurve von Zwicker genau im Bereich des Maskierers seinen höchsten Wert. Zwar steigt sie nochmals an, dies jedoch nur bei sehr hohen Störpegeln, während das Verhalten bei Slawin bei allen Störpegeln ähnlich ausfällt. Zusätzlich ist zu beobachten, dass bei Slawin, um das Testsignal zu hören, ein viel höherer Pegel eingestellt werden muss als bei Zwicker, obwohl der Störpegel geringer ist. Beide, sowohl Slawin als auch Zwicker, geben jedoch als Störsignal einen Ton an.

Wie kann dieser Unterschied erklärt werden?

Zusätzlich wurde aus den obigen Kurven folgende Vermutung abgeleitet:

Durch die Verdeckungseffekte sind starke Beeinträchtigungen der Sprachverständlichkeit zu erwarten, da genau im für Kommunikation wichtigen Frequenzbereich (800Hz bis 2500Hz) [19] starke Verdeckungserscheinungen auftreten. In welchem Maße können dadurch entstehende Probleme und Effekte (z.B. tieffrequentes Dröhnen [25]) in Bezug auf die Sprachverständlichkeit wissenschaftlich nachgewiesen und begründet werden?

Diese beiden aufgeworfenen Fragen sollten mit Hilfe geeigneter Hilfsmittel und Messmethoden geklärt werden. Dafür wurden zunächst Messungen von Mithörschwellen durch unterschiedliche Störsignale durchgeführt. Anschließend sollten, ebenfalls mit verschiedenen Störungen, Silbenverständlichkeitstests Aufschluss über die oben aufgeworfenen Probleme geben.

## 5.2. Aufbau für die Hörschwellenmessungen

### 5.2.1. Anforderungen an die Messumgebung

Die Umgebung, in der gemessen werden sollte, musste strenge Kriterien erfüllen. Da bei diesen Tests mit Sinussignalen gearbeitet wurde, war ein freies Schallfeld zwingend vorgegeben [3].

In diesem wurde zunächst der Grundgeräuschpegel kontrolliert, der bei komplett eingeschaltetem Messsystem (Computer und Verstärker) vorhanden war. Es durfte auch teilweise auftretender Umgebungslärm (Lastwagen, Sägen im Nachbarlabor usw.) nicht bemerkbar werden. Zusätzlich musste überprüft werden, ob die notwendigen Anforderungen an das Schallfeld [3] eingehalten wurden. Diese lauten:

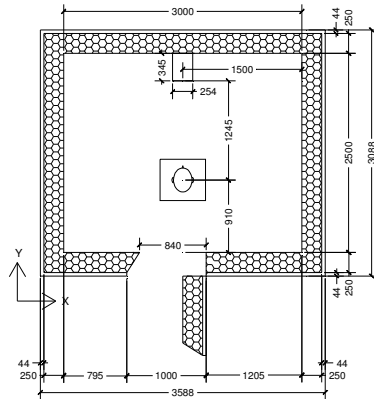
Wenn Proband und der Probandenstuhl aus dem Schallfeld entfernt sind, dürfen sich die vom Lautsprecher erzeugten Schalldruckpegel an den Positionen 0,15m vom Bezugspunkt auf diesen Achsen um folgende Maximalwerte unterscheiden:

- Achse links – rechts:  $f \leq 4\text{kHz}$   $\pm 1\text{dB}$
- Achse oben – unten:  $f \leq 4\text{kHz}$   $\pm 1\text{dB}$
- Achse vorne – hinten:  $f \leq 4\text{kHz}$   $\pm 1\text{dB}$  vom theoretischen Wert nach dem inversen Abstandsgesetz

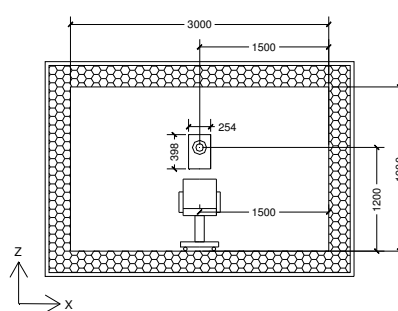
Diese Anforderungen können nur in einem reflexionsarmen Raum erfüllt werden.

### 5.2.2. Messumgebung Audiokabine

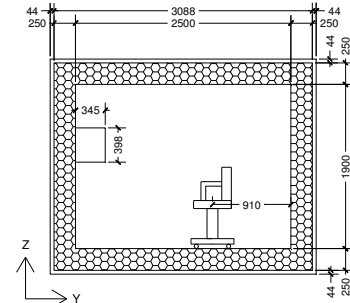
Im Rahmen einer Diplomarbeit wurde im Jahr 2000 eine reflexionsarme Audiokabine entwickelt und gebaut, die speziell für solche audiometrischen Untersuchungen genutzt werden soll [44]. Diese besitzt folgenden Aufbau:



**Abbildung 64:**  
Grundriss der Audiokabine



**Abbildung 65:**  
Seitenansicht X-Z



**Abbildung 66:**  
Seitenansicht Y-Z

Zur reflexionsarmen Gestaltung der Wandflächen sind diese komplett mit Absorbentien ausgestattet. Diese Absorbentien sind im gesamten Hörbereich wirksam, also auch bei tiefen Frequenzen. In den Ecken werden Breitband-Kompakt-Absorbentien (BKA) [26] dazu verwendet tieffrequente Raummoden<sup>4</sup> zu bedämpfen.

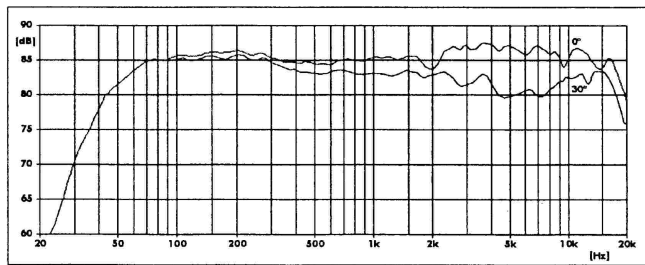
Diese Kabine erfüllte in früheren Messungen (damalige Untersuchungen im Diffusfeld) die in den Normen [27] geforderten Auflagen bezüglich der Isotropie<sup>5</sup> sehr gut. Allerdings konnten die Homogenitätsanforderung nicht ganz eingehalten werden. Im Bereich der Übernahmefrequenz des eingebauten Zweiwegelautsprechers kam es zu Pegeldifferenzen im Bereich von  $\pm 4,5\text{dB}$ . Aus diesem Grund wurde im jetzigen Aufbau (Freifeldsimulation) von Anfang an ein koaxialer Lautsprecher eingebaut. Dieser wurde in Kopfhöhe des sitzenden Probanden, also auf ca. 120cm Höhe abgehängt. Dabei lief die Bezugsachse durch den Bezugspunkt [3]. Der Abstand zwischen Lautsprecher und Proband betrug 125cm.

<sup>4</sup> Raummoden treten dann auf, wenn eine halbe Wellenlänge oder ganzzahlige Vielfache davon mit mindestens einer Raumdimension übereinstimmen. Die Folge davon ist eine starke Ortsabhängigkeit der Schalldruckpegel im Raum.

<sup>5</sup> Wird das Schallfeld am Referenzpunkt in 18 verschiedene Richtungen gemessen, darf die maximale Pegeldifferenz nicht mehr als 5dB betragen.

### 5.2.3. Verstärker und Lautsprecher

Der verwendete koaxiale Lautsprecher *ME100* der Firma *musikelectronic geithain GmbH* besitzt die Abmessungen von 39,8 x 25,4 x 34,5 cm. Der maximale Schalldruckpegel in 1m Entfernung beträgt 106dB, wobei der Übertragungsbereich von 50Hz bis 20kHz einen Frequenzgang mit  $\pm 3\text{dB}$  aufweist (Abb. 67). Damit besitzt dieser Lautsprecher die idealen Voraussetzungen für die Messungen.



**Abbildung 67: Freifeldübertragungskurve des ME100 [28]**



**Abbildung 68: Sicherung der Verstärkereinstellung**

Zur Verstärkung der Signale vor der Lautsprecherwiedergabe wurde eine Endstufe *SA 240* der Firma *Klein&Hummel* eingesetzt. Diese besitzt einen sehr hohen Signal-Rauschabstand von 113dB. Die Aussteuerung des Verstärkers wurde einmal bei der Kalibrierung der Signale eingestellt und anschließend nicht mehr verändert.

### 5.2.4. Computer zur Software-Steuerung

Um exakte Messergebnisse und auch die Reproduzierbarkeit der Ergebnisse zu sichern, wurden die Hörschwellentests computerunterstützt durchgeführt. Im Versuch wurde dafür ein Pentium II mit 133MHz verwendet. Der Arbeitsspeicher betrug 128MB. Das dafür entwickelte Messprogramm selbst wird im folgenden Kapitel genauer vorgestellt.

Um einen genügend großen Dynamikbereich nutzen zu können, musste im Computer eine qualitativ hochwertige Soundkarte eingebaut sein – im aktuellen Fall die Soundkarte *Siena* von *SEK'D*. Diese besitzt acht getrennte Mono- bzw. vier getrennte Stereoausgänge. Die möglichen Sample-Raten liegen zwischen 11kHz und 96kHz. Der Signal-Rausch-Abstand der Karte beträgt 95dB und der AD/DA-Wandler arbeitet mit 24Bit.

Beim Aufbau des Messsystems stellte sich jedoch heraus, dass es mit dieser Soundkarte nicht möglich ist, die Ausgänge unabhängig voneinander ein- und auszuschalten. Wenn das Test-

signal (Dauer 1s) auf Kanal 1 abgeschaltet wurde, verschwand gleichzeitig das Störsignal, das dauerhaft von Kanal 3 abgespielt werden sollte.

Dieses Problem war mit der zur Verfügung stehenden Soundkarte nicht zu lösen, so dass zum Abspielen der Störsignale ein zweites Abspielgerät benutzt werden musste. Da die Störsignale zu diesem Zeitpunkt bereits als Datei für den Computerbetrieb vorlagen, wurde ein zweiter PC benutzt. Auch die hier benützte Soundkarte *EWS 64 XL* der Firma *Terratec Promedia* wies gute technische Daten auf: Sie besitzt zwei analoge Stereo-Ausgänge, sowie zwei digitale Ausgänge nach S/PDIF Standard. Der Signal-Rauschabstand beträgt, nahezu gleich gut wie bei der Siena-Soundkarte, 93dB. Die maximale Samplerate liegt bei 48kHz, der AD/DA-Wandler arbeitet mit 16Bit. Damit ist die Auflösung der Rausch-Störsignale auf 16Bit beschränkt, was jedoch kein Problem darstellte, da diese Signale alle relativ laut abgespielt wurden.



**Abbildung 69: Aufbau der Geräte im Vorraum des Freifeldraums**

Die Aufzeichnung der Probandenreaktionen auf die Testsignale wurde mit einem Joypad durchgeführt. Dieses (*Saitek P 120*), das per Tastendruck bedient wurde, lieferte entsprechende Signale an das Messprogramm. Dafür war im "Testsignal-Computer" eine zusätzliche Joystickkarte integriert.

#### 5.2.5. Messgeräte

Um alle Pegel exakt kalibrieren zu können, wurde ein entsprechendes Messgerät verwendet: der *Real Time Analytator RTA 840* (Version 1.01) der Firma *Norsonic*. Die damit bestimmten Werte konnten abgespeichert und weiterverarbeitet werden.

Als Messmikrofon wurde ein  $\frac{1}{2}$ "-Freifeld-Mikrofon, *Type 4190* mit dem dazugehörigen *Vorverstärker Type 2619* benutzt. Sollten besonders niederpegelige Werte unterhalb 0dB gemessen werden, wurde das hochempfindliche 1"-Kondensator-Mikrofon *Type 4179* mit dem Vor-

verstärker Type 2660 verwendet. Vor jeder Messung wurden diese Mikrofone mit dem *Sound Level Calibrator, Type 4231* kalibriert.

Alle Mikrofone, Mikrofonvorverstärker und der Kalibrator sind Geräte der Firma *Brüel&Kjaer*.

### 5.3. Mess-Signale für die Hörschwellentests

Bei den durchgeführten Versuchen wurden sowohl Ruhehörschwellen als auch Mithörschwellen bestimmt. Daher mussten entsprechende Nutz- und Störsignale generiert und kalibriert werden.

#### 5.3.1. Nutzsignale

Als Testsignale, mit denen die Hörschwellen der Testpersonen bestimmt werden sollten, wurden reine Sinustöne verwendet. Die Frequenzen reichten dabei von 40Hz bis 4kHz mit einer Schrittweite von jeweils einer Terz. Bei 4kHz wurde die Messung abgebrochen, da in diesem Bereich bereits alle erwarteten Effekte aufgetreten sein sollten. Zusätzlich konnten eventuell auftretende Schwierigkeiten mit der Einhaltung der in der Norm [3] geforderten Schallfeldeigenschaften umgangen werden.

Damit ergaben sich folgende Prüffrequenzen (Angabe jeweils in Hz):

// 40 / 50 / 63 / 80 / 100 / 125 / 160 / 200 / 250 / 315 /  
/ 400 / 500 / 630 / 800 / 1k / 1,25k / 1,6k / 2k / 2,5k / 3,15k / 4k //

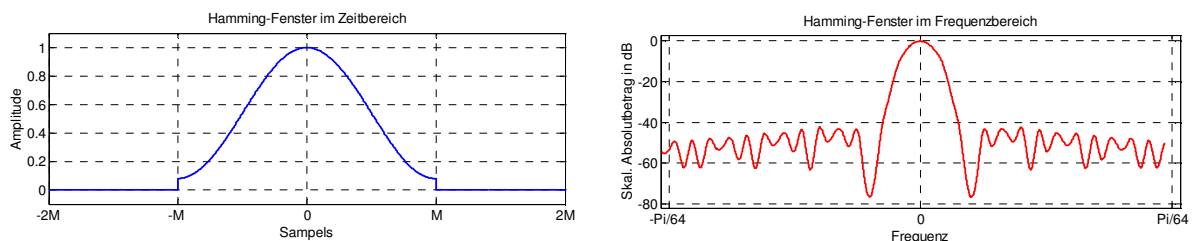
Für alle Prüffrequenzen mussten die Testsignale mit Pegeln zwischen -10dB und +84dB mit einer Schrittweite von 1dB zur Verfügung gestellt werden (insgesamt also  $95 \cdot 21 = 1995$  einzelne Signale). Um diesen hohen Dynamikbereich realisierbar zu machen, wurden die Signale mit 24Bit aufgelöst, d.h. es standen insgesamt  $2^{24}$  Stufen zur Verfügung. Mit dieser Auflösung wurde ein theoretischer Dynamikbereich von  $24 \cdot 6,02 = 144dB$  erzielt. Dadurch war es mit Hilfe des Bearbeitungsprogramms *Samplitude*<sup>6</sup> möglich, die Pegel in 1dB-Schritten um insgesamt 50dB zu reduzieren. 50dB entsprechen 9Bit ( $50/6,02$ ), es standen also immer noch genügend Stufen ( $24 - 9 = 15\text{Bit} \rightarrow 2^{15}$  Stufen) für die Signaldarstellung zur Verfügung.

---

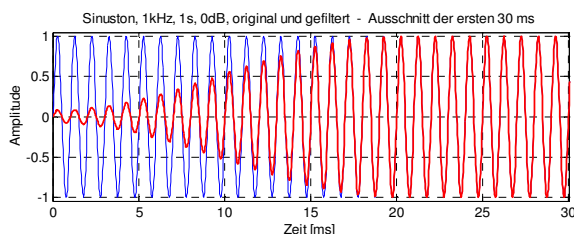
<sup>6</sup> *Samplitude classic 7.0* der Firma *magix*

Die Signale mit hohen Pegeln wurden dann bei Vollaussteuerung des Verstärkers wiedergegeben. Um die leiseren Signale zu erhalten, wurden die gleichen Audiodateien wie für laute Pegel abgespielt, dabei aber der hardwareseitige Verstärker um 40dB gedämpft. Dadurch wurde auch das Rauschen reduziert. Ein weiterer Vorteil dieser Signal-Verstärker-Kombination bestand darin, dass nur knapp die Hälfte der benötigten Nutzsignale kalibriert werden musste. Bei den restlichen Signalen wurde lediglich die hardwareseitige Verstärkung um einen konstanten Betrag reduziert, so dass hier einige Stichproben zur Kontrolle ausreichten.

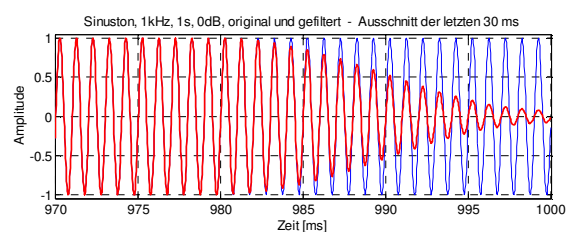
Die Testsignale waren ca. 1s lang. Um beim Ein- und Ausschalten Knackser zu vermeiden, wurde der Sinuston mit einem Hamming-Fenster (Abb.70) jeweils 20ms ein- und ausgeblendet (Abb. 71 und 72).



**Abbildung 70: Darstellung des Hamming-Fensters**



**Abbildung 71: Eingeblandetes Signal**



**Abbildung 72: Ausgeblendeter Testton**

### 5.3.2. Störsignale

Als Störsignale waren sowohl konventionelle Rauschsignale (werden ab jetzt als Rauschen bezeichnet) als auch Low-noise noise (Bezeichnung LNN) geplant.

Bei den Rauschsignalen wurde zwischen Terzrauschen (Terzgefiltertes Rosa Rauschen) und Tiefpassrauschen unterschieden.

Das Terzrauschen mit den Mittenfrequenzen 63Hz, 125Hz, 250Hz und 500Hz wurde auf die Pegel 30dB, 50dB und 70dB eingemessen. Bei den Mittenfrequenzen von 63Hz und 125Hz

wurde der Pegelbereich noch nach oben auf 80dB erweitert, bei den Mittenfrequenzen 250Hz und 500Hz nach unten auf 20dB. Damit lagen alle Störpegel über der Ruhehörschwelle.

Beim Tiefpassrauschen wurden die Grenzfrequenzen auf 100Hz, 200Hz und 510Hz (also eine, zwei und fünf Frequenzgruppenbreiten) festgesetzt. Die Pegel betragen dabei 35dB, 50dB und 70dB, bei der 100Hz-Grenzfrequenz zusätzlich 80dB. Bei diesem Signaltyp musste der Unterschied zwischen Schallpegel und Schallintensitätsdichtepegel (Dichtepegel) beachtet werden (siehe Punkt 3.6.1. *Schalldruckpegel und Dichtepegel*). Um vergleichbare Ergebnisse zu erhalten, wurde hier der Dichtepegel konstant gehalten [11]. Mit den Gleichungen 2, 3 und 4 konnten die tatsächlich benötigten Pegelwerte berechnet werden:

$$L_{100,35} = 35\text{dB} = [l_{\text{WR}} + 10 \cdot \lg(f_o - f_u)] \text{ dB} = [l_{\text{WR}} + 10 \cdot \lg(100\text{Hz} - 20\text{Hz})] \text{ dB}$$

$$L_{100,35} = [l_{\text{WR}} + 19.0] \text{ dB}$$

$$l_{\text{WR},35} = 35\text{dB} - 19\text{dB} = 16\text{dB}$$

$$L_{200,35} = 16\text{dB} + 10 \cdot \lg(200\text{Hz} - 20\text{Hz}) \text{ dB} = 16\text{dB} + 22.5\text{dB} = 38.5\text{dB}$$

$$L_{510,35} = 16\text{dB} + 10 \cdot \lg(510\text{Hz} - 20\text{Hz}) \text{ dB} = 16\text{dB} + 26.9\text{dB} = 42.9\text{dB}$$

$$L_{100,50} = 50\text{dB} = l_{\text{WR}} + 19.0\text{dB}$$

$$\rightarrow l_{\text{WR},50} = 50\text{dB} - 19\text{dB} = 31\text{dB}$$

$$\rightarrow L_{200,50} = 31\text{dB} + 22.5\text{dB} = 53.5\text{dB}$$

$$\rightarrow L_{510,50} = 31\text{dB} + 26.9\text{dB} = 57.9\text{dB}$$

$$L_{100,70} = 70\text{dB} = l_{\text{WR}} + 19.0\text{dB}$$

$$\rightarrow l_{\text{WR},70} = 70\text{dB} - 19\text{dB} = 51\text{dB}$$

$$\rightarrow L_{200,70} = 51\text{dB} + 22.5\text{dB} = 73.5\text{dB}$$

$$\rightarrow L_{510,70} = 51\text{dB} + 26.9\text{dB} = 77.9\text{dB}$$

$$L_{100,80} = 80\text{dB} = l_{\text{WR}} + 19.0\text{dB}$$

$$\rightarrow l_{\text{WR},80} = 80\text{dB} - 19\text{dB} = 61\text{dB}$$

$$\rightarrow L_{200,80} = 61\text{dB} + 22.5\text{dB} = 83.5\text{dB}$$

$$\rightarrow L_{510,80} = 61\text{dB} + 26.9\text{dB} = 87.9\text{dB}$$

Zur Anregung musste Gleichmäßig Verdeckendes Rauschen benutzt werden. Da dieses aber bei tiefen Frequenzen (näherungsweise bis 510Hz) gleich Weißem Rauschen ist, konnte der Dämpfungsfaktor  $a_{\text{GVR}}$  (siehe Punkt 3.6.2. *Breitbandrauschen*) vernachlässigt und Weißes Rauschen als Signaltyp verwendet werden.

Alle Rauschsignale wurden vollständig digital mit dem Programm *CoolEdit Pro* (Version 1.2) der *SyntrilliumSoftware Corporation* erzeugt. Die notwendige Filterordnung wurde berechnet [27].



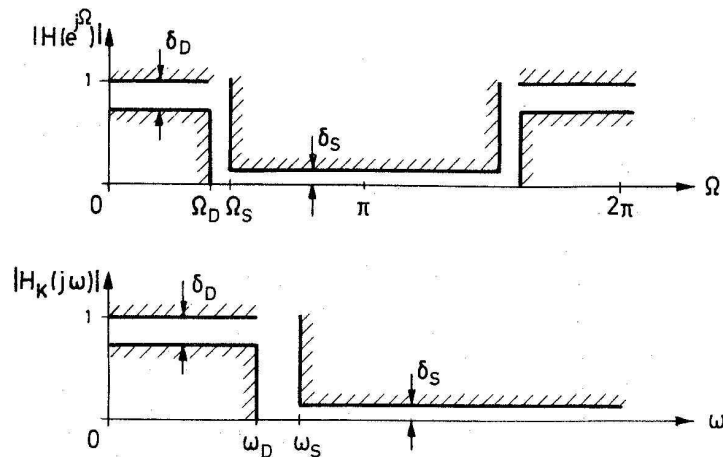


Abbildung 73: Toleranzschema für Filterdesign [30]

Das Toleranzschema (Abb. 73) wird durch die folgenden Parameter festgelegt:

$f_D = 194\text{Hz}$ ,  $f_S = 206\text{Hz}$ ,  $\delta_D = 0,1$  und  $\delta_S = 0,1$ . Für die Abtastfrequenz gilt  $f_A = 44100\text{Hz}$ .

Daraus wird die normierte Grenzfrequenz  $\Omega_S$  und  $\Omega_D$  bestimmt:

$$\Omega_D = \frac{2\pi \cdot f_D}{f_A} = 0,0276 \quad \Omega_S = \frac{2\pi \cdot f_S}{f_A} = 0,0294 \quad \text{Gleichung 32}$$

Dieses Toleranzschema muss nun in Hinblick auf den Entwurf eines kontinuierlichen Systems vorverzerrt werden:

$$\omega_D = \tan(\Omega_D/2) = 0,0138 \quad \omega_S = \tan(\Omega_S/2) = 0,0147 \quad \text{Gleichung 33}$$

Normiert auf die Durchlass-Grenzfrequenz ergibt sich

$$\omega'_D = 1 \quad \omega'_S = \omega_S / \omega_D = 1,0619 \quad \text{Gleichung 34}$$

Damit lässt sich die erforderliche Filterordnung  $n$  festlegen:

$$n \geq \frac{\lg(\Delta_2/\Delta_1)}{\lg(\omega'_S)} \quad \text{Gleichung 35}$$

mit

$$\Delta_1 = \frac{\sqrt{2\delta_D - \delta_D^2}}{1 - \delta_D} \quad \text{und} \quad \Delta_2 = \frac{\sqrt{1 - \delta_S^2}}{\delta_S} \quad \text{Gleichung 36}$$

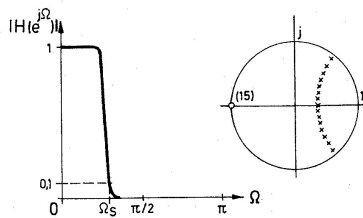
Mit den oben angegebenen Werten ergaben sich daraus folgende Ergebnisse:

$$n_{\text{Butterworth}} = 51$$

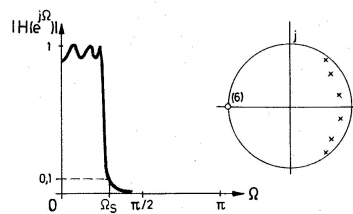
$$n_{\text{Tschebyscheff}_1} = 11$$

$$n_{\text{Tschebyscheff}_2} = 11$$

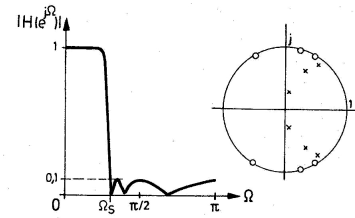
Wie gut sichtbar, benötigt man beim Butterworth-Filter eine wesentlich höhere Filterordnung als beim Tschebyscheff-Filter, da der Tschebyscheff-Filter gegenüber dem Butterworth-Filter einen viel steileren Übergang vom Durchlass- in den Sperrbereich besitzt (bei gleicher Filterordnung  $n$ ) [30].



**Abbildung 74:**  
Butterworth-Filter,  
 $n = 15$  [30]



**Abbildung 75:**  
Tschebyscheff Typ I,  
 $n = 6$  [30]



**Abbildung 76:**  
Tschebyscheff Typ II,  
 $n = 6$  [30]

Aus diesem Grund wurde ein Tschebyscheff-Filter vom Typ II (Abbildung 76) mit der Filterordnung  $n = 11$  (siehe Berechnungen oben) für die Generierung der Störsignale verwendet.

Die LNN-Signale wurden äquivalent zu den Rausch-Signalen behandelt. Auch hier wurde zwischen Terz- und Tiefpassrauschen unterschieden. Das Terzrauschen wurde mit den Mittenfrequenzen 125Hz und 250Hz generiert, beim Tiefpassrauschen wurden die Grenzfrequenzen 100Hz und 200Hz gewählt. Die Pegel wurden (wie bei den herkömmlichen Rauschsignalen) jeweils auf 30dB und 50dB festgelegt, um die Ergebnisse mit den "normalen Rausch-Ergebnissen" vergleichen zu können.

Diese Testgeräusche wurden nach den Formeln 8 und folgende und mit den Werten aus den Tabellen der Phasenwinkel von Pumplin und Hartmann [15] mit *MATLAB*<sup>7</sup> erzeugt. Um dabei den obigen Signalen im Spektralbereich möglichst ähnlich zu werden, wurden folgende Einstellungen für die Generierung gewählt:

<sup>7</sup> *MATLAB* Version 6.5, Release 13 der Firma *MathWorks, Inc.*

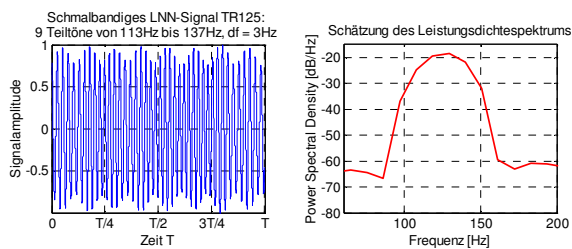
Terzrauschen,  $f_m = 125\text{Hz}$ : 9 Teiltöne,  $d_f = 3\text{Hz}$ , Frequenzbereich von 113Hz bis 137Hz

Terzrauschen,  $f_m = 250\text{Hz}$ : 21 Teiltöne,  $d_f = 3\text{Hz}$ , Frequenzbereich von 220Hz bis 280Hz

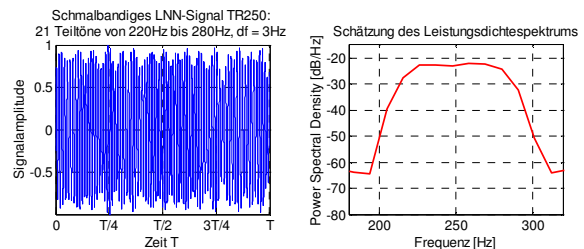
Tiefpassrauschen,  $f_g = 100\text{Hz}$ : 31 Teiltöne,  $d_f = 3\text{Hz}$ , Frequenzbereich 3Hz bis 93Hz

Tiefpassrauschen,  $f_g = 200\text{Hz}$ : 60 Teiltöne,  $d_f = 3\text{Hz}$ , Frequenzbereich 3Hz bis 180Hz

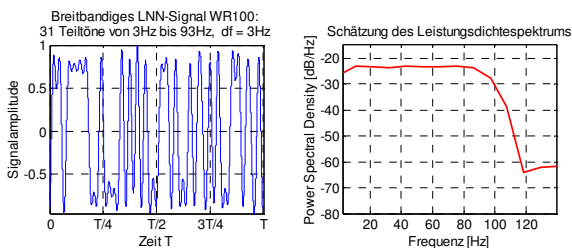
Damit ergaben sich folgende Signalformen:



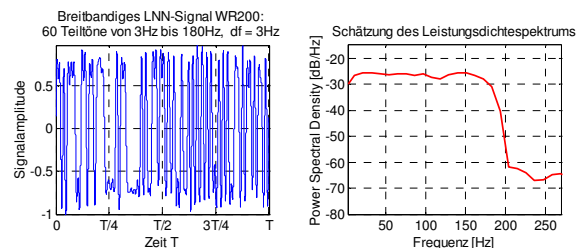
**Abbildung 77: LNN-Terzrauschen  
mit  $f_m = 125\text{Hz}$**



**Abbildung 78: LNN-Terzrauschen  
mit  $f_m = 250\text{Hz}$**

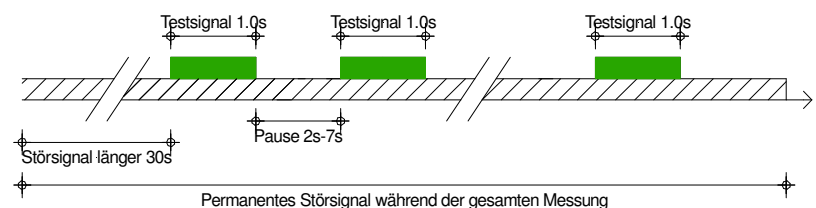


**Abbildung 79: LNN-Tiefpaßrauschen  
mit  $f_g = 100\text{Hz}$**



**Abbildung 80: LNN-Tiefpaßrauschen  
mit  $f_g = 200\text{Hz}$**

Sämtliche Störgeräusche wurden bei den Testverfahren ständig im Hintergrund abgespielt.



**Abbildung 81: Darstellung des zeitlichen Messablaufs**

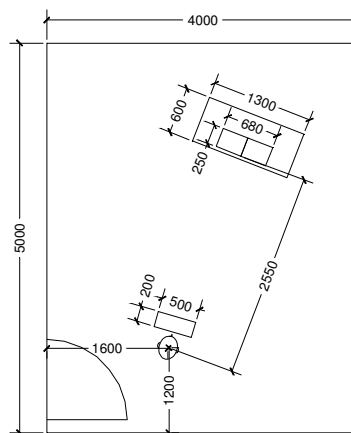
Dadurch war eine Störsignallänge von ca. 25 – 30 min notwendig. Die kürzer generierten Signale wurden dafür entsprechend oft geloopt, wobei beachtet wurde, dass diese Wiederholungen von den Testpersonen nicht bemerkt werden konnten.

## 5.4. Messaufbau für die Verständlichkeitstests

Als Verständlichkeitstest wurde ein Logatomtest mit jeweils 50 sinnlosen Einsilbern geplant. Um auch hier ein zuverlässiges Ergebnis zu erhalten, mussten mehr Testpersonen als bei den Hörschwellenversuchen eingeplant werden, da die subjektive Verständlichkeit vorgesprochener Silben sehr viel unterschiedlicher ausfällt als Effekte bei Hörschwellenmessungen.

### 5.4.1. Messumgebung Tieffrequenz-Labor

Als Messumgebung wurde das Tieffrequenz-Labor (T-Labor) [31] gewählt. Damit war eine definierte Umgebung gewährleistet. Gleichzeitig war es möglich, durch einen einfachen und schnellen Umbau von verschiedenen Absorberplatten eine neue Umgebungssituation zu simulieren. Somit konnten mehrere Messumgebungen in einem Raum realisiert werden.

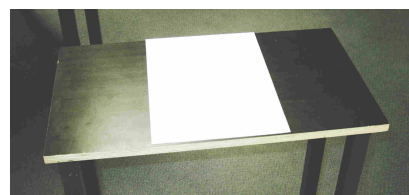


**Abbildung 82: Grundriss T-Labor**

Lautsprecher und Proband wurden wie in Abbildung 82 abgebildet im Raum platziert. Dabei wurden zwei Lautsprecher verwendet: Von einem wurde das Störsignal abgespielt, vom zweiten die Logatomlisten (Nutzsignal). Beide wurden möglichst nah zueinander aufgestellt, da der ungünstigste Beschallungsfall dann vorliegt, wenn sich die Nutz- und Störschallquelle an der gleichen Position befinden (siehe 2.3.4. *Der Cocktailparty-Effekt*). Dadurch sind keine räumlichen Unterscheidungsmerkmale der Schallquellen vorhanden [18]. Dieser Fall sollte untersucht werden.



**Abbildung 83: Lautsprecheraufstellung**



**Abbildung 84: Schreibtischchen für Probanden**

Der (effektive) Hallradius  $r_H$  ist definiert mit  $r_H = 0,057 \cdot \sqrt{\frac{V}{T}}$  bzw.  $r_{H,eff} = 0,057 \cdot \sqrt{\gamma} \cdot \sqrt{\frac{V}{T}}$ .

Dabei sind für  $V$  das Raumvolumen [ $m^3$ ], für  $T$  die Nachhallzeit [s] und für  $\gamma$  der Bündelungsgrad der Schallquelle einzusetzen. Im T-Labor ergaben sich die Größenordnungen:

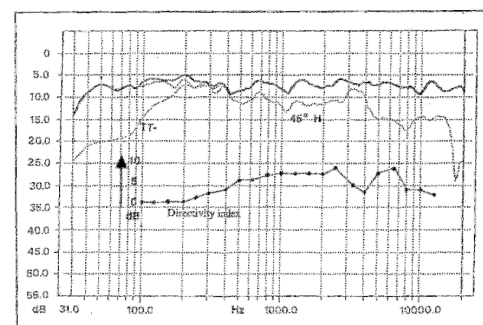
$$\begin{array}{llll} r_{H,Leer} = 0,23m & r_{H,VPR} = 0,56m & r_{H,BKA} = 0,80m & r_{H,Caruso} = 0,65m \\ r_{H,Leer,eff} = 0,54m & r_{H,VPR,eff} = 1,31m & r_{H,BKA,eff} = 1,86m & r_{H,Caruso,eff} = 1,50m \end{array}$$

Mit einem Abstand von 255cm zum Lautsprecher wurde der Proband außerhalb des (effektiven) Hallradius platziert. Es wurde von ihm verlangt, die gehörten Silben aufzuschreiben. Um dabei zu verhindern, dass sich die Testperson beim Schreiben mit einem Block auf den Knien zusammenkrümmt, und sich der Kopf dadurch nicht mehr in einer Höhe mit den Lautsprechern (120cm Höhe) befindet, wurde ein kleines Tischchen zur Verfügung gestellt (Abb. 84), dessen Schallfeldbeeinflussung vernachlässigt werden konnte.

#### 5.4.2. Signaldarbietung

Um den Aufbau für die Hörschwellenmessung bei Durchführung der Sprachverständlichkeitstests noch nicht zu zerstören (sind Lautsprecher und Probandenstuhl aus der Messumgebung entfernt, müssten für eine evtl. zusätzlich benötigte Messung alle Signale neu kalibriert werden), wurden für diese Untersuchungen andere Lautsprecher verwendet.

Die hier verwendeten Boxen vom Typ *RL904* stammten zwar auch von der Firma *musikelectronic geithain*, waren aber aktiv ausgeführt. Sie besaßen die Gehäuseabmessungen von 39 x 25 x 34 cm und einen maximalen Schalldruckpegel von 106dB in 1m Entfernung. Der Übertragungsbereich reichte von 40Hz bis 20kHz mit einem Frequenzgang von  $\pm 3$ dB und lieferte damit ideale Eigenschaften. Durch die aktive Ausführung wurde kein zusätzlicher Verstärker benötigt, die Lautsprecher konnten direkt an die Wiedergabegeräte angeschlossen werden.



**Abbildung 85: Frequenzgang des RL904 bei  $r = 2m$  [29]**

Dafür wurde für die Testsignale – die Logatomlisten – ein DAT-Recorder der Firma *Sony* benutzt, für die Störsignale der "Störsignal-Computer" der Hörschwellenmessung.

Um sicherzustellen, dass alle Signale in jeder Raumkonfiguration den Testpersonen immer gleichlaut vorgespielt werden, wurden die Schalldruckpegel an der Hörposition mit dem *RTA 840* (Firma *Norsonic*) bestimmt, zur Messung wurde ein  $\frac{1}{2}$ "-Mikrofon vom *Type 4190* der Firma *Briuel&Kjaer* verwendet.

#### 5.4.3. Geplante Raumauskleidungen

Bei den Sprachverständlichkeitstests mit Logatomen waren vier Raumvarianten geplant.

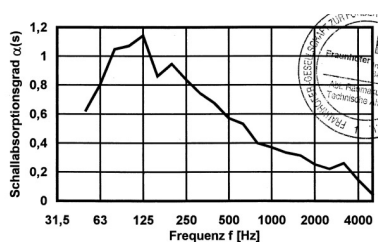
Zunächst sollte im unbehandelten Raum der hallige Ausgangszustand untersucht werden.

Dann wurden für einen ersten Versuch spezielle Tiefenabsorber eingebaut. Diese Verbund-Platten-Resonatoren (VPR) (Abb. 86) sind vor allem bis 500Hz sehr gut wirksam [32].

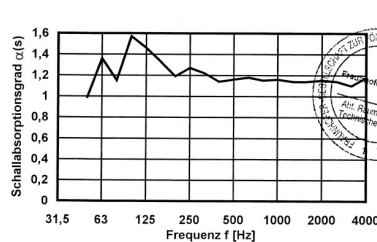
In diesen Versuchen wurden dafür 18 VPR-Module mit einer jeweiligen Fläche von  $1,5\text{m}^2$  und einer Tiefe von 10cm in den Raum eingebracht. Diese Anzahl wurde im Voraus mit einer Software des Fraunhofer-Instituts speziell für solche Abschätzungen berechnet. Die Befestigung wurde mit Rechteckwinkeln ausgeführt.

Dem gegenüber standen Versuche mit Breitband-Kompakt-Absorbern (BKA) [26]. Diese setzen sich aus VPR-Modulen mit einem darüber aufgebracht Höhenabsorber zusammen und liefern über den gesamten Frequenzbereich sehr hohe Absorptionswerte (Abb. 87).

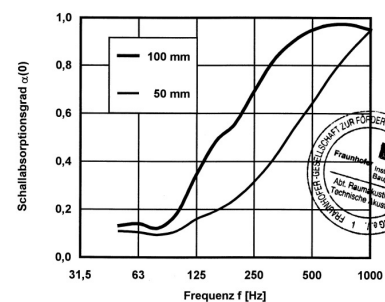
Im Versuch wurden diese BKAs realisiert, indem auf die VPR-Absorber Caruso-ISO-Bond-Absorptionsplatten (Polyfasermaterial mit einem Wirkungsbereich ab 500Hz aufwärts) aufgeklebt wurden. Diese besitzen eine Tiefe von 5cm und eine Dichte von  $20\text{kg/m}^3$ .



**Abbildung 86:**  
Absorptionskurve VPR [29]



**Abbildung 87:**  
Absorptionskurve BKA [30]



**Abbildung 88:**  
Absorptionskurve Caruso [31]

Als vierte Auskleidung wurde der Raum mit herkömmlichen Caruso-Absorptionsplatten ausgestattet. Diese besitzen ebenfalls eine Grundfläche von  $1,5\text{m}^2$ , jedoch eine Tiefe von nur 5cm. Die Dichte dieser Platten beträgt  $40\text{kg/m}^3$ , dadurch wirken diese Absorber ab 250Hz sehr gut (siehe Abb. 88). Da die Anschraub-Winkel auf die Dicke der VPR-Module von 10cm ausgelegt waren, konnten sie in diesem Fall nicht verwendet werden. Deshalb wurden die

Caruso-Platten mit doppelseitigem Klebeband so an der Wand befestigt, dass sich dahinter kein Luftvolumen bilden konnte.

## 5.5. Test-Signale für Verständlichkeitsuntersuchungen

### 5.3.1. Logatomlisten

Die Logatome, die die Testpersonen bei den Versuchen verstehen sollten, mussten bei jeder Testperson exakt gleich vorgesprochen werden. Deshalb wurden diese Silben einmal von einem ausgebildeten männlichen Sprecher in einem Sprecherraum vorgelesen und dabei aufgenommen. Diese Listen wurden auf DAT-Kassette kopiert und für diese Versuche zur Verfügung gestellt [20]. Es handelte sich um insgesamt 12 verschiedene Listen mit jeweils 50 Logatomen. Dabei waren die einzelnen Logatome, wie im Theorieteil gefordert, in kurze Sätze eingebunden. Diese Listen sind als Anhang an die Arbeit beigefügt. Es musste beachtet werden, dass jede Logatomliste einer Testperson nur ein einziges mal vorgespielt werden konnte. Selbst wenn zwischen den Versuchen mit der gleichen Liste eine größere Zeitspanne liegt, kann sich der Proband – zumindest unbewusst – an einzelne Silben erinnern und das Ergebnis wird verfälscht [20].

Die Logatomlisten wurden mit einem A-bewerteten Schalldruckpegel von 63dB am Abhörplatz abgespielt. Dadurch waren alle Silben akustisch gut hörbar und man befand sich laut Abbildung 50 immer noch im Bereich der höchsten Verständlichkeit.

### 5.3.2. Störsignale

Die Störsignale wurden alle mit einem Schalldruckpegel von 61dB(A) an der Hörposition abgespielt. Damit waren sie etwas leiser als die Nutzsignale, die dadurch noch gut hörbar waren. Trotzdem war eine deutliche Beeinflussung der Verständlichkeit zu erwarten (Abb. 50).

Die Signalform war dabei unterschiedlich bandbegrenztes Rosa Rauschen, das digital generiert wurde. Zunächst sollte der Einfluss von tieffrequentem Rauschen untersucht werden, deshalb wurde als erstes Signal bandbegrenztes Rosa Rauschen mit einer Grenzfrequenz von  $f_0 = 250\text{Hz}$  gewählt. Dem gegenüber stand Rosa Rauschen von 20Hz bis 20kHz.

Die Ergebnisse aus diesen beiden Versuchen wurden mit den Werten verglichen, die ohne Störsignal erreichbar waren.

## 6. Mess-Software Eartect2003

Für die Messungen in wissenschaftlichen Versuchen sind vor allem eine eindeutige Auswertung und damit die Reproduzierbarkeit der Ergebnisse von entscheidender Wichtigkeit. Deshalb wurde bei den geplanten Hörschwellenmessungen ein dafür entwickeltes Software-Programm verwendet, das über Tastendruck eines Joypads gesteuert wurde. Dieses wurde speziell für diese Anwendung entwickelt und besitzt den Namen Eartect2003, der sich aus den Wörtern "ear" und "detect" zusammensetzt.

### 6.1. Eigenschaften des vorhandenen Programms

Eine erste Version dieser Software wurde bereits 2000 im Rahmen einer Diplomarbeit angefertigt, um Ruhehörschwellen in einer reflexionsarmen Umgebung zu bestimmen [44]. Dieses Programm besaß folgende Eigenschaften:

#### 6.1.1. Allgemeine Daten

Als Programmiersprache wurde C++ Builder gewählt. Damit ist das Programm auf Standard-PCs lauffähig. Die Steuerung des Programms konnte dabei sowohl am Computer selbst auf einer graphischen Bedienoberfläche mittels Maus, als auch über ein Joypad erfolgen.

Es registrierte dabei auch die Reaktion der Probanden auf die Prüfsignale. Dafür waren drei Tasten mit Funktionen belegt: *Programm starten*, *Prüfsignal gehört* und *Programm abbrechen*. Nach [27] darf die Antworttaste keine störenden Geräusche erzeugen. Dies wurde gewährleistet, da die Druckknöpfe sehr leicht ansprechen und ein eventuelles Knacken immer außerhalb der Zeitschranken (siehe 5.1.2. *Zeitlicher Ablauf des Messverfahrens*) lag.

#### 6.1.2. Zeitlicher Ablauf des Messverfahrens

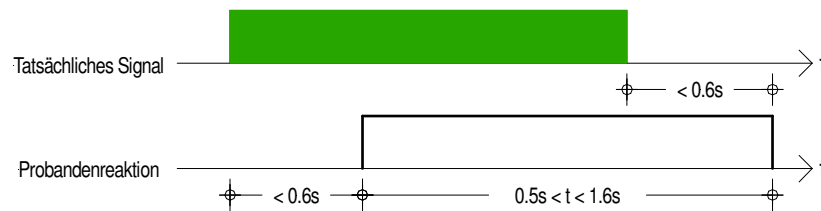
Bei der Messung der Hörschwelle konnte gewählt werden, ob die gesamte Kurve im definierten Frequenzbereich aufgenommen werden soll oder nur eine einzelne Frequenz. Wurde die gesamte Kurve gewählt, führte das Programm eigenständige Messungen in allen definierten Frequenzen durch. Dabei wurde folgender Ablauf eingehalten:

Der Proband startet, sobald er zur Messung bereit ist, das Programm. Anschließend muss er sich nur noch auf das Hören der Testsignale konzentrieren. Dabei muss er



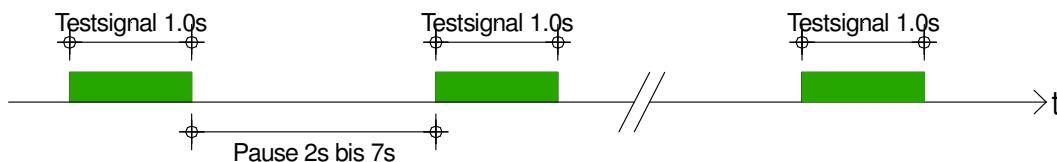
- die Taste drücken, sobald er das Testsignal hört,
- die Taste gedrückt halten, solange er das Testsignal hört und
- die Taste sofort wieder loslassen, sobald das Signal nicht mehr zu hören ist.

Nur wenn alle drei dieser zeitlichen Kriterien erfüllt waren (Abb. 89), wurde das Signal vom Computer als "gehört" gewertet. Wurde der "Signal-gehört"-Knopf gedrückt, obwohl kein Signal anliegt, hatte dies keine Folgen.



**Abbildung 89: Notwendige Probandenreaktion um das Signal als "gehört" zu werten**

Die Pausen zwischen den einzelnen Testsignalen waren zufallsgesteuert (Abb. 90). Dadurch war es nicht möglich, durch systematisches Drücken das Ergebnis zu beeinflussen. Es konnte nur dann gedrückt werden, wenn auch tatsächlich etwas gehört wurde, das System war nicht überlistbar.



**Abbildung 90: Pausen zwischen aufeinanderfolgenden Signalen**

Damit sich die Testperson auf das Signal einstellen konnte, wurde die Hörschwellenbestimmung bei jeder Frequenz mit einem noch gut hörbaren Pegel (25dB) begonnen. Dieser wurde bis zu dreimal vorgespielt. Alle anderen Pegel wurden danach nur einmal getestet.

Reagiert der Proband auf das Signal und drückte die Taste, so wurde der Signalpegel verringert, ansonsten erhöht (siehe Abb. 91). Dadurch musste die Testperson zwar öfter innerhalb der Bestimmung einer Frequenz das Joypad drücken, aber das Ergebnis wurde viel genauer gegenüber der Methode "Drücken, sobald das Signal erstmals gehört wird".

Der Algorithmus war so ausgelegt, dass er die Hörschwelle möglichst genau bestimmt, dabei aber den Proband nicht unnötig belastet. Fehler durch Ermüdung sollten ausgeschlossen werden.

Waren alle Frequenzen durchlaufen, wurden die Ergebniswerte am Computerbildschirm angezeigt. Wenn sich dabei herausstellte, dass zwischenzeitliche Probleme aufgetreten sind, konnten nochmals einzelne Frequenzen nachgemessen werden. Ansonsten konnten die Messergebnisse der Hörschwelle als Textdatei abgespeichert und anschließend ausgewertet werden.

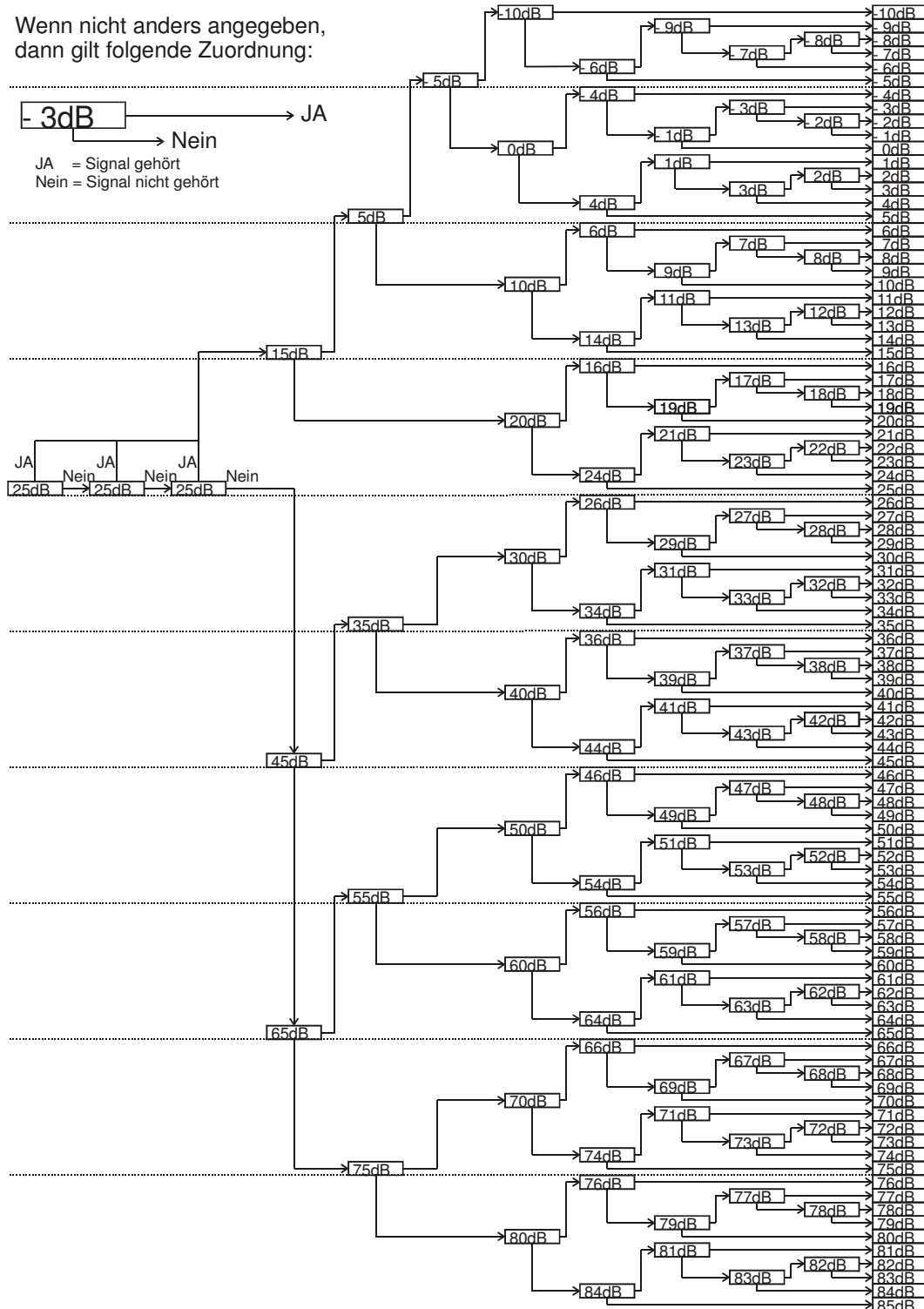


Abbildung 91: Algorithmus der Hörschwellenbestimmung

## 6.2. Notwendige Änderungen an der Software

Für die jetzt durchzuführenden Messungen wurde das Programm an die neuen Bedingungen und Anforderungen angepasst.

Die wichtigste Änderung dabei war, dass für die Hörschwellenbestimmung nicht mehr Terzrauschen als **Testsignal** verwendet wurde, sondern **Sinustöne**. Dadurch war vorgegeben, dass diese nur noch mit einem Lautsprecher wiedergegeben werden (nicht mehr durch vier), so dass die Messumgebung die Anforderungen eines freien Schallfeldes erfüllte [3].

Der **Frequenzbereich**, in dem die Hörschwelle bestimmt werden sollte, wurde auf **40Hz bis 4kHz** festgelegt. So waren auftretende Effekte im tieffrequenten Bereich in jedem Fall erkennbar. Nach oben wurde die Messreihe bei 4kHz abgebrochen, da der frequenzhöchste Maskierer mit 510Hz bandbegrenzt war. Damit schien der gewählte Bereich ausreichend groß. Zusätzlich konnten mit dieser Begrenzung auftretende Probleme bei der Einhaltung der in der Norm [3] geforderten Eigenschaften des Schallfeldes (Homogenität) umgangen werden. Dieser festgelegte Frequenzbereich wurde in **Terzschritten** durchlaufen. Dadurch erhielt man 21 Werte pro Messreihe. Dies bestätigte die Richtigkeit der Begrenzung des Frequenzbereiches nach oben hin, um bei der Messung Ermüdungserscheinungen der Testpersonen zu verhindern. Bereits jetzt wurde eine Messdauer von ca. 25min pro Hörschwelle erreicht.

Die **Prüfsignale** wurden dabei **zunächst von 1kHz aufwärts bis 4kHz** abgespielt, **dann von 1kHz abwärts bis 40Hz**. Bei 1kHz wurde die Prüfung wiederholt und die Ergebnisse gemittelt [3].

Wie in Punkt 6.1.2. *Zeitlicher Ablauf des Messverfahrens* beschrieben, betrug der erste gemessene Pegel jeder Frequenz 25dB. Dieser liegt jedoch im tieffrequenten Bereich nahe an der Ruhehörschwelle [35] bzw. knapp darunter. Deshalb wurde der **Gewöhnungspegel** bei **50Hz, 63Hz und 80Hz auf 45dB** erhöht, bei **40Hz auf 65dB**. Diese Werte wurden so gewählt, dass der bereits vorhandene Messalgorithmus beibehalten werden konnte. Die Schwierigkeit lag dabei darin, einen genügend hohen Einstiegspegel zu erreichen, aber gleichzeitig die notwendigen leisen Pegel messen zu können, ohne dass das Programm vorher abbricht (siehe Abb. 91). Mit den jetzt gewählten Werten war dafür eine gute Möglichkeit gefunden worden.

Zusätzlich wurde geändert, dass das **Testsignal ein zweites mal gleichlaut abgespielt** wird, wenn der Proband dieses beim ersten mal nicht hört. Erst wenn auch das mal zweite keine Reaktion erfolgt, wird der Signalpegel erhöht. Dadurch ergaben sich zwei Vorteile: Zum ei-

nen wurde die Messgenauigkeit nochmals erhöht. Vor allem bei ungeübteren Testpersonen kann dies beachtliche Unterschiede bewirken. Zum anderen war gleichzeitig die Definition der Hörschwelle [3] exakt erfüllt: Hörschwelle = 50% richtige Antworten bei einer vorgegebenen Anzahl von Versuchen.

Um sowohl die Ruhehörschwelle als auch die Mithörschwelle bei sehr hohen Störpegeln mit demselben System messen zu können, wurde ein großer Dynamikbereich benötigt. Dieser wurde durch zwei verschiedene Dynamikeinstellungen erreicht: zunächst wurden die Testsignale bereits unterschiedlich laut abgespeichert. Anschließend wurden sie mit **zwei unterschiedlichen hardwareseitigen Verstärkereinstellungen** abgespielt, so dass sich die einzelnen Bereiche ergänzten.

### 6.3. Eigenschaften des adaptierten Programms

Mit diesen Änderungen war das Programm *Eartect2003* sehr gut für die vorgesehenen Aufgabenstellungen geeignet.

Die äußere Form des Programms war leicht zu bedienen und der Fortschritt der Messung auch außerhalb des Messraumes jederzeit mitzuverfolgen (siehe Abb. 92 und 93). Dadurch konnte die Richtigkeit der Ergebnisse gewährleistet und im Bedarfsfall sofort eine zusätzliche Kontrollmessung einer einzelnen Frequenz durchgeführt werden.

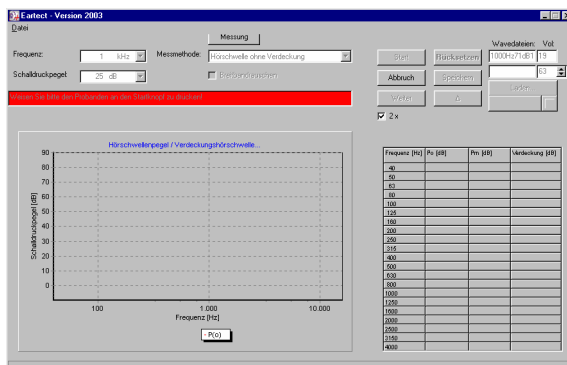


Abbildung 92: Initialisierung des Messprogramms *Eartect2003*

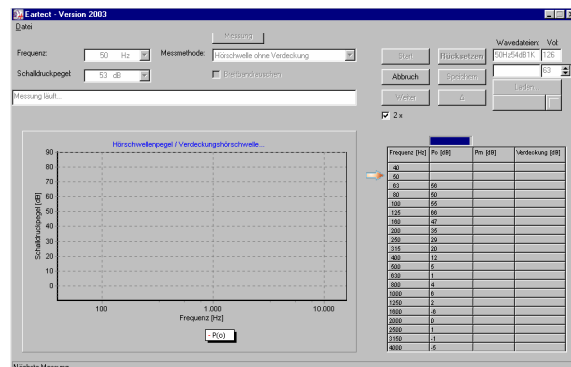


Abbildung 93: Einfache Überwachung des Messablaufs und Anzeige der Ergebnisse

Allerdings musste auf die Auswahl der Testpersonen geachtet werden. Diese durften zum einen keine Anzeichen einer Ohrenerkrankung erkennen lassen [3], zum anderen mussten sie an das Prüfverfahren gewöhnt werden.

Trotz des einfachen Messablaufs (nach dem Starten des Algorithmus muss nur noch eine Taste gedrückt werden, diese aber nach bestimmten Vorgaben) lieferten im Durchschnitt die ersten beiden Hörkurven lediglich Näherungswerte. Die Testpersonen drückten entweder zu spät, liesen die Taste zu früh los oder reagierten z.T. überhaupt nicht, obwohl das Signal gehört wurde. Deshalb mussten, bevor die eigentlichen Versuche beginnen konnten, Übungsmessungen durchgeführt werden. Dabei bot sich an, einen Durchlauf zur reinen Gewöhnung zu verwenden, den zweiten Test jedoch bereits zur Bestimmung der Ruhehörschwelle zu nutzen. Dadurch konnte sofort festgestellt werden, ob die Person für die weitere Untersuchung in Frage kommt oder ob Hörschäden vorlagen. Für diese Einteilung reichten Näherungswerte. War kein abnormales Hörverhalten sichtbar, konnte anschließend mit den eigentlichen Messungen begonnen und exakte Messkurven ermittelt werden.

## 7. Versuchsdurchführung Hörschwellenmessung

### 7.1. Vorbereitungen für Hörschwellenversuche

Bevor die Versuche der Bestimmung der unterschiedlichen Mithörschwellen beginnen konnten, musste die Messumgebung kontrolliert werden.

Die Audiokabine wurde, wie in Punkt 5.2.2. *Messumgebung Audiokabine* beschrieben, aufgebaut. Allerdings wurde für die Kontrollmessungen der Probandenstuhl entfernt [3] und an dessen Stelle in Kopfhöhe (Höhe 120cm) ein Messmikrofon gebracht.



Abbildung 94: Messmikrofon in Kabine

#### 7.1.1. Messumgebung Audiokabine

Zunächst musste die Frage geklärt werden, ob in der Audiokabine tatsächlich Hörschwellenmessungen durchgeführt werden können, ohne dass Umgebungslärm hörbar wird. Dies musste tagsüber geprüft werden, da hier der größte Lärmpegel vorhanden war. Schwierigkeiten wurden hier nicht erwartet, da bereits im Jahr 2000 mit dieser Audiokabine Hörtests durchgeführt wurden, und alle Anforderungen an dieses Kriterium erfüllt wurden. Umso mehr überraschte folgendes Messergebnis (Abb. 95):

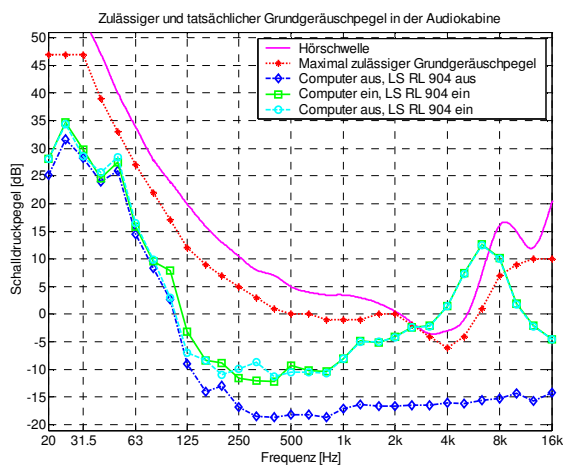


Abbildung 95: Grundgeräuschpegel mit LS RL904

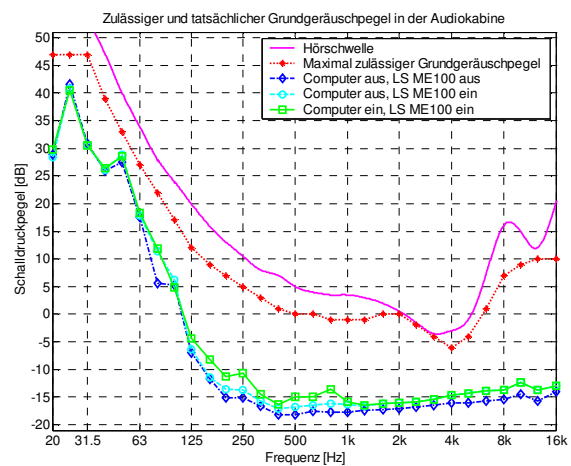
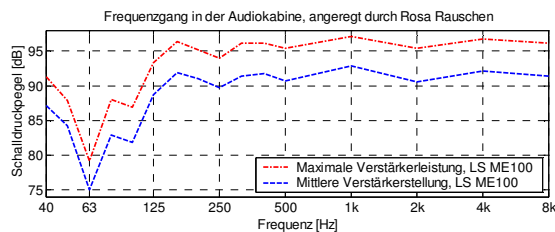


Abbildung 96: Grundgeräuschpegel mit LS ME100

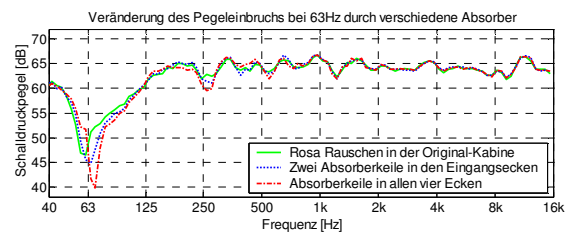
Der Grundgeräuschpegel lag sowohl über den durch die Norm [3] vorgegebenen zulässigen Werten, als auch über der genormten Ruhehörschwelle. Diese Überschreitung betrug z.T. bis zu 12dB.

Wie in der Graphik zu sehen, war dieser starke Anstieg sowohl bei ein- als auch bei ausgeschaltetem Computersystem vorhanden. Deshalb konnte nur der verwendete aktive Lautsprecher *RL904* (siehe Abb. 85) dafür verantwortlich sein. Eine entsprechende Nachfrage beim Hersteller ergab, dass es in diesem Frequenzbereich zu Geräuscentwicklungen durch die Elektronik kommt und damit dieses Messergebnis durchaus realistisch ist. Um den Lautsprecher jedoch als Geräuschquelle ausschließen zu können, wurde ein anderer, passiver Lautsprecher, der *ME100* der Firma *musikelectronic geithain* (siehe Abb. 67 und Punkt 5.2.3. *Verstärker und Lautsprecher*) anstelle der aktiven Box verwendet. Damit ergaben sich die tadellosen Messergebnisse aus Abbildung 96.

Von diesem neuen Lautsprecher wurde nun der Frequenzgang gemessen:



**Abbildung 97: Frequenzgang, gemessen mit Rosa Rauschen**



**Abbildung 98: Veränderung des Einbruchs durch Absorberkeile in den Ecken**

Dabei ergab sich bei der Frequenz 63Hz ein Pegel einbruch von über 15dB, obwohl der Lautsprecher einen Frequenzgang von  $\pm 3\text{dB}$  im Bereich von 40Hz bis 20kHz aufweisen sollte! Dieser Einbruch konnte nur auf modale Einflüsse zurückgeführt werden. Um dies nachzuweisen wurden folgende Berechnungen durchgeführt:

Die Audiokabine besitzt Innenabmessungen von  $l_x = 3\text{m}$ ,  $l_y = 2,5\text{m}$  und  $l_z = 1,9\text{m}$ .

Für einen quaderförmigen Raum mit schallharten Begrenzungsflächen und den Abmessungen  $l_x$ ,  $l_y$  und  $l_z$  (Größenangabe in [m]) ergeben sich die Eigenfrequenzen:

$$f_{n_x, n_y, n_z} = \frac{c}{2} \cdot \sqrt{\left(\frac{n_x}{l_x}\right)^2 + \left(\frac{n_y}{l_y}\right)^2 + \left(\frac{n_z}{l_z}\right)^2} \quad \text{Gleichung 37}$$

mit  $n_x$ ,  $n_y$  und  $n_z = 0, 1, 2, 3, \dots, n$ .

Damit ergibt sich die niedrigste Eigenfrequenz des Raumes allgemein und im Fall der Audiokabine zu

$$f_{1,1,1} = \frac{c}{2} \cdot \sqrt{\left(\frac{1}{l_x}\right)^2 + \left(\frac{0}{l_y}\right)^2 + \left(\frac{0}{l_z}\right)^2} \text{ Hz} = \frac{343}{2} \cdot \sqrt{\left(\frac{1}{3}\right)^2} \text{ Hz} = 57 \text{ Hz} \quad \text{Gleichung 38}$$

Betrachtet man den Fall, dass nur zwei gegenüberliegende, schallharte Wände reflektieren, so treten folgende Eigenfrequenzen der Axialmoden auf:

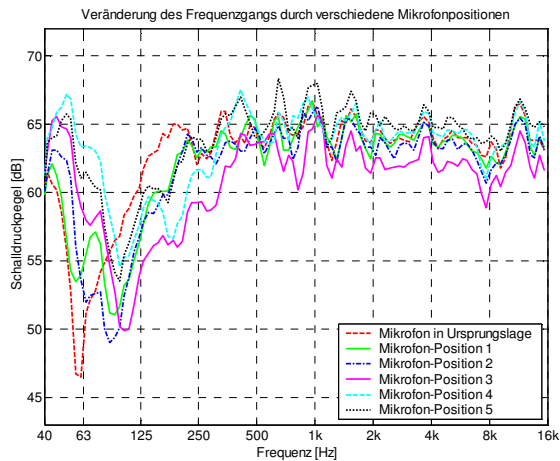
$$\begin{aligned} f_{n_x} &= \frac{c}{2} \cdot \frac{n_x}{l_x} \text{ Hz} = \frac{343}{2} \cdot \frac{1}{3,0} \text{ Hz} = 57 \text{ Hz} && \text{in } x\text{-Richtung} \\ f_{n_y} &= \frac{c}{2} \cdot \frac{n_y}{l_y} \text{ Hz} = \frac{343}{2} \cdot \frac{1}{2,5} \text{ Hz} = 68 \text{ Hz} && \text{in } y\text{-Richtung} \\ f_{n_z} &= \frac{c}{2} \cdot \frac{n_z}{l_z} \text{ Hz} = \frac{343}{2} \cdot \frac{1}{1,9} \text{ Hz} = 90 \text{ Hz} && \text{in } z\text{-Richtung} \end{aligned} \quad \text{Gleichung 39}$$

Vergleicht man diese Ergebnisfrequenzen mit dem Frequenzgang in Abbildung 97, so kann man erkennen, dass die Resultate gut mit den gemessenen Werten übereinstimmen. Die Raummoden in X- und Y-Richtung wirken sich offensichtlich besonders störend aus, aber auch die Mode in Z-Richtung ist nicht stark genug bedämpft und deutlich ausgeprägt. Die Absorption der eingebauten BKA-Module [26] reichte aber offenbar nicht aus, um in diesem winzigen Raum Freifeldbedingungen zu erreichen. Deshalb wurden zusätzliche Absorberkeile in den Raumecken angebracht, die den modalen Schallfeldeinfluss verringern sollten. Doch auch diese brachten keine Verbesserung der Situation (Abb. 98).

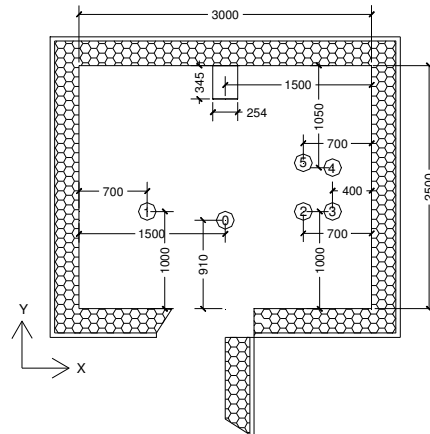
Deshalb wurde versucht, diesen Einfluss mit einer veränderten Abhörposition zu umgehen. Dafür wurde der Frequenzgang an insgesamt sechs verschiedenen Standpunkten gemessen.

Wie jedoch in Abbildung 99 zu sehen ist, konnte der Pegleinbruch nie in ausreichendem Maße verhindert werden: die Einbruchfrequenz verschob sich je nach Abhörposition mehr oder weniger nach oben, verringerte sich dabei aber nur unwesentlich.



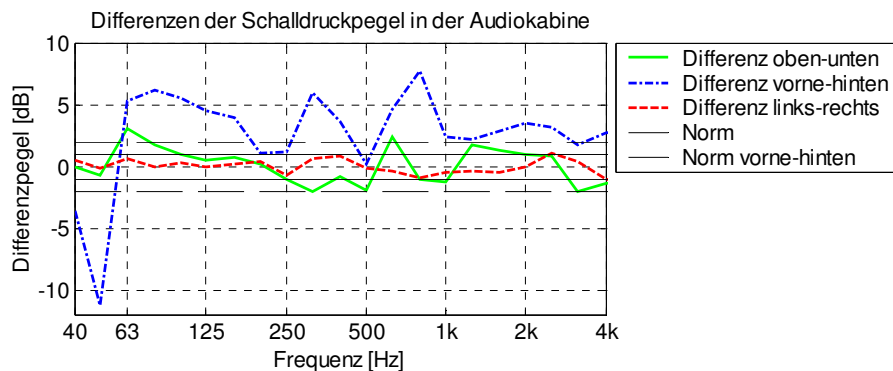


**Abbildung 99: Veränderung des Einbruchs durch unterschiedliche Messpositionen**



**Abbildung 100: Lagepunkte der verschiedenen Messpositionen**

Wie wirkte sich diese Messumgebung auf die Homogenität des Schallfeldes aus? Diese ist das einzige Kriterium, das laut Norm [3] bei Versuchen mit reinen Sinustönen erfüllt sein muss, dafür sind diese Anforderungen (siehe 5.2.1 Anforderungen an die Messumgebung) aber auch sehr streng. Die Homogenität, die an der ursprünglichen Abhörposition gemessen wurde, konnte die Forderungen bei weitem nicht erfüllen:



**Abbildung 101: Homogenität des Schallfeldes in der Audiokabine**

Benötigt wurden  $\pm 1\text{dB}$  bei einem Abstand von 30cm der Messpunkte zueinander. Bei der Differenz vorne-hinten ergaben sich unter Anwendung des inversen Abstandsgesetz  $\pm 2,1\text{dB}$  als Forderung:

$$L = L_0 + 20 \cdot \lg\left(\frac{r}{r_0}\right) \text{dB} = L_0 + 20 \cdot \lg\left(\frac{1,10}{1,25}\right) \text{dB} = L_0 + 1,1\text{dB} \quad \text{Gleichung 40}$$

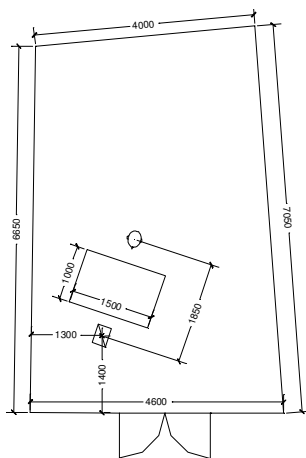
In der Audiokabine hingegen traten Maximalwerte von +8dB/-11dB auf. Auch mit einer veränderten Abhörposition würden sich die Werte nicht wesentlich verbessern, da sich der Fre-

quenzgang an den neuen Positionen nicht stark ändert (siehe Abb. 99). Daher sind diese Resultate inakzeptabel. Alle Ergebnisse, die mit diesem Messaufbau (nur ein Lautsprecher in der reflexionsarmen Umgebung) in dieser Messumgebung erzielt werden, wären automatisch anfechtbar. Deshalb musste eine besser geeignete Messumgebung gefunden werden. Dafür bot sich der neu errichtete Halbfreifeldraum an.

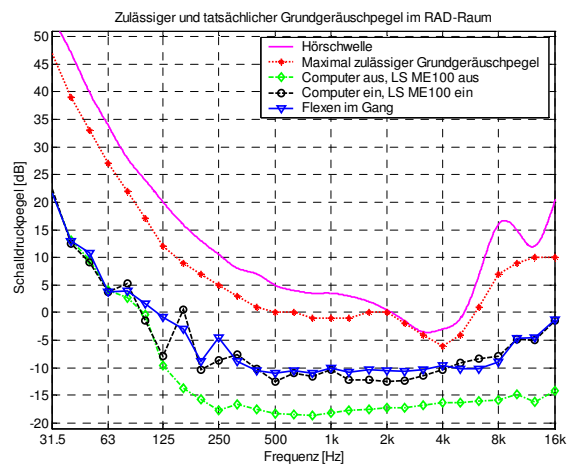
### 7.1.2. Testmessungen im Halbfreifeldraum

Der Halbfreifeldraum, ursprünglich als Hallraum ausgeführt, wurde 1998 in die jetzige Form umgebaut. Er besitzt schiefwinklig zueinander stehende Wände. Mit einer Grundfläche von knapp  $40\text{m}^2$  und einer mittleren Raumhöhe von ca.  $6\text{m}$  ergibt sich ein freies Raumvolumen von  $236\text{m}^3$  (Abb. 102).

Anstelle der konventionellen Keilabsorber erhielt dieser Raum an den Seiten- und Deckenwänden eine neuartige Absorberauskleidung aus Verbund-Platten-Resonatoren (VPR) (siehe auch Abb. 86), die eine Bautiefe von nur  $25\text{cm}$  erreichen und Messungen bis zu einer Grenzfrequenz von  $50\text{Hz}$  erlauben [36]. Der Fußboden ist schallhart ausgeführt.



**Abbildung 102: Grundriss Halbfreifeldraum**

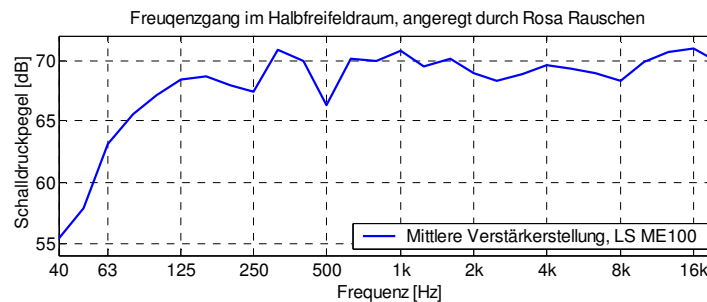


**Abbildung 103: Grundgeräuschpegel im Halbfreifeldraum**

Dieser Halbfreifeldraum besitzt jedoch den Nachteil, dass er vom restlichen Gebäude nicht körperschallentkoppelt ist. Die Umgebungslautstärke ist jedoch oft recht hoch, so dass geprüft werden musste, wie stark sich dieser Lärm in den Messraum überträgt. Wie jedoch die Messungen ergaben, war die Angst eines zu hohen Grundgeräuschpegels unbegründet. Selbst bei extrem lauten Störgeräuschen außerhalb des Raumes konnten die Anforderungen leicht eingehalten werden (Abb. 103).

Um bei den Messungen in diesem Halbfreifeldraum Reflexionen vom schallharten Boden zu vermeiden, wurde beschlossen zwei Caruso-Absorptionsplatten mit der jeweiligen Dicke von 10cm (siehe Abb. 88) übereinander zu legen und diese zwischen Lautsprecher und Abhörposition zu platzieren. Der Lautsprecher selbst wurde wieder in einer Höhe von 120cm aufgebaut (Ohrhöhe einer sitzenden Versuchsperson), der Abstand zur Abhörposition betrug 185cm (siehe Abb. 102). Dieser Zwischenraum entspricht der nächstmöglichen Sitzposition, in der die Beine der Probanden noch bequem vor den Absorberplatten stehen konnten.

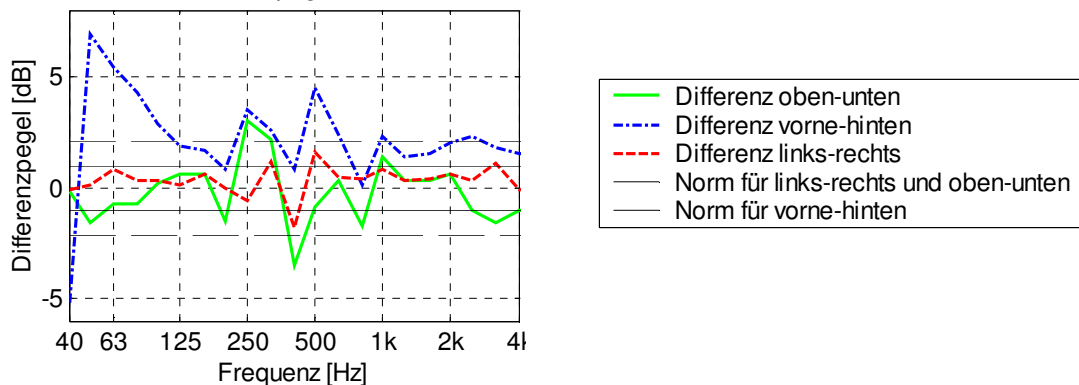
Mit dieser Aufstellung ergab sich an der Abhörposition folgender Frequenzgang des Lautsprechers ME100:



**Abbildung 104: Frequenzgang im Halbfreifeldraum**

Wie gut zu sehen ist, weist der Frequenzgang im tieffrequenten Bereich – ähnlich wie in der Audiokabine – einen starken Pegel einbruch von ca. 15dB auf. Bei der anschließenden Prüfung der Homogenität des Schallfeldes war ebenfalls festzustellen, dass die Anforderungen mit dieser Messanordnung ebenfalls nicht eingehalten werden konnten (Abb. 105).

Differenzen der Schalldruckpegel im Halbfreifeldraum



**Abbildung 105: Homogenität des Schallfeldes im Halbfreifeldraum**

Auch eine veränderte Messposition im Raum verbesserte das Ergebnis nicht ausreichend. Deshalb wurde beschlossen, die Versuche im großen Freifeldraum durchzuführen.

## 7.1.3. Messumgebung Freifeldraum

Der große Freifeldraum des IBP Stuttgart bietet für die geplanten Versuche ideale Bedingungen. Er besitzt Abmessungen von 9,3m x 10,4m x 10,3m, wodurch sich ein Raumvolumen von 996m<sup>3</sup> ergibt [37].

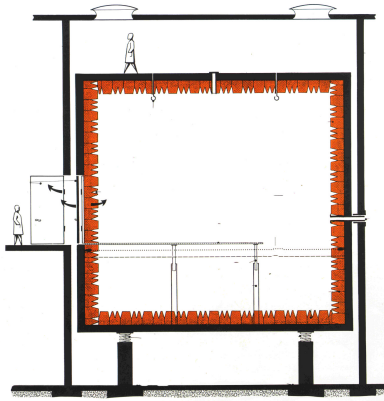


Abbildung 106:  
Querschnitt Freifeldraum [37]

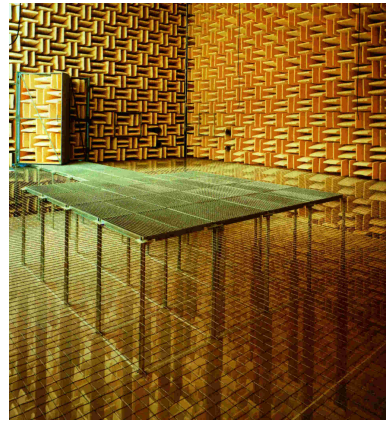


Abbildung 107:  
Großer Freifeldraum des IBP [37]

Die Grenzfrequenz dieses Raumes liegt bei 80Hz. Aufgrund der Größe und des Abstandes zu den Begrenzungsflächen können jedoch problemlos Messungen bis 40Hz durchgeführt werden. Der Lautsprecher wurde dafür auf einem Ständer auf 120cm Höhe aufgebaut, der Abstand zur Abhörposition betrug 145cm. Um ein versehentliches Verrutschen des Lautsprecherständers oder des Probandenstuhls zu verhindern, wurden beide mit Kabelbindern an den Gitterplatten fixiert. Alle anderen nicht benötigten Gitterplatten wurden entfernt, um Reflexionen daran zu vermeiden (siehe Abb. 108).



Abbildung 108: Messaufbau im Freifeldraum

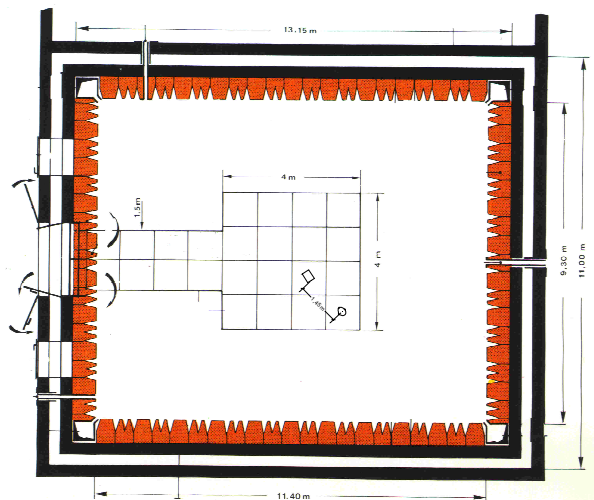


Abbildung 109: Positionierung des Messaufbaus

Die verwendeten Kabel wurden durch eine Einlassöffnung in den Vorraum des Freifeldraums geführt und dort mit den entsprechenden Geräten verbunden. Die Öffnung selbst wurde mit Terostat, einem dauerelastischen Dichtkit, verschlossen und damit ein Eindringen von Störschall verhindert (siehe Abb. 69). Der exakte Messaufbau im Vorraum wurde bereits in Punkt 5.2. *Aufbau für die Hörschwellenmessung* beschrieben, deshalb kann hier darauf verzichtet werden.

Der Freifeldraum ist vom restlichen Gebäude körperschallentkoppelt. Deshalb wurde von einem sehr geringen Grundgeräuschpegel ausgegangen, was auch entsprechende Messungen bestätigten:

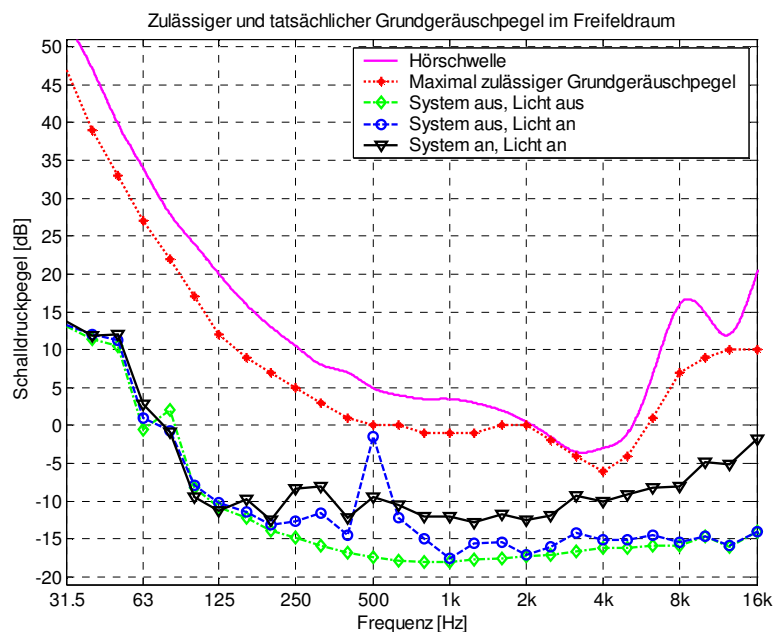


Abbildung 110: Grundgeräuschpegel im Freifeldraum

Der Ausreißer der blau gestrichelten Kurve kann nicht genau erklärt werden. Allerdings befinden sich direkt vor dem Freifeldraum Kühlaggregate, die sich von Zeit zu Zeit einschalten. Eventuell kommt dieser Anstieg des Grundgeräuschpegels davon. Allerdings ist auch dann der gemessene Pegel noch innerhalb des erlaubten Bereichs, so dass dies kein Problem darstellte.

Bei der anschließenden Überprüfung der weiteren Anforderungen an die Messumgebung wurde zunächst der Frequenzgang des Lautsprechers *ME100* gemessen: das hochempfindliche 1"-Messmikrofon *Brüel&Kjaer, Type 4179* (siehe 5.2.5. *Messgeräte*) befand sich am Abhörreferenzpunkt, der Probandenstuhl war noch nicht eingebaut. Dabei ergab sich folgende Messkurve:

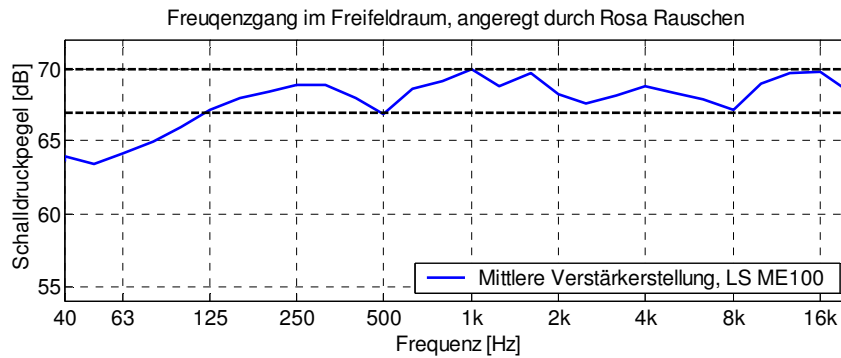


Abbildung 111: Frequenzgang des Lautsprechers ME100 im Freifeldraum

Das Ergebnis lieferte gute Werte, die mit den Angaben aus dem Datenblatt übereinstimmen: im gesamten Übertragungsbereich von 50Hz bis 20kHz beträgt der maximale Schwankungsbereich  $\pm 3\text{dB}$ , im Frequenzbereich von 125Hz bis 20kHz sogar nur  $\pm 1,5\text{dB}$ . Dieses Resultat wirkte sich auch auf die Homogenität des Schallfeldes aus:

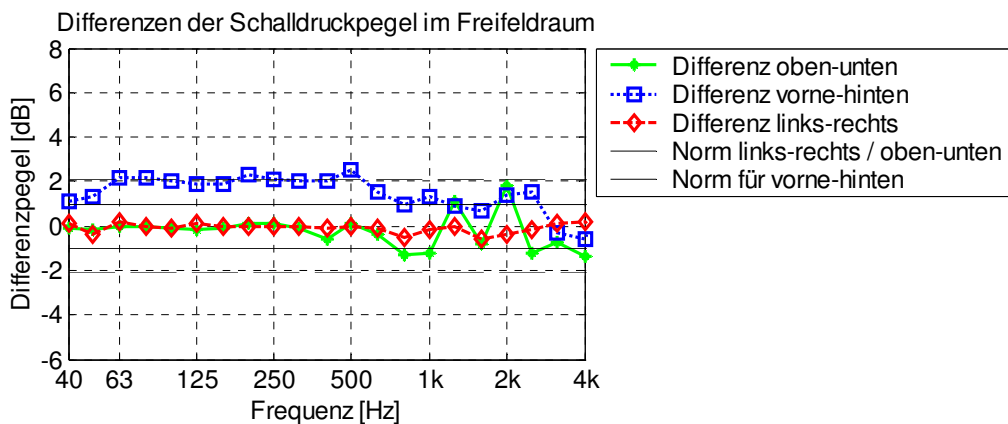


Abbildung 112: Homogenität des Schallfeldes im Freifeldraum

Die Anforderungen der Norm [3], dass sich die Schallpegel 15cm vom Referenzpunkt entfernt um maximal 1dB ändern dürfen, konnten nahezu im gesamten untersuchten Frequenzbereich eingehalten werden. Es traten lediglich zwei geringe Überschreitungen der grünen durchgezogenen Linie bei 800Hz und 2kHz auf, die sich auf die Differenz oben-unten bezieht. Den Versuchspersonen wird bei den Messungen jedoch erklärt, dass sie während der Messung ruhig sitzen müssen, so dass eine Auf- und Abwärtsbewegung um  $\pm 15\text{cm}$  auszuschließen ist. Deshalb wurden die geringen Überschreitungen vernachlässigt. Die wichtige Kurve links-rechts hingegen blieb immer innerhalb der geforderten Grenzen. Damit wurde als endgültige Messumgebung für die Versuche zur Bestimmung von Mithörschwellen der große Freifeldraum des IBP Stuttgart festgelegt.

## 7.2. Ergebnisse der Messungen mit konventionellen Störsignalen

Zunächst sollten die unterschiedlichen Versuchsergebnisse von Slawin und Zwicker untersucht werden. Da sowohl Slawin als auch Zwicker Mithörschwellen ermittelt haben, bei denen Töne als Störsignal verwendet wurden, wurde in der Versuchsreihe ebenfalls mit diesem Signaltyp begonnen.

### 7.2.1. Störsignal Sinuston

Für diese Messung wurde ein 200Hz-Sinuston auf die Pegel 30dB, 50dB und 70dB am Hörort eingemessen und bei den Versuchen dauerhaft im Hintergrund abgespielt.

Dabei wurde zunächst die Hörschwelle bei beidohrigem Hören bestimmt (Abb. 113). Dieser Versuch wurde an nur zwei normalhörenden Probanden durchgeführt, so dass die Ergebnisse lediglich als Näherungen betrachtet werden können. Jedoch traten Trends und Effekte bereits hier gut sichtbar auf, so dass diese ersten Messungen als ausreichend gewertet wurden.

Wie in der Graphik zu sehen, steigt die Mithörschwelle in der Umgebung des Störtones stark an und erreicht bei der Störtonfrequenz  $f = 200\text{Hz}$  ihr Maximum, das knapp unter dem Pegel des Störsignals liegt. Zum höherfrequenten Bereich hin sinkt die Mithörschwelle jedoch wieder ab und bildet sich auf die Ruhehörschwelle zurück. Die Messergebnisse verhalten sich also in keinsten Weise so, wie es von Slawin veröffentlicht wurde und steigen zum hochfrequenten Bereich hin nicht an.

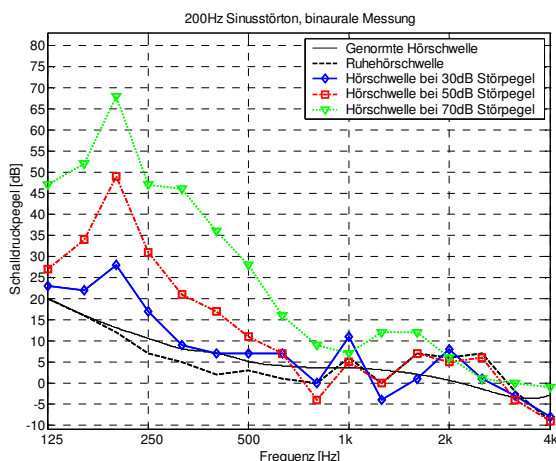


Abbildung 113: Mithörschwellen durch 200Hz-Sinuston, binaural gemessen

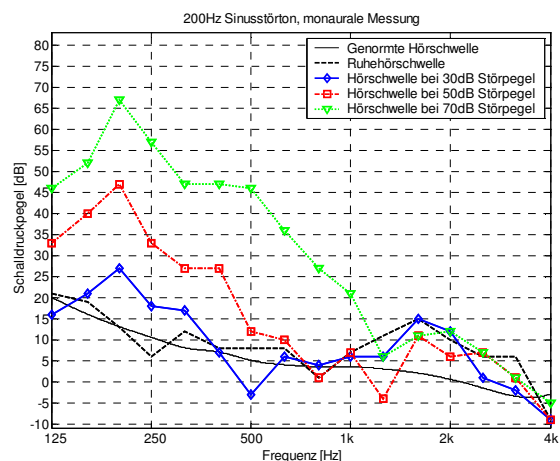


Abbildung 114: Mithörschwellen durch 200Hz-Sinuston, monaural gemessen

Um auszuschließen, dass eventuelle Effekte nur bei monauralem Hören auftreten, wurde die gleiche Messung mit einohrigem Hören durchgeführt. Dafür verwendeten die Versuchsperso-

nen in einem Ohr einen Gehörschutz, der die eintreffende Schalle um 20dB reduzierte. Damit konnte nur noch das nicht verschlossene Ohr die Signale wahrnehmen und bestimmte alleine die Mithörschwelle. Jedoch traten auch bei dieser Messung keine Effekte im kHz-Bereich auf (Abb. 114). Zwar verbreiterte sich die obere Flanke der Mithörschwelle zu höheren Frequenzen hin, das Slawin-Resultat mit einer ansteigenden Mithörschwelle zu hohen Frequenzen konnte aber nicht nachvollzogen werden.

Aus diesem Grund wurde beschlossen weitere Versuche mit unterschiedlichen Störsignalen durchzuführen.

### 7.2.2. Störsignal Terzrauschen

Im nächsten Fall, der untersucht wurde, wurden die Mithörschwellen durch Terzrauschen bestimmt. Diese Messungen wurden wieder binaural durchgeführt, da sich beim vorhergehenden Versuch kein gravierender Unterschied zwischen den monauralen und binauralen Resultaten ergeben hat. Zusätzlich ist die binaurale Abhörsituation für die Probanden angenehmer, da während der Messung kein Ohrstöpsel verwendet werden muss.

Aus Zeitgründen wurde in diesem Fall nur eine Versuchsperson untersucht. Es wurden mit dieser zwei Einzelmessungen durchgeführt und gemittelt, so dass die vorliegenden Kurven lediglich als vorläufiges Ergebnis mit Trenderkenntung zu werten sind.

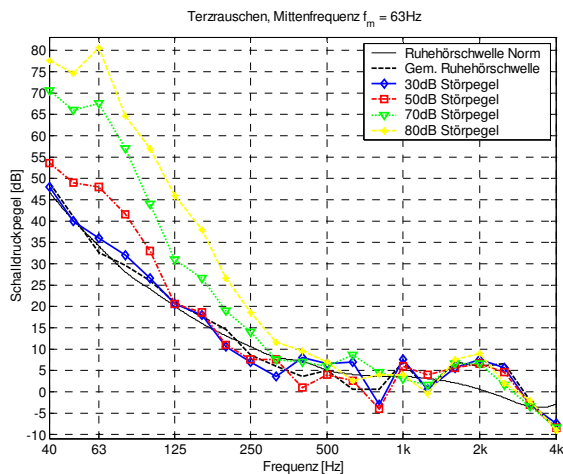


Abbildung 115: Terzrauschen,  $f_m = 63\text{Hz}$

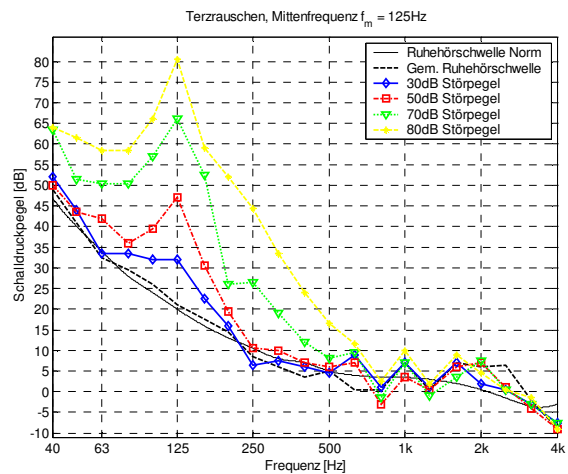
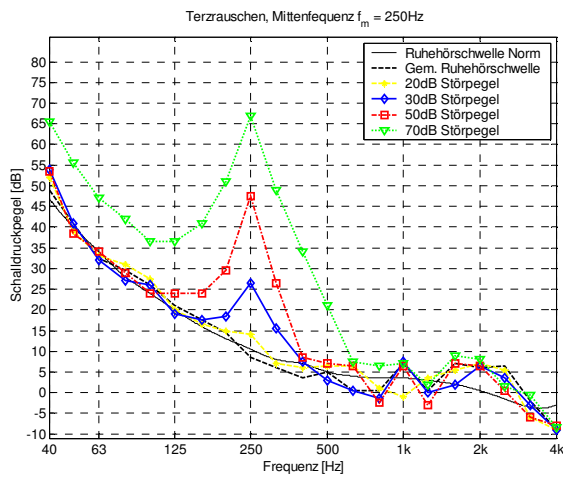
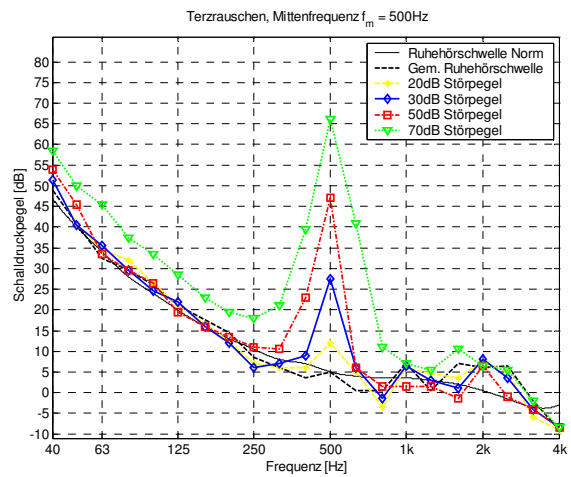


Abbildung 116: Terzrauschen,  $f_m = 125\text{Hz}$



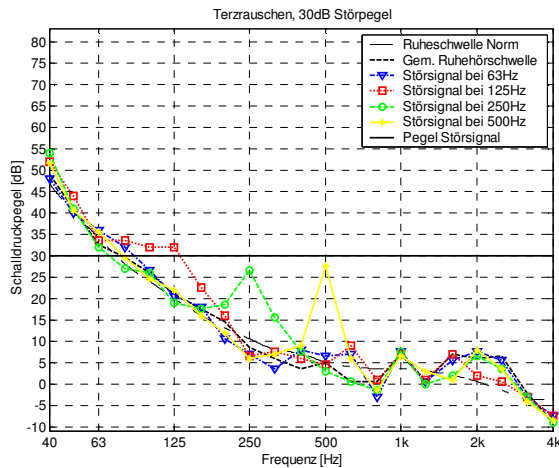
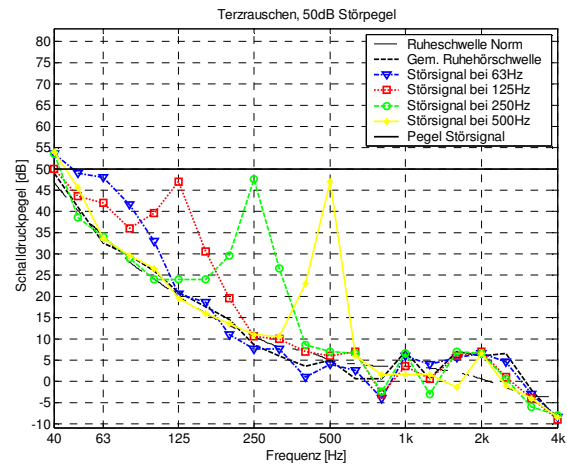
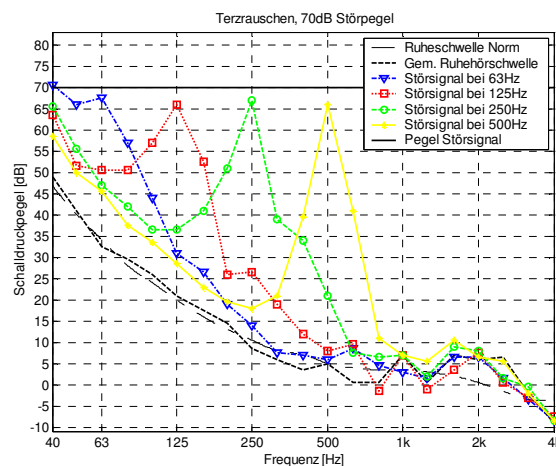
Abbildung 117: Terzrauschen,  $f_m = 250\text{Hz}$ Abbildung 118: Terzrauschen,  $f_m = 500\text{Hz}$ 

In allen untersuchten Varianten zeigen die Mithörschwellen einen steilen Anstieg der unteren Flanke (unten in Bezug zur Frequenz) bis zur Mittenfrequenz des Störsignals. Die etwas flacher abfallende obere Flanke kann vor allem bei den beiden tieferfrequenten Rauschsignalen ( $f_m = 63\text{Hz}$  und  $f_m = 125\text{Hz}$ ) beobachtet werden.

Weiter ist zu sehen, dass bei Darstellung mit einer logarithmisch aufgetragenen Frequenzachse die Mithörschwelle umso schmaler wird, je höherfrequent das Störsignal wird. Wird der Bereich jedoch in Frequenzgruppen ausgedrückt, so erhält man gleiche Ergebnisse für alle Störfrequenzen. Der Bereich, in dem Verdeckungserscheinungen vorhanden sind, reicht bei einer Störung mit  $f_m = 125\text{Hz}$  und einem Schalldruckpegel von  $L = 70\text{dB}$  von  $63\text{Hz}$  bis  $500\text{Hz}$ , hat also Auswirkungen über viereinhalb Frequenzgruppen (siehe 3.5. *Frequenzgruppen und Tonheit z*). Bei einem Störsignal mit der Mittenfrequenz  $f_m = 500\text{Hz}$  (Schalldruckpegel wieder  $L = 70\text{dB}$ ) werden Frequenzen von  $250\text{Hz}$  bis  $800\text{Hz}$  maskiert, also ca. fünfeinhalb Frequenzgruppen. Bei diesen durchschnittlich fünf Frequenzgruppen, in denen Verdeckung auftritt, werden eine bis zwei durch die aufsteigende Flanke übertönt, während etwa drei bis vier Frequenzgruppen von der abfallenden Flanke betroffen sind.

Auffallend bei den ermittelten Hörkurven ist, dass bei allen vier Fällen ab einem Störpegel von  $70\text{dB}$  die Mithörschwelle im tieffrequenten Bereich nicht mehr auf die Ruhehörschwelle zurückgeht, sondern parallel zu dieser nach oben verschoben wird. Bei den niedrigeren Störpegeln bis einschließlich  $50\text{dB}$  ist der Effekt nicht zu beobachten.

Ansatzweise kann auch der in der Literatur beschriebene Rückgang des Maximums der Hörschwelle zu höheren Frequenzen hin beobachtet werden (Abb. 119 bis 121). Dieser Rückgang liegt laut Zwicker [11] im Bereich von 3dB, der im gemessenen Fall bestätigt werden kann.

Abbildung 119: Terzrauschen,  $L = 30\text{dB}$ Abbildung 120: Terzrauschen,  $L = 50\text{dB}$ Abbildung 121: Terzrauschen,  $L = 70\text{dB}$ 

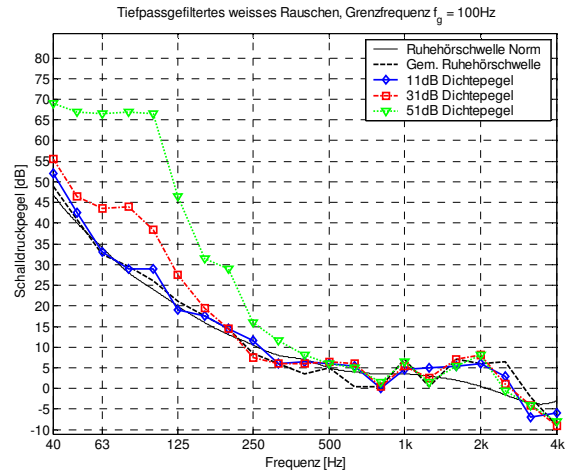
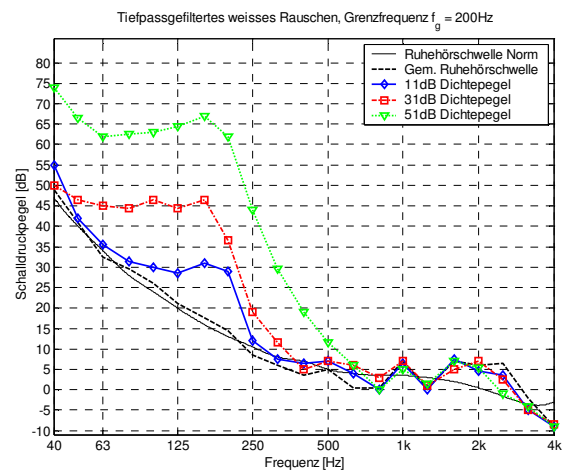
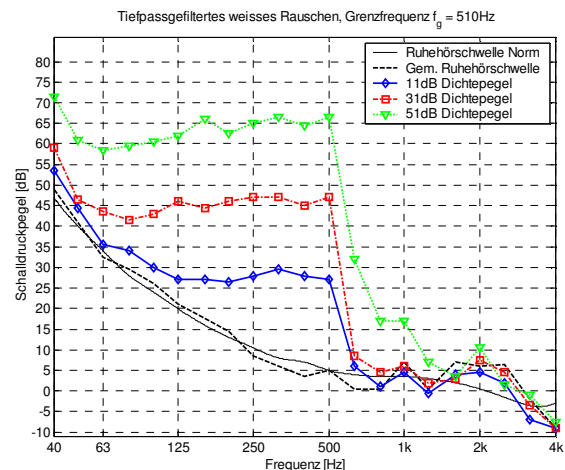
### 7.2.3. Störsignal Tiefpassrauschen

Weiters wurde als Störsignal Tiefpassrauschen mit den Grenzfrequenzen 100Hz, 200Hz und 510Hz (also eine, zwei und fünf Frequenzgruppen) untersucht. Bei der zweimaligen Messung eines Probanden ergaben sich folgende Mithörschwellen:

Bei diesen Versuchen wird der Dichtepiegel der Störsignale, wie in 3.6.1. *Schalldruckpegel und Dichtepiegel* und 5.3.2. *Störsignale* beschrieben, konstant gehalten. Dabei entspricht ein Dichtepiegel von  $l = 11\text{dB}$  beim Tiefpassrauschen mit der Grenzfrequenz  $f_g = 100\text{Hz}$  einem Schalldruckpegel von  $L = 30\text{dB}$ , bei den beiden höheren Grenzfrequenzen  $f_g = 200\text{Hz}$  und  $f_g = 510\text{Hz}$  beträgt der entsprechende Schalldruckpegel  $33,5\text{dB}$  bzw.  $37,9\text{dB}$  (siehe 5.3.2. *Störsignale*). Äquivalent dazu ergeben sich  $31\text{dB}$  Dichtepiegel bei  $50\text{dB}$  Schalldruckpegel und  $51\text{dB}$  Dichte bei  $70\text{dB}$  Schalldruckpegel (Grenzfrequenz  $f_g = 100\text{Hz}$ ).

Gut zu erkennen ist der annähernd horizontale Verlauf aller Ergebniskurven. Die enthaltenen Schwankungsbewegungen können auf die fehlende Mittelung mehrerer Probanden zurückgeführt werden.

An der Stelle der Grenzfrequenz des Tiefpassrauschens fällt die Mithörschwelle stark ab und geht auf die Ruhehörschwelle zurück. Dieser Abfall ist gut mit den fallenden Flanken der Terzrausch-Kurven vergleichbar.

Abbildung 122: Tiefpassrauschen,  $f_g = 100\text{Hz}$ Abbildung 123: Tiefpassrauschen,  $f_g = 200\text{Hz}$ Abbildung 124: Tiefpassrauschen,  $f_g = 510\text{Hz}$

Interessant ist auch, wie sich die unterschiedliche Bandbreite der Störsignale auf die Hörschwelle auswirkt:

Es ist festzustellen, dass die Mithörschwelle umso niedriger wird, je breiter das Störsignal ist. Besonders gut kann dies in Abbildung 127 beobachtet werden: während die Mithörschwelle bei Tiefpassrauschen bis 100Hz (Dichtepegel 51dB) konstant bei 67dB Schalldruckpegel liegt, beträgt sie mit der Grenzfrequenz 200Hz 63dB und bei  $f_g = 510\text{Hz}$  nur noch 60dB.

Diese Verringerung der Mithörschwelle fällt dabei umso stärker aus, je höher der Dichtepegel des Störsignals gewählt wird: Die Mithörschwellen der Grenzfrequenzen 100Hz und 200Hz liegen bei einem Dichtepegel von 31dB bis zur Grenzfrequenz  $f_g = 100\text{Hz}$  noch exakt übereinander (Abb. 126), während ab einem Dichtepegel von 31dB die Mithörschwelle wie gerade beschrieben, umso höher wird, je schmaler die Bandbreite des Störsignals gewählt wurde.

Inwieweit dieser Effekt ein zufälliges Resultat einer "Einzelmessung" ist (es wurden nur zwei Messergebnisse einer einzigen Versuchsperson gemittelt), oder ob er auch noch nach der Mittelung mehrerer Versuchspersonen vorhanden ist, kann an dieser Stelle nicht geklärt werden.

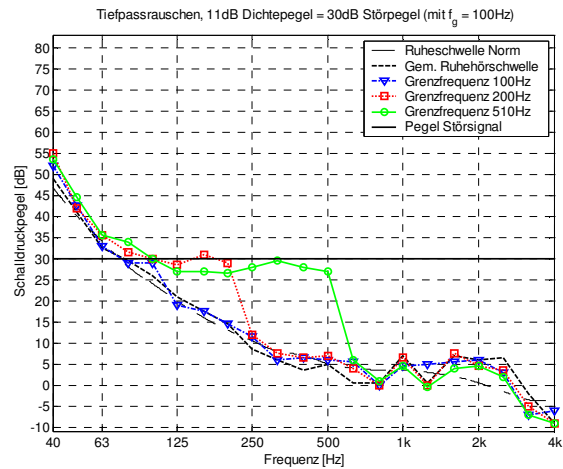


Abbildung 125: Tiefpassrauschen,  $l = 11\text{dB}$

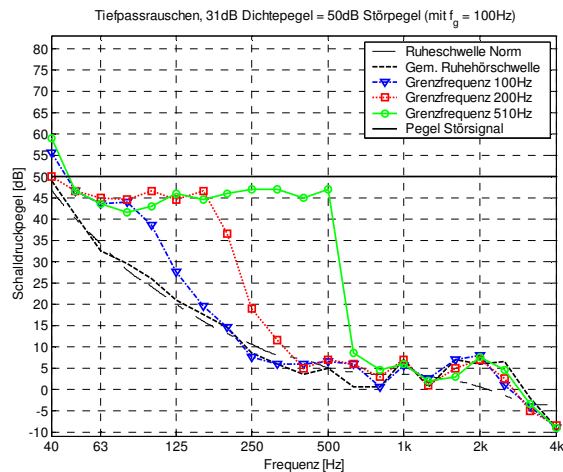


Abbildung 126: Tiefpassrauschen,  $l = 31\text{dB}$

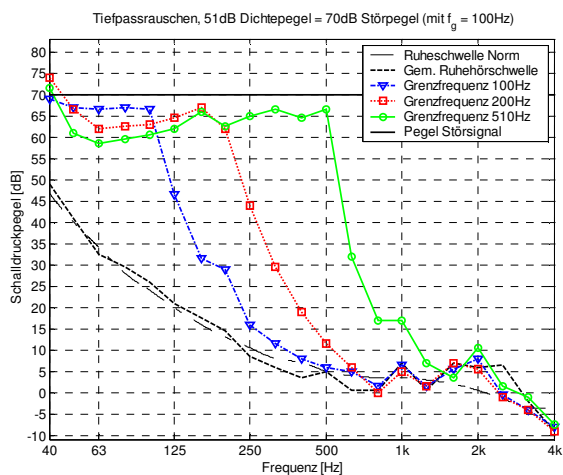


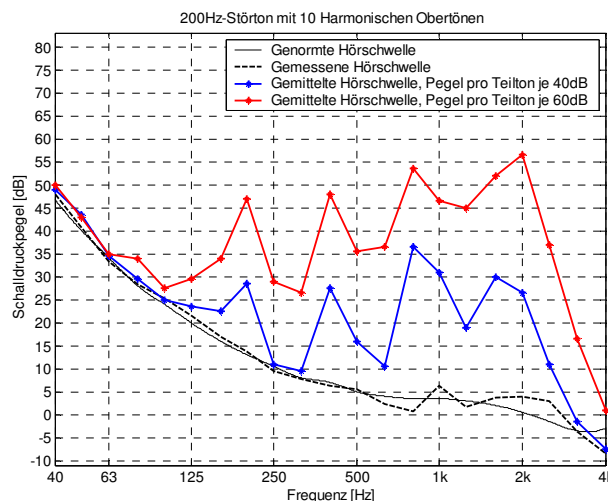
Abbildung 127: Tiefpassrauschen,  $l = 51\text{dB}$

## 7.2.4. Störsignal Tongemisch

Bis zu diesem Zeitpunkt konnten alle theoretischen Darstellungen von Zwicker [11] sehr gut nachvollzogen werden, während für die Aussagen von Slawin (5.1. Ausgangslage) keine Bestätigung gefunden wurde. Da Slawin bereits bei Erscheinen der deutschen Übersetzung seines Buches verstorben war und auch das Datum der ersten Ausgabe sehr lange zurückliegt (Erste Auflage im Jahr 1960), konnten keine Personen mehr gefunden werden, die Angaben zu diesem Problem machen konnten.

Deshalb wurde nun untersucht, ob ein Übersetzungsfehler bzw. ein Missverständnis Ursache der völlig unterschiedlichen Ergebnisse von Slawin und Zwicker sein könnten:

Bis jetzt wurde davon ausgegangen, dass mit der Bezeichnung *Ton* ein reiner Sinuston gemeint ist. Nach den vorliegenden Resultaten ist jedoch zu vermuten, dass Slawin von einem Tongemisch, also von einem musikalischen Ton, spricht, wenn er Hörschwellen mit Tönen als Störsignal präsentiert. Um dies zu bestätigen wurde ein 200Hz-Störton erzeugt, der zehn harmonische Obertöne besitzt. Diese besitzen jeweils die gleiche Amplitude wie der Grundton. Damit sind Spektralanteile der Frequenzen 200Hz, 400Hz, 600Hz, 800Hz, 1kHz, 1,2kHz, 1,4kHz, 1,6kHz, 1,8kHz, 2kHz und 2,2kHz vorhanden. Die Phasenlagen der einzelnen Teiltöne sind zufällig aus dem gesamten Wertebereich gewählt. Damit ergaben sich folgende Mithörschwellen:



**Abbildung 128: 200Hz-Störton mit zehn harmonischen Obertönen gleicher Amplitude**

Diese Messung wurde mit zwei Versuchspersonen jeweils zweimal durchgeführt, so dass insgesamt vier Ergebnisse gemittelt wurden.

Wie deutlich zu erkennen ist, ergeben sich zwei Mithörschwellen, die jeweils aussehen, als hätte man mehrere Hörschwellen von Schmalbandrauschen unterschiedlicher Frequenz übereinandergelegt: im Bereich der vorhandenen Spektralkomponenten steigt die Mithörschwelle stark an, erreicht an der Störfrequenz ihr Maximum und fällt nach oben hin wieder ab, bis erneut ein Anstieg durch die nächste Spektralkomponente folgt. Wären also noch mehr Frequenzen im Störsignal enthalten, wären die Einbuchtungen weniger stark ausgeprägt und würden bei entsprechender Dichte der Einzelfrequenzen sogar ganz verschwinden.

Zusätzlich kann bei dieser Messkurve ein Ansteigen der Verdeckung zu höheren Frequenzen hin beobachtet werden.

#### 7.2.5. Allgemeine Erkenntnisse aus den Hörversuchen mit Zufallsrauschen

Damit ist diese Versuchsanordnung mit einem Tongemisch als Störsignal die einzige, bei der der von Slawin beschriebene Effekt einer zu hohen Frequenzen ansteigenden Mithörschwelle tendenziell nachvollzogen werden konnte. Die Verdeckungseffekte reichen bis in den kHz-Bereich hinein. Dies ist allerdings nur der Fall, da in diesem Frequenzbereich auch Störkomponenten vorhanden sind. Es muss also davon ausgegangen werden, dass sich die Aussagen von Slawin, dass ein Ton ein Ansteigen der Mithörschwelle bis in den kHz-Bereich hinein verursacht, nicht auf ein Sinusstörsignal (Ton im physikalischen Sinn) beziehen, sondern auf ein Tongemisch, einen Ton im musikalischen Sinn.

Mit diesem Hintergrund sind die Aussagen von Slawin und Zwicker nicht mehr als gegensätzlich zu bewerten, sondern beide Autoren kommen zu den gleichen Ergebnissen. Die anfängliche Widersprüchlichkeit ist auf ein Missverständnis der nicht eindeutigen Versuchsbeschreibungen bei Slawin zurückzuführen, das aber durch entsprechende Messungen aufgeklärt werden konnte.

Zusammenfassend wurden damit folgende Erkenntnisse aus den Hörversuchen gewonnen:

- Ein Sinuston maskiert höhere Frequenzen eher als tiefere
- Je größer die Intensität, desto breiter ist der Maskierungsbereich
- Enges Bandrauschen maskiert in etwa wie ein Sinuston
- Breitbandrauschen kann alle Frequenzen maskieren, ein Ansteigen des Breitbandrauschens (weißes Rauschen) um 10 dB hebt auch die Mithörschwelle um 10 dB
- Harmonische Töne, deren Frequenzen eng nebeneinander liegen, verdecken eher als weit auseinander liegende

Diese Aussagen werden mit den folgenden Graphiken nochmals verdeutlicht:

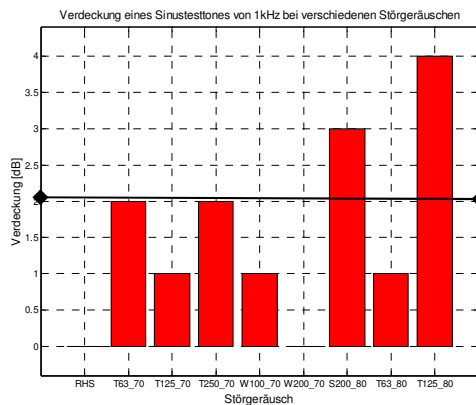


Abbildung 129: Nutzsignal: Sinuston 1kHz

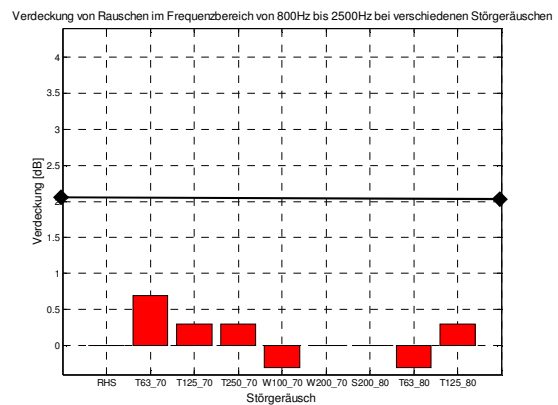


Abbildung 130: Nutzsignal: Rosa Rauschen  
800 – 2500Hz

Zusätzlich zu den Messergebnissen wurde in den beiden obigen Graphiken der Bereich der Messungenauigkeit (ca. 2dB) als schwarzer Balken eingezeichnet. Bei den Achsenbeschriftungen wurden folgende Abkürzungen verwendet:

RHS	→	Ruhehörschwelle
T	→	Terzrauschen
W	→	Weißes Tiefpassrauschen
S	→	Sinuston
T63_70	→	Terzrauschen mit der Mittenfrequenz 63Hz und einem Schalldruckpegel von 70dB
W100_80	→	Tiefpassrauschen mit der Grenzfrequenz 100Hz und einem Pegel von 80dB

Man sieht, dass, obwohl die Störsignale durchwegs Pegel von 70dB bzw. 80dB aufweisen, die maximale Erhöhung der Hörschwelle lediglich 4dB beträgt. Dabei ist besonders zu beachten, dass aufgrund fehlender Mittelungen mehrerer Testpersonen (zwei Messungen an einem Probanden) die Messungenauigkeit noch relativ hoch ist. Es sind also nur zwei Störsignale geprüft worden, die einen minimalen Verdeckungseffekt hervorrufen. Diese schon geringe Auswirkung der Störsignale wird nochmals geringer, wenn das Nutzsignal verbreitert wird (Vergleich Abb. 129 mit Abb. 130). Mit einem breitbandigen Rauschsignal im Frequenzbereich von 800Hz bis 2500Hz, treten keinerlei Maskierungseffekte auf.

Zusätzlich ist wichtig, wie viele Anteile des Nutzsignals in der gleichen Frequenzgruppe liegen wie die Maskierfrequenzen. Frequenzen, die innerhalb liegen werden verdeckt, die Frequenzen außerhalb werden nicht beeinflusst.

### 7.3. Messungen mit Low-noise noise

Äquivalent zu den oben beschriebenen Versuchen wurden Messungen mit Low-noise noise als Störsignal (siehe 3.9. *Verdeckung durch Low-noise noise*) durchgeführt. Dabei wurden die Auswirkungen an vier Testpersonen untersucht. Die Vergleichsmessungen mit herkömmlichen Störsignalen wurden mit diesen vier Testpersonen wiederholt, um die Ergebnisse gegenüberstellen zu können.

#### 7.3.1. Störsignal Terzrauschen

Für diese Untersuchungen wurde Low-noise noise mit Weißem Rauschen verglichen, da bei beiden Signaltypen alle vorhandenen Spektralanteile die gleiche Amplitude besitzen. Beide Maskierer wurden auf Terzrauschen bandbegrenzt, die Mittenfrequenzen wurden auf 125Hz und 250Hz festgelegt (siehe 5.3.2. *Störsignale*). Damit kann mit Gleichung 15 festgestellt werden, dass sowohl das Terzrauschen mit der Mittenfrequenz 125Hz, als auch mit  $f_m = 250\text{Hz}$  als schmalbandig zu bewerten sind. Bei der Versuchsdurchführung wurden die Signale mit gleichem Schalldruckpegel ( $L = 50\text{dB}$  bzw.  $L = 70\text{dB}$ ) abgespielt.

Es wurde festgestellt, dass, wie erwartet (3.9.3. *Verdeckungseffekte bei Verwendung von LNN als Maskierer*), die Hörschwelle durch Zufallsrauschen höher liegt als diejenige, die durch Low-noise noise erzeugt wird. Die Erhöhung liegt dabei im Bereich von 3dB bis 7dB, der Durchschnittswert beträgt 4,25dB. Der Frequenzbereich, in dem nennenswerte Verdeckungserscheinungen auftreten, wird jedoch durch die Verwendung von LNN nicht eingeschränkt, es verändert sich nur die Stärke der Maskierung.

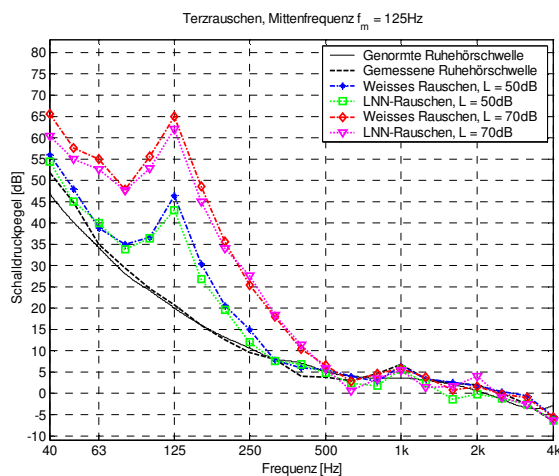


Abbildung 131: Terzrauschen,  $f_m = 125\text{Hz}$

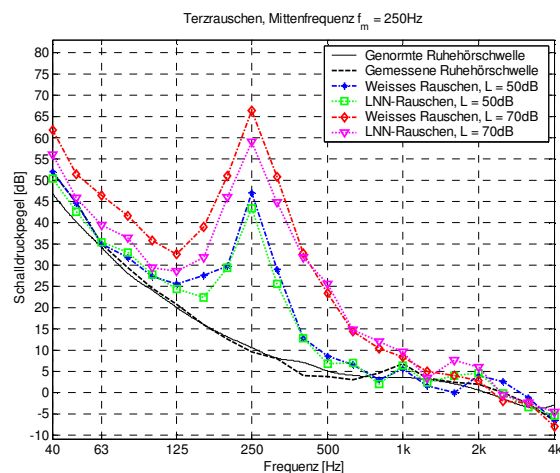


Abbildung 132: Terzrauschen,  $f_m = 250\text{Hz}$



Äquivalent zum Störsignal Zufalls-Terzrauschen, kann auch hier eine steil ansteigende untere Flanke und eine etwas flacher abfallende obere Flanke beobachtet werden.

Auch fällt auf, dass, wie bei der Verdeckung durch Schmalbandrauschen, bei hohen Störpegeln die Mithörschwelle im tieffrequenten Bereich nicht mehr ganz auf die Ruhehörschwelle zurückgeht. Dies ist umso stärker zu sehen, je tieffrequenter das Störsignal bei gleichen Signalpegeln eingespielt wird. Da in diesem Versuch die Ergebnisse von vier Probanden gemittelt wurden, kann dies nicht mehr auf ein untypisches Hörverhalten einer Einzelperson zurückgeführt werden.

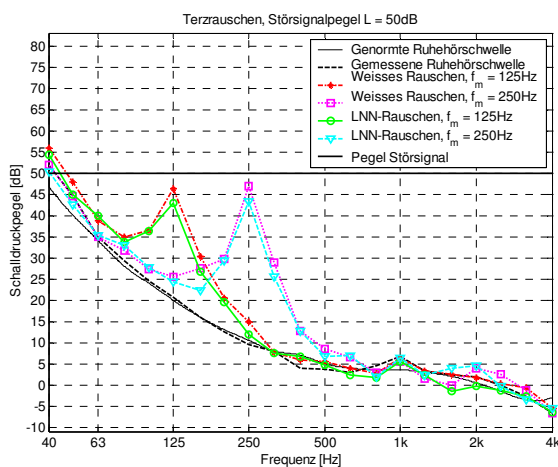


Abbildung 133: Terzrauschen, L = 50dB

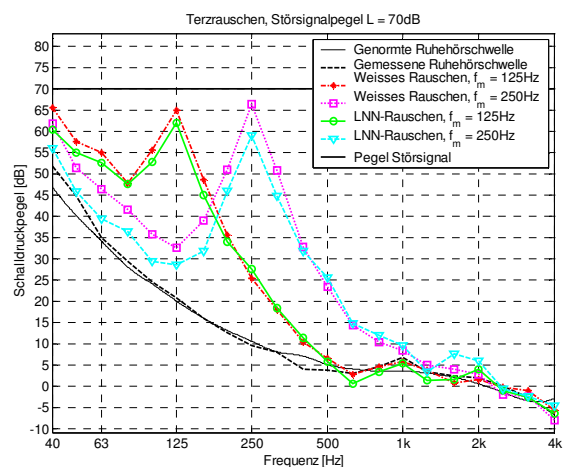


Abbildung 134: Terzrauschen, L = 70dB

Zusätzlich ist, vor allem in Abbildung 134, der Rückgang des Maximums der Hörschwelle zu sehen. Während mit der Mittenfrequenz des Störsignals von 125Hz das Nutzsignal noch mit 62dB abgespielt werden muss, um wahrgenommen zu werden (durchgezogene grüne Kurve), sind bei der Mittenfrequenz von 250Hz bereits 59dB ausreichend (gestrichelte blaue Kurve). Dieser Rückgang ist offensichtlich unabhängig davon, ob als Störsignal Zufallsrauschen oder LNN verwendet wird, da er äquivalent zum zufälligen Terzrauschen ca. 3dB beträgt.

### 7.3.2. Störsignal Tiefpassrauschen

Wenn als Störsignal tiefpassgefiltertes Weißes Rauschen bzw. tiefpassgefiltertes Low-noise noise verwendet wird, erhält man die Ergebniskurven der Abbildungen 135 und 136. Es ist gut zu sehen, dass bis auf die beiden Signale mit  $f_g = 100\text{Hz}$  und  $l = 51\text{dB}$  die Hörschwellen durch Weißes Rauschen und die Kurven durch LNN mit nahezu identischen Werten knapp übereinander liegen. Die Differenz zwischen den einzelnen Messergebnissen beträgt maximal 2,5dB.

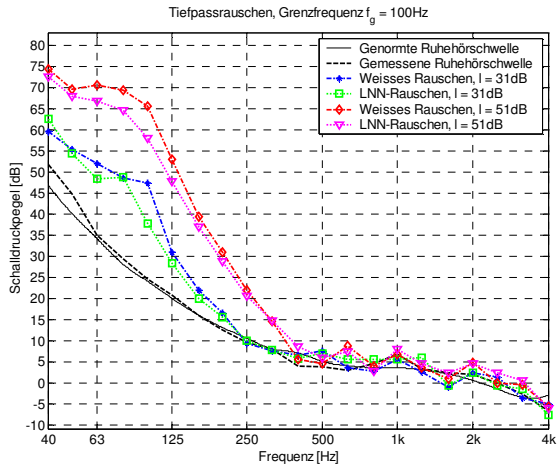


Abbildung 135: Tiefpassrauschen,  $f_g = 100\text{Hz}$

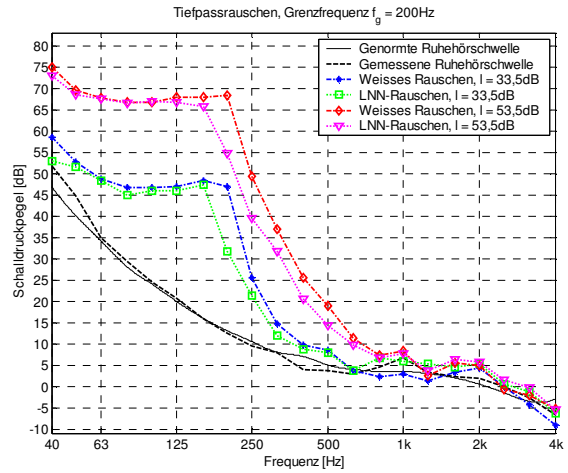


Abbildung 136: Tiefpassrauschen,  $f_g = 200\text{Hz}$

Warum sich dieses Ergebnis jedoch nicht bei allen Kurven einstellt, obwohl der Dichtepiegel der Signale immer konstant gehalten wurde, kann nicht mit Sicherheit gesagt werden. Es ist aber davon auszugehen, dass bei einer großen Anzahl von Versuchspersonen die Differenz bei  $f_g = 100\text{Hz}$ ,  $l = 51\text{dB}$  wesentlich geringer ausgefallen bzw. nicht mehr aufgetreten wäre.

Ein Unterschied tritt jedoch zur Messung mit Zufallsrauschen auf:

Es ist bei der Verwendung von LNN nicht mehr erkennbar, dass mit einer Verbreiterung des Störsignals die Mithörschwelle absinkt. Die Mithörschwelle liegt jeweils auf dem gleichen Pegelniveau knapp unter dem Störsignal, egal ob die Grenzfrequenz  $f_g = 100\text{Hz}$  oder  $f_g = 200\text{Hz}$  beträgt (Abb. 137 und 138). Wenn jedoch tiefpassgefiltertes Weißes Rauschen abgespielt wird, sinkt die Mithörschwelle bei tiefen Frequenzen wieder ab, je breiter das Störgeräusch gefiltert wurde.

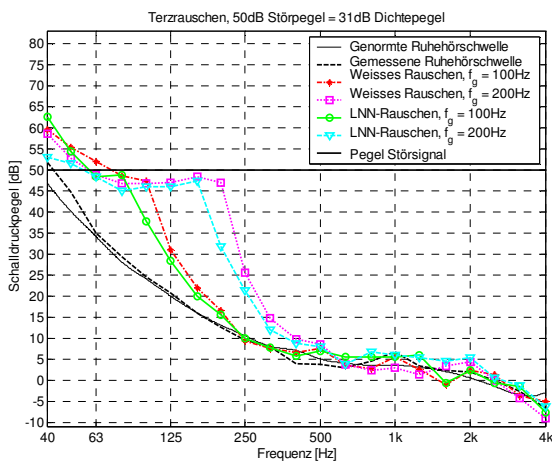


Abbildung 137: Tiefpassrauschen,  $l = 31\text{dB}$

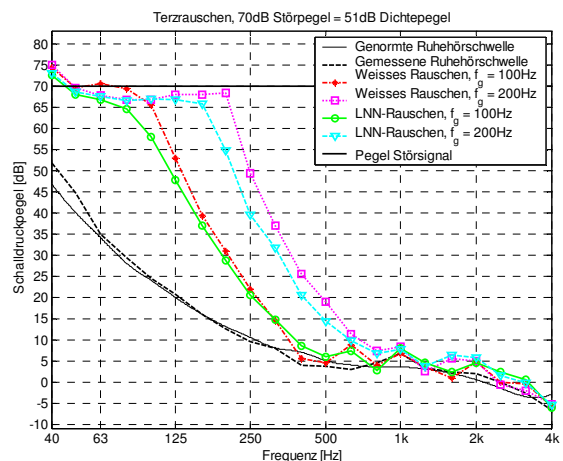


Abbildung 138: Tiefpassrauschen,  $l = 51\text{dB}$

Durch dieses erneute Auftreten des Phänomens kann nun davon ausgegangen werden, dass der bei Zufallsrauschen beschriebene Effekt der Verringerung der Mithörschwelle (7.2.3. *Störsignal Tiefpassrauschen*) kein zufälliges Resultat war. Vielmehr konnte er auch bei mehreren Versuchspersonen beobachtet werden (vorher: Mittelung zweier Ergebnisse mit einer Testperson, jetzt vier Ergebnisse von vier Probanden) und wurde dadurch in seiner Wirkung bestätigt.

### 7.3.3. Zusammenfassung der Ergebnisse und Fazit

Die Signalform Low-noise noise ist vor allem wegen ihrer Eigenschaft der geringstmöglichen Amplitudenschwankung für Untersuchungen im Akustik- und Psychoakustikbereich von großem Interesse.

Der größte Unterschied zum zufälligen Weißen Rauschen mit statistisch verteilten Phasenlagen der vorhandenen Spektralkomponenten liegt bei den durchgeführten Messungen vor allem in einer Verringerung der Hörschwelle bei den schmalbandigen Störgeräuschen. Bei den Untersuchungen mit breitbandigen Störgeräuschen ist bei fast allen Resultaten kein wesentlicher Unterschied zur Hörschwelle mit Weißem Rauschen zu beobachten. Mit diesen Ergebnissen ist eine gute Ausgangslage für die Voraussagbarkeit zukünftiger Versuche mit Low-noise noise geschaffen worden.

Der große Vorteil von Low-noise noise gegenüber Weißem Rauschen liegt darin, dass bei nahezu äquivalenten Ergebnissen wie mit Zufallsrauschen, durch LNN eine genau definierte Signalform vorliegt. Diese kann jederzeit exakt reproduziert werden und unterliegt keinerlei statistischen Schwankungen. Deshalb sollte für zukünftige Untersuchungen bevorzugt diese Signalform verwendet werden.

## 8. Versuche zur Ermittlung der Sprachverständlichkeit

### 8.1. Messumgebung Tieffrequenz-Labor

Alle Versuche zur Ermittlung der Sprachverständlichkeit wurden, wie bereits in Kapitel 5.4. *Messaufbau für die Verständlichkeitstests* beschrieben, im Tieffrequenz-Labor (T-Labor) durchgeführt. Um diese Messumgebung exakt zu definieren, wurden zu Beginn jeder Versuchsreihe die Nachhallzeit an der Abhörposition, die Transferfunktion des Raumes und die Raumimpulsantwort ermittelt.

#### 8.1.1. Die Nachhallzeit

Die Nachhallzeit an der Abhörposition wurde mit dem *Real Time Analyser RTA 840* von *Norsonic* bestimmt. Dafür wurde der Raum mit Rosa Rauschen (Signalquelle: Rauschgenerator des *RTA*) angeregt. Die nach der Beendigung dieses Signales auftretende Abklingkurve wurde am Bezugspunkt mit dem  $\frac{1}{2}$ "-Messmikrofon Type 4190 von Brüel&Kjaer aufgenommen. Zusätzlich wurde sie aus der ebenfalls ermittelten Impulsantwort (siehe 8.1.3. *Die Raumimpulsantwort*) berechnet. Da diese beiden unterschiedlich bestimmten Nachhallzeiten nahezu identische Werte aufwiesen, werden hier zur besseren Übersicht nur die berechneten Werte graphisch dargestellt. Diesen Ergebnissen wurde die Early Decay Time (EDT) gegenübergestellt.

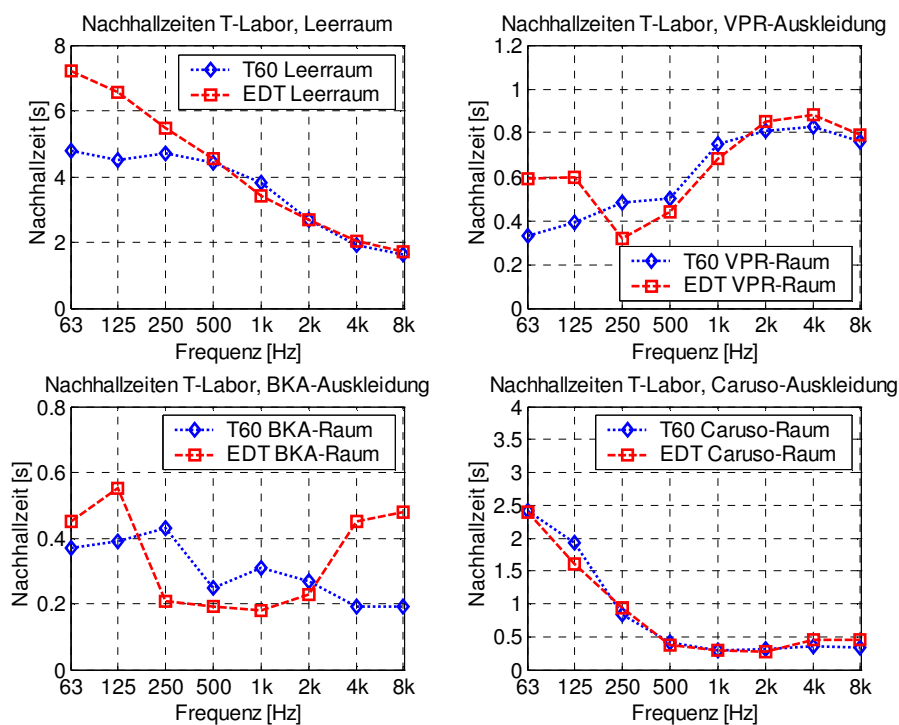


Abbildung 139: Nachhallzeiten der vier simulierten Räume

Im leeren T-Labor ist eine sehr hallige Ausgangssituation vorgegeben. Die Nachhallzeit steigt für alle Frequenzen unter 4kHz auf Werte über zwei Sekunden, unterhalb 1kHz sogar auf über vier Sekunden. Werden in diesen kahlen Raum die Verbund-Platten-Resonatoren (VPR, siehe 5.4.3. *Geplante Raumauskleidungen*) eingebaut, sinkt die Nachhallzeit über den gesamten Frequenzbereich rapide ab. Im tieffrequenten Bereich bis 500Hz werden nur noch Werte von 0,4s bis 0,6s gemessen, oberhalb 1kHz 0,8s bis 1,0s. Dieser Anstieg der Nachhallzeit ab 500Hz aufwärts wird durch die Verwendung von Breitband-Kompakt-Absorber-Modulen (BKA, siehe Punkt 5.4.3 *Geplante Raumauskleidungen*) verhindert. Im gesamten untersuchten Frequenzbereich von 63Hz bis 8kHz werden mit diesen Absorbereinbauten Nachhallzeiten von 0,25s bis 0,45s gemessen. Lediglich die Anfangsnachhallzeit (EDT), die im Pegelbereich von 0dB bis -10dB ausgewertet wird, steigt im tieffrequenten und im hochfrequenten Bereich auf etwas höhere Werte. Werden Caruso-Absorber-Platten (siehe 5.4.3. *Geplante Raumauskleidungen*) verwendet, bleibt die Nachhallzeit ab einer Frequenz von 500Hz aufwärts konstant bei 0,4s. Bei den Frequenzen unterhalb steigt sie allerdings stark an und beträgt bei 63Hz ca. 2,5s.

### 8.1.2. Die Transferfunktion

Die Transferfunktion des Raumes wurde mit einem am Fraunhofer-Institut entwickelten System ermittelt: dem IBP-Analysator. Dabei wird der Raum mit einem MLS-Signal (Maximum Length Sequence) angeregt. Dieses Signal ist eine periodische, zweiwertige Pseudozufallsfolge mit der Periodendauer

$$L = 2^m - 1 \quad \text{Gleichung 41}$$

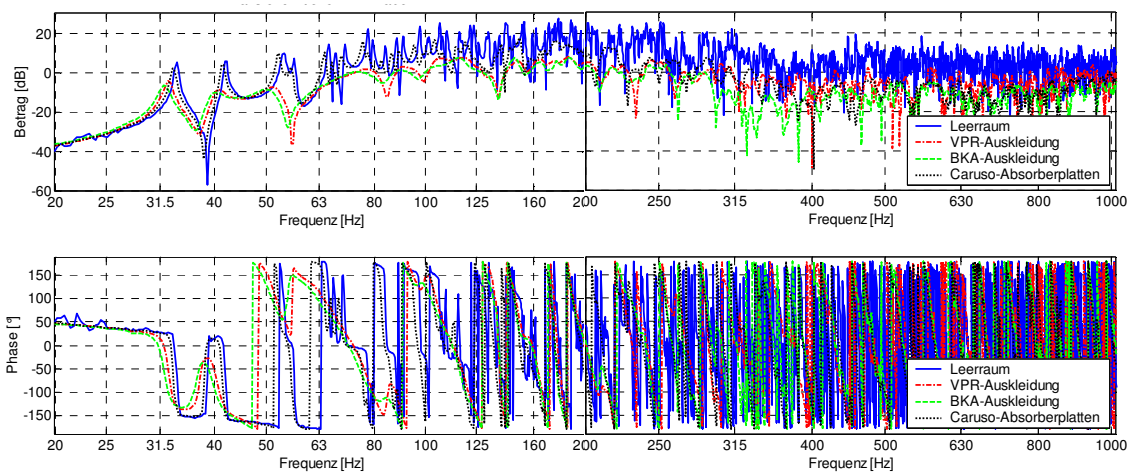
und ermöglicht ein sehr energiereiches Messverfahren. Der theoretische Dynamikgewinn durch die Verwendung dieses Signals gegenüber einer Messung mit einem Einzelimpuls beträgt

$$D_{\text{Gewinn}} = 10 \cdot \log L \approx 10 \cdot \log 2^m = m \cdot 3 \quad [\text{dB}] \quad \text{Gleichung 42}$$

Bei einer typischen MLS-Länge von  $m = 17$  ergibt sich ein theoretischer Dynamikgewinn von über 50dB. In der Praxis erreicht man durch nicht ganz ideale Messsysteme bis zu 40dB [22].

Dieses Erreger-Signal wird direkt am Lautsprecher in einer Raumecke abgegriffen und zusätzlich in der schräg gegenüberliegenden Ecke mit einem  $\frac{1}{2}$ "-Mikrofon von *Brüel&Kjaer* gemessen. Damit berechnet man die Differenz aus Eingangs- und Ausgangssignal einer

Raumdiagonale und erhält so den Raumeinfluss auf das Signal. Dies entspricht auch der Darstellung des modalen Schallfeldes (raumspezifische Pegleinbrüche).



**Abbildung 140: Transferfunktionen des unterschiedlich ausgekleideten T-Labors**

Bei der Darstellung der Transferfunktion ist ein möglichst glatter Verlauf der Messkurven erwünscht [38]. In diesem Fall wird das Signal annähernd linear übertragen und ein unerwünschtes tieffrequentes "Dröhnen" des Raumes [39] tritt nicht auf.

Wie in der obigen Graphik gut zu sehen ist, ergeben sich jedoch im leeren Versuchsraum starke Einbrüche und Überhöhungen (blaue, durchgezogene Kurve), die Werte bis zu +25dB/-55dB erreichen. Diese verringern sich bei der Verwendung von Caruso-Absorptionsplatten im tieffrequenten Bereich nur unwesentlich, im höherfrequenten Bereich ist der Verlauf aber deutlich gerader (schwarze gepunktete Linie). Eine Glättung bereits im tieffrequenten Bereich und damit ein Verhindern des "Dröhrens" erreicht man durch die VPR- bzw. durch die BKA-Module (rote Strich-Punkt-Kurve bzw. grün gestrichelte Linie). Die Einbrüche treten bedeutend weniger oft auf und sind auch wesentlich schwächer ausgeprägt. Zusätzlich werden mit diesen Absorbern auch im höherfrequenten Bereich akzeptable Werte erreicht.

### 8.1.3. Die Raumimpulsantwort

Die Raumimpulsantwort ist die Grundlage der raumakustischen Messtechnik. Aus ihr lassen sich die Gütemaße und die Nachhallzeit des Raumes berechnen [22].

Im T-Labor wurde diese Messung mit dem System *Monkey Forest*<sup>8</sup> durchgeführt, und die Impulsantwort mit einem MLS-Signal (siehe 8.1.2. *Die Transferfunktion*) unter Anwendung der Korrelationsmesstechnik bestimmt. Der große Vorteil dieser MLS-Korrelationstechnik liegt darin, dass, wie bei der Transferfunktion, nicht das Erregersignal einem idealen Impuls gleichen muss, sondern die Autokorrelationsfunktion des Erregersignals. Diese Autokorrelationsfunktion besitzt die Form

$$\Phi_{xx}(\tau) = \lim_{T \rightarrow \infty} \frac{1}{T} \cdot \int_0^T x(t) \cdot x(t + \tau) dt \quad \text{Gleichung 43}$$

Für ein ideales Zufallssignal, das ja keinen inneren Zusammenhang besitzt, gilt

$$\Phi_{xx}(0) \neq 0 \quad \Phi_{xx}(\tau \neq 0) = 0 \quad \text{Gleichung 44}$$

Durch die Verwendung dieser Technik erreicht man – identisch zur Transferfunktion – einen Dynamikgewinn von  $D_{\text{Gewinn}} = m \cdot 3$  [dB] und damit einen sehr großen Signal-Rausch-Abstand. Zusätzlich fallen durch die Berechnung über die Kreuzkorrelation Anteile, die nicht mit dem MLS-Signal korrelieren, heraus. Da natürlich vorkommende Störgeräusche nie mit dem pseudozufälligen MLS-Signal korrelieren, ist diese Messung unanfällig gegen alle natürlichen Störgeräusche [22].

Bei der durchgeführten Messung wurde  $m = 20$  gewählt, und der Raum mit einer MLS der Länge  $2^{20}$  Samples angeregt. Die Systemantwort wurde mit einem *ITA-Kunstkopf digital* gemessen und anschließend zum Messgerät *Monkey Forest* übermittelt. Hier wurde die Systemantwort mit dem bekannten Erregersignal kreuzkorreliert. Das Ergebnis war die periodische Impulsantwort (PIR), woraus sich alle weiteren Größen berechnen lassen [22].

Die in der folgenden Abbildung dargestellte Quadrierung des Schalldrucks entspricht einer Gleichrichtung, die auch vom Ohr beim Übergang vom mechanischen in den neuronalen Teil durchgeführt wird. Der Nachteil besteht darin, dass schwächere Reflexionen gegenüber den stärkeren weniger in Erscheinung treten, was nicht dem menschlichen Hören entspricht [22].

---

<sup>8</sup> Audio Measuring System *Monkey Forest* der Firma *Audio & Acoustics Consulting Aachen*, Version 3.5D

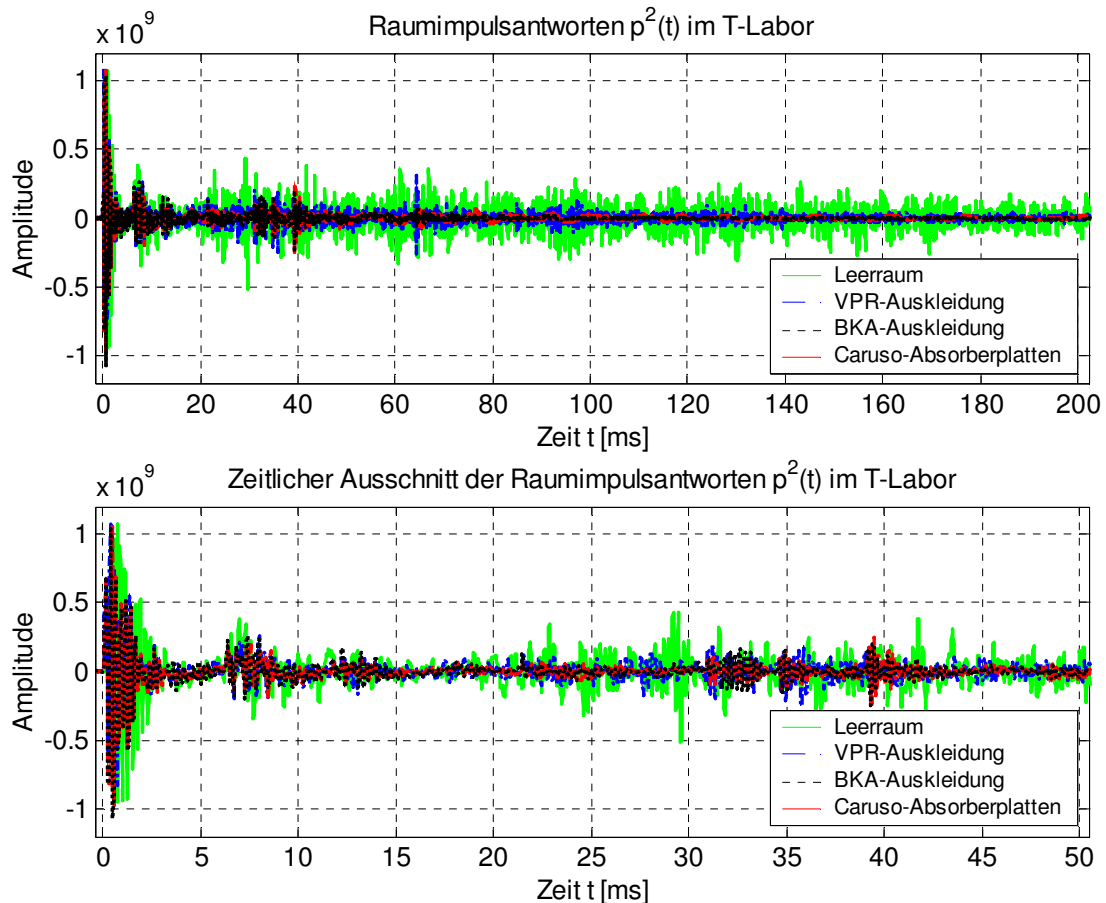


Abbildung 141: Raumimpulsantworten an der Abhörposition

Gut zu erkennen sind die unterschiedlichen Ausprägungen der Raumimpulsantworten in den verschiedenen Räumen. Die Energie-Zeit-Kurven bei eingebauten Absorbern sehen sich dabei sehr ähnlich: Nach dem Direktschall treffen einige Anfangsreflexionen bis 50ms an der Abhörposition ein (Abb. 141 unten). Nach diesen nimmt die Schallenergie konstant ab, ohne dass weitere Überhöhungen oder Einbrüche auftreten (Abb. 141 oben). Eine Ausnahme bildet dabei der VPR-Raum. Hier tritt im Nachhall bei ca. 65ms eine deutliche Reflexion auf. Diese ist jedoch nicht als problematisch zu werten, da sie einzeln auftritt und nicht mit weiteren Reflexionen korreliert. Im Leerraum treten mehrere und stärker ausgeprägte Anfangsreflexionen als in den bedämpften Räumen auf. Zusätzlich dazu nimmt die Energie während der Nachhallzeit nur langsam ab und weist leichte Einbrüche und Überhöhungen auf. Inwieweit sich diese raumakustischen Gegebenheiten auf die Verständlichkeit von Sprache auswirken, sollen die folgenden Logatom-Tests klären.

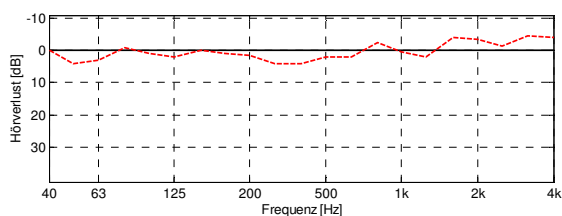


## 8.2. Durchführung des Logatomtests

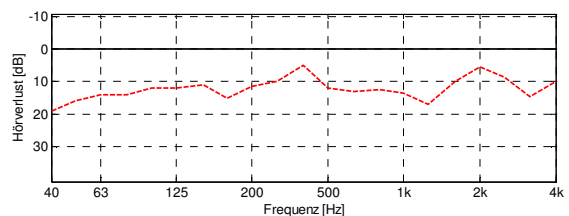
Um die Silbenverständlichkeit in den einzelnen Räumen subjektiv zu ermitteln, wurden Logatom-Tests an insgesamt 15 Testpersonen durchgeführt. Zusätzlich wurde von allen Probanden die Ruhehörschwelle ermittelt, um neben Alter und Muttersprache bei der Auswertung auch das Kriterium "Hörschäden" bewerten zu können. Die vollständigen Kurven sind im Anhang an diese Arbeit angefügt, ausgewählte Einzelbeispiele sind in den folgenden Abbildungen dargestellt.

Die Messung wurde so geplant, dass alle Testpersonen in jedem der vier Räume nacheinander drei Logatomlisten zu hören bekamen und die verstandenen Wörter auf vorgefertigte Auswertbögen schreiben mussten. Dabei wurde eine Logatomliste in ruhiger Umgebung abgespielt, eine mit dem Störgeräusch Tiefpassrauschen bis 250Hz und eine mit einem breitbandigen Rauschen im Frequenzbereich von 20Hz bis 20kHz (siehe 5.3.1. *Nutzsignale* und 5.3.2. *Stör-signale*). Dadurch ergaben sich Versuchsdauern von zehn bis zwölf Minuten, so dass Ermüdungserscheinungen bei den Probanden ausgeschlossen werden konnten.

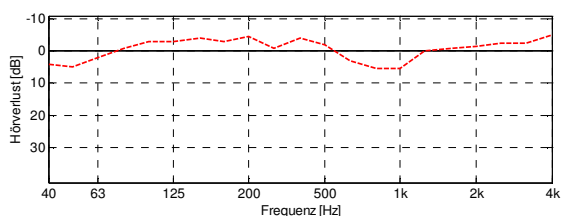
Die Testpersonen wurden nach den oben genannten Kriterien bei der Auswertung in vier verschiedene Gruppen eingeteilt:



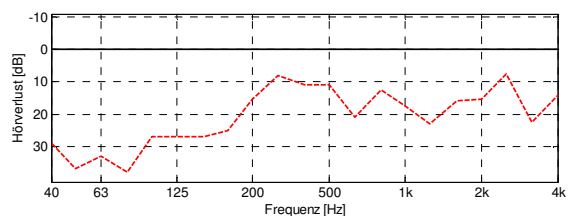
**Abbildung 142: Gruppe 20- bis 50-Jährige, Proband 1**



**Abbildung 143: Gruppe Über-50-Jährige, Proband 10**



**Abbildung 144: Gruppe Nicht-Muttersprachler, Proband 12**



**Abbildung 145: Gruppe Hörgeschädigte, Proband 15**

*Gruppe 20 bis 50 Jahre:* Neun normalhörende Probanden mit einem Alter zwischen 20 und 50 Jahren (siehe Abb. 142).

*Gruppe Über-50-Jährige:* Zwei Personen über 50 Jahre, bei denen altersbedingte Hörverluste vorhanden sind (Abb. 143).

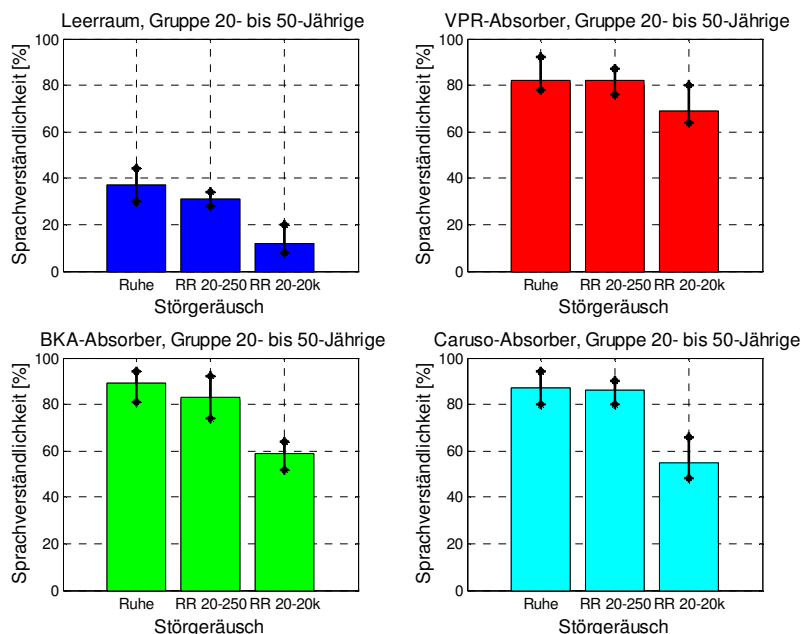
*Gruppe Nicht-Muttersprachler:* Drei normalhörende Probanden, deren Muttersprache nicht deutsch ist (türkisch, chinesisch und ungarisch), die aber gut Deutsch sprechen und verstehen (siehe Abb. 144).

*Gruppe Hörgeschädigte:* Eine Testperson über 50 Jahre, die zusätzlich zum altersbedingten Hörverlust auf einem Ohr nahezu taub ist (Abb. 145).

### 8.2.1. Auswirkungen der Störgeräusche

Die Auswirkungen der einzelnen Versuchsaufbauten auf die Probanden werden im Folgenden besonders ausführlich am Beispiel der Gruppe *20- bis 50-Jährige* aufgezeigt. Dies ist die einzige Gruppe, in der eine Mittelung mit mehr als drei Personen durchgeführt werden konnte (neun Probanden) und deshalb am aussagekräftigsten ist.

Zunächst werden die Auswirkungen der Störgeräusche untersucht. Dazu wurde das Versuchsergebnis als Balkendiagramm dargestellt (Abb. 146), in das zusätzlich die Streuung der Messergebnisse mit eingetragen wurde.



**Abbildung 146: Auswirkungen der Störgeräusche auf die Silbenverständlichkeit innerhalb eines Raumes**

Im leeren, halligen T-Labor ist bereits ohne Störgeräusch eine schlechte Silbenverständlichkeit von lediglich 37% vorhanden. Diese steigt in allen anderen Raumvarianten mit eingebauten Absorbern durchgehend auf sehr gute Ergebnisse von über 80%. Dabei ist der Bereich der

Streuung besonders bemerkenswert: die Verständlichkeit von 80% wurde nur von einer Testperson im VPR-Raum unterschritten, alle anderen Ergebnissen lagen über dieser Grenze.

Wird zusätzlich zu den Logatomen die verstanden werden müssen, ein tieffrequentes Störgeräusch (Rosa Rauschen, tiefpassgefiltert mit der Grenzfrequenz  $f_g = 250\text{Hz}$ ) dauerhaft im Hintergrund abgespielt, verschlechtert sich die Silbenverständlichkeit in allen vier Räumen nicht bzw. nur unwesentlich. Dies ist nachvollziehbar, da die Störung in einem für Sprache untergeordneten Frequenzbereich auftritt (siehe *Kapitel 4* und Untersuchungen *Kapitel 7*).

Erst wenn auch die wichtigen Frequenzanteile (800Hz bis 2500Hz) beeinträchtigt werden (breitbandiges Rosa Rauschen von 20Hz bis 20kHz), nimmt die Verständlichkeit um durchschnittlich 20% ab. Dadurch erreicht man im Leerraum nur noch 12% Silbenverständlichkeit, für die Praxis indiskutabel. In den bedämpften Räumen werden Werte von 55% bis 69% erreicht und können damit nach *Abbildung 47* mit befriedigend bis gut bewertet werden.

Diese Ergebnisse können mit Hilfe entsprechender Diagramme in etwa vorausgesagt werden: Die Silbenverständlichkeit im Ruheraum liegt laut *Abbildung 50* beim Idealwert 96%. Für das tieffrequente Störsignal wird davon ausgegangen, dass sich auf Grund fehlender Spektralanteile im für Sprache wichtigen Frequenzbereich oberhalb 800Hz keine Auswirkungen für die Verständlichkeit ergeben, sie sich also trotz Störung nicht reduziert. Dem gegenüber wirkt sich das breitbandige Störgeräusch deutlich aus: mit einem Signalpegel von 63dB und einem Störpegel von 61dB ergeben sich laut *Abbildung 50* etwa 64% Verständlichkeit. Man sieht, dass im Ruheraum eine sehr gute Verständlichkeit vorausgesagt wird, die mit dem breitbandigen Störgeräusch etwas absinkt. Man erreicht aber immer noch gute Werte (vergleiche *Abb. 47*). Damit sind lediglich grobe Abschätzungen für Räume mit guter Hörsamkeit durchführbar, genaue Werte, v.a. für die Differenzierung unterschiedlicher Raumbedingungen jedoch unmöglich.

In den anderen untersuchten Testgruppen wurden äquivalente Ergebnisse erzielt: Die tieffrequente Störung hat keine oder nur minimale Auswirkungen auf die Verständlichkeit, während sie durch das breitbandige Störgeräusch deutlich vermindert wird.

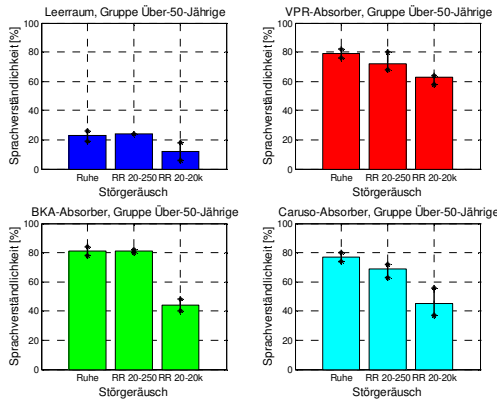


Abbildung 147: Gruppe Über-50-Jährige

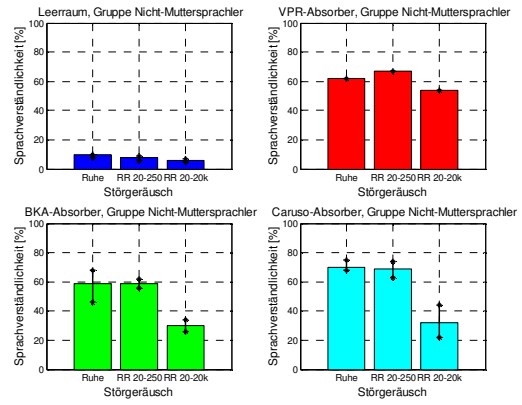


Abbildung 148: Gruppe Nicht-Muttersprachler

Besonders stark wirkt sich dabei die Reduktion bei den Personen, deren Muttersprache nicht deutsch ist (siehe Abbildung 148) und beim hörgeschädigten Proband (Abb. 149) aus. In diesen Gruppen werden nur noch Verständlichkeitswerte bis maximal 30% erzielt, so dass Kommunikation mit anderen Menschen in dieser Umgebung für sie kaum mehr möglich ist.

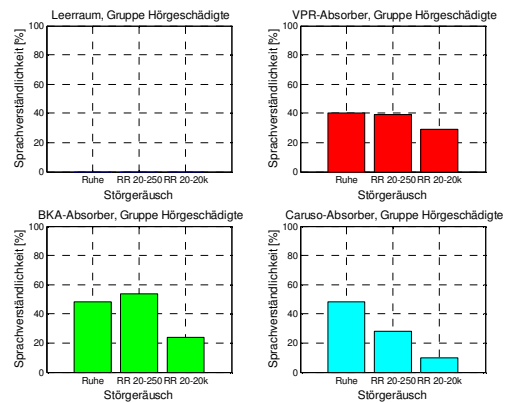
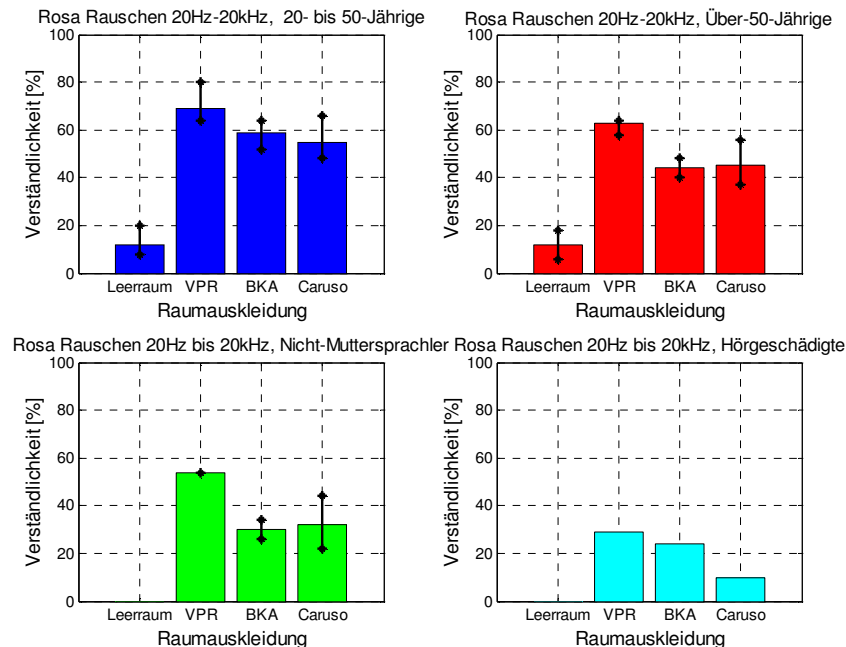


Abbildung 149: Gruppe Hörgeschädigte

Diese Ergebnisse decken sich mit früheren Veröffentlichungen des IBP. Sie sind deshalb sehr wichtig, da bereits knapp 20% der Bundesbürger (60% der Menschen über 65 und 4% derjenigen zwischen 15 und 35 Jahren) schwerhörig sind [2].

Allerdings muss bei den Schlussfolgerungen beachtet werden, dass in diesem Versuch lediglich drei Einzelwerte gemittelt wurden (Gruppe Nicht-Muttersprachler) bzw. nur eine Person in der untersuchten Gruppe enthalten ist (Hörgeschädigte). Eine vom Trend abweichende Auswertung (z.B. durch Unkonzentriertheit bei Versuchsdurchführung) führt sofort zu stark veränderten Ergebnissen, die daher mit Vorsicht interpretiert werden müssen. Eine Trenderkennung und die Bestätigung der bisherigen Resultate (Gruppe der normalhörenden 20- bis 50-Jährigen und frühere Veröffentlichungen) ist jedoch sehr gut möglich.

Zur Verdeutlichung wird hier nochmals dargestellt, wie sich die Ergebnisse der verschiedenen Versuchsgruppen beim breitbandigen Störgeräusch Rosa Rauschen 20Hz bis 20kHz ähneln (Abb. 150):



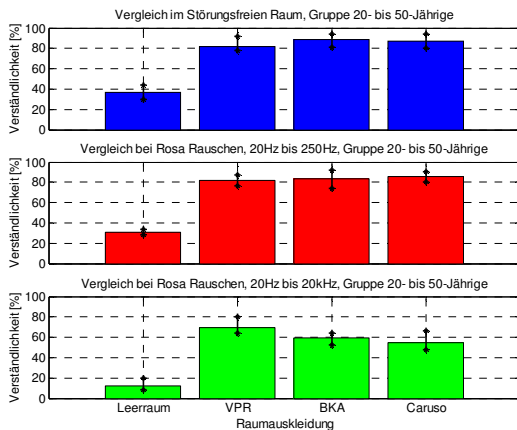
**Abbildung 150: Silbenverständlichkeit in den verschiedenen Versuchsgruppen beim breitbandigem Störgeräusch Rosa Rauschen 20Hz bis 20kHz**

Bei der subjektiven Befragung aller Versuchspersonen wurde durchgängig erklärt, dass das tieffrequente Geräusch bedeutend weniger störend erlebt wurde als das breitbandige Rauschen. Es wurde als "nicht besonders störend", "weniger unangenehm" und oft auch "leiser" empfunden, obwohl beide Störsignale gleich laut waren. Diese Aussagen wurden durch die obigen Messergebnisse bestätigt.

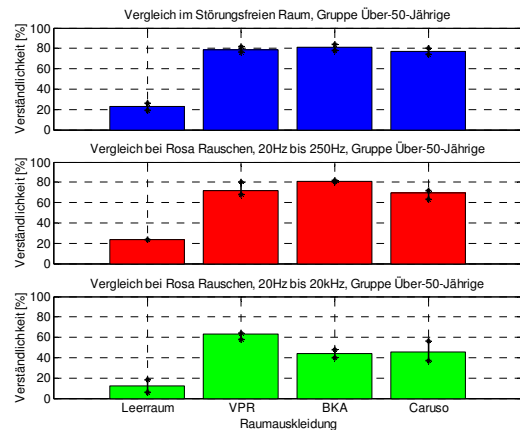
### 8.2.2. Auswirkungen der Raumzustände

Nun sollen die Einflüsse der verschiedenen Raumauskleidungen auf die Messergebnisse untersucht werden: welche Unterschiede ergeben sich durch tieffrequente, hochfrequente und breitbandige Absorption der Schallereignisse?

Dafür werden die Ergebnisse der vier verschiedenen Räume bei gleicher Störung (Ruhe, Tiefpassrauschen oder Rosa Rauschen) in jeweils einem Teildiagramm zusammengefasst.



**Abbildung 151: Silbenverständlichkeit, Gruppe 20- bis 50-Jährige**



**Abbildung 152: Silbenverständlichkeit, Gruppe Über-50-Jährige**

Deutlich erkennbar ist auch hier wieder die oben beschriebene Tatsache, dass im ungestörten Zustand alle Räume mit Absorberauskleidung ähnliche Werte für Verständlichkeit aufweisen, die durch zusätzlich vorhandenes Tiefpassrauschen nicht oder nur unwesentlich reduziert werden (erstes (blaues) und zweites (rotes) Teildiagramm).

Eine Überraschung ergab sich jedoch bei der Auswertung der Ergebnisse mit Verwendung des breitbandigen Rauschens: Werden, während das Störgeräusch (20Hz bis 20kHz) abgespielt wurde, vor allem die tiefen Frequenzen bis 500Hz bedämpft und die höheren Frequenzanteile nicht verändert (VPR-Raum), war weiterhin eine gute Verständlichkeit im Bereich von 64% bis 69% vorhanden (siehe Abb. 151 und 152). Vermutet wurde eigentlich, dass bei zusätzlicher Bedämpfung der hohen Frequenzanteile die Verständlichkeit wieder zunimmt. Die Frequenzen bis 250Hz bewirken keine Änderung der Verständlichkeit. Werden die durch das breitbandige Rauschen neu dazugekommenen Spektralanteile über 250Hz von entsprechenden Absorbern (BKA bzw. Caruso) bedämpft, erzwingt man ähnliche Bedingungen wie mit dem tieffrequenten Störersignal und die Verständlichkeit müsste noch bessere Werte annehmen.

Diese Erwartung bestätigte sich jedoch nicht. Im Gegenteil verringerte sich die Verständlichkeit durch die Verwendung von breitbandig absorbierenden BKA-Modulen (siehe Abb. 87) um 10% bis 20% (siehe Abb. 151 und 152, untere grüne Teildiagramme) auf 59% bei der normalhörenden Gruppe bzw. auf 44% bei den über 50-Jährigen. Es ergaben sich auch kaum andere Werte, wenn durch die Verwendung von Caruso-Absorbern die Störung nur hochfrequent absorbiert wurde: 55% bei den 20- bis 50-Jährigen, 45% bei den älteren Testpersonen.

Dies bedeutet: im ungestörten Zustand und bei der Verwendung von tieffrequentem Störmaterial ergibt sich kein Unterschied welche Raumauskleidung verwendet wird, die Ergebnisse

sind nahezu identisch. Wird jedoch ein Störgeräusch mit Spektralkomponenten über den gesamten Hörbereich eingespielt, ist die Verständlichkeit stark abhängig vom Absorbertyp. Wird das Signal breitbandig bedämpft (BKA) ergeben sich ähnliche Werte wie mit Absorbieren, die erst im höherfrequenten Bereich wirken (Caruso): die Verständlichkeit sinkt um gut 30% gegenüber dem ungestörten Raum. Wird jedoch nur tieffrequent bedämpft, steigt die Verständlichkeit wieder an und erreicht gute Werte, die ca. 15% unter denen des Ruheraumes liegen. Diese Ergebnisse, die sich auch in den Gruppen der Nicht-Muttersprachler und der Hörgeschädigten (Abb. 153 und 154) bestätigen, unterstützen die bisher oft bezweifelten Aussagen in Veröffentlichungen, dass ein Anstieg der Nachhallzeit zu höheren Frequenzen hin eher toleriert werden kann als ein Abfall [25].

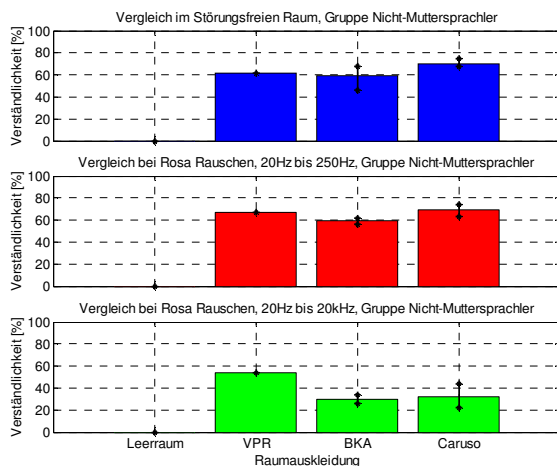


Abbildung 153: Silbenverständlichkeit, Gruppe Nicht-Muttersprachler

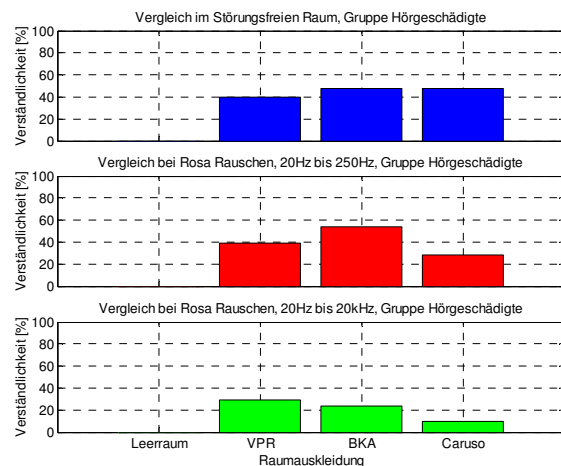


Abbildung 154: Silbenverständlichkeit, Gruppe Hörgeschädigte

Zusätzlich zu diesem Messergebnis beurteilten nahezu alle Probanden die Testumgebung mit **VPR-Absorbieren** als die angenehmste. Die *Behaglichkeit* während des Versuchs wurde mit *gut* angegeben und die *Verständlichkeit* der Logatome *hoch* eingeschätzt.

Dem gegenüber wird die Situation mit eingebauten **Caruso-Absorbieren** deutlich schlechter bewertet. Alle Testpersonen waren sich einig, dass sie sich in diesem bedämpften Raum nach dem unbedämpften Raum am *unwohlsten* gefühlt haben, dass der Versuch in dieser Umgebung *sehr unbehaglich* war und dass die *Verständlichkeit sehr schlecht* ist. Dabei war auffällig, dass die Verständlichkeit durchwegs schlechter eingeschätzt wurde als sie tatsächlich gemessen wurde. Diese Aussagen können als Verunsicherung darüber gewertet werden, ob man selbst beim Sprechen gut vom Gegenüber gehört wird. Der Lombard-Effekt (siehe Kapitel 2.4. *Anpassung und Cocktailparty-Effekt*) ist deshalb vermutlich bei den Räumen mit BKA-

Auskleidung und Caruso-Absorbern stärker ausgeprägt als mit VPR-Modulen, da der Störpegel breitbandig angehoben ist.

Bei der Verwendung von breitbandig absorbierenden **BKA-Modulen** wurden *sehr unterschiedliche Angaben* gemacht. Dabei unterschieden sich vor allem die Aussagen der 20- bis 50-Jährigen und der Über-50-Jährigen von denen der ausländischen Testpersonen. Während die ersten beiden Gruppen die Behaglichkeit zwar zwischen VPR und Caruso einstuften, dabei aber tendenziell eher gute Noten vergaben (eine Testperson befand diesen Raum sogar als den angenehmsten), wurden von den Personen mit nicht-deutscher Muttersprache schlechte Urteile abgegeben.

Die subjektiv empfundene Sprachqualitätsverbesserung beruhte im Wesentlichen auf der Reduzierung des Nachhalls im tieffrequenten Bereich (sogenanntes "Dröhnen"). Für diese Erkenntnisse wurden bereits bei einer früheren raumakustischen Untersuchung [40] Hinweise gefunden, die jetzt untermauert und bestätigt wurden. Die subjektiv empfundene Sprachqualität ist in Räumen mit Nachhallzeiten, die zu hohen Frequenzen hin ansteigen, am günstigsten.

### 8.3. Vergleich der subjektiven Ergebnisse mit berechneten Werten

Da es aufwändig ist Logatomtests durchzuführen, ist es sehr interessant, inwieweit sich die durch die subjektiven Versuche ermittelten Ergebnisse auch durch objektive Messungen vorhersagen lassen und in welchem Genauigkeitsbereich diese Prognosen liegen.

Deshalb wurden nach den in Kapitel 4.3. *Berechnung der Sprachverständlichkeit durch physikalische Parameter* beschriebenen Vorgehensweisen aus der Impulsantwort und den Nachhallzeiten folgende Größen berechnet:

- Schwerpunktzeit TS
- Objektive Silbenverständlichkeit v
- Artikulationsindex AI
- Deutlichkeitsgrad  $D_{50}$  nach Thiele
- Deutlichkeitsmaß  $C_{50}$
- Artikulationsverlust  $AI_{\text{cons}}$
- Sprachübertragungsindex STI
- Schneller Sprachübertragungsindex RASTI



Dabei ist zu beachten, dass die Objektive Silbenverständlichkeit  $v$  die einzige Größe ist, bei der das Störsignal in das Endergebnis miteinbezogen werden kann. Nur hier können Werte sowohl für den ungestörten Raum, als auch für die Situationen mit Störgeräusch ermittelt werden. Bei allen anderen Größen, kann lediglich eine Prognose für die ruhige Messumgebung aufgestellt werden. Zusätzlich dazu muss der Verständlichkeitsverlust durch die Störgeräusche ermittelt und mit diesem das störungsfreie Ergebnis gewichtet werden. Wie stark die Reduktion ausfällt, kann entweder mit dem Artikulationsindex AI berechnet werden oder für Störsignale, die dem Nutzsignal ähnlich sind, aus Abbildung 50 abgelesen werden.

### 8.3.1. Schwerpunktzeit TS

Bei der Schwerpunktzeit TS kann für Werte unter 70ms von guter Sprachverständlichkeit ausgegangen werden. Dies bestätigt sich auch mit den Ergebnissen des T-Labors:

Im Leerraum beträgt die Schwerpunktzeit 184ms, liegt also deutlich über dem gewünschten Grenzwert. Die Hörsamkeit des Raumes wird deshalb mit schlecht bewertet, was durch die subjektiven Tests bestätigt wurde. In den mit Absorbern ausgekleideten Räumen liegen die Schwerpunktzeiten durchwegs unterhalb der geforderten 70ms:  $TS_{VPR} = 47\text{ms}$ ,  $TS_{BKA} = 22\text{ms}$  und  $TS_{Caruso} = 37\text{ms}$ . Deshalb kann in allen drei Fällen von gleichmäßig guter Verständlichkeit ausgegangen werden. Diese Werte wurden ebenfalls durch die Logatomtests belegt.

### 8.3.2. Deutlichkeitsmaß $C_{50}$

Da das Deutlichkeitsmaß  $C_{50}$  die Pegeldifferenz der Schallenergie vor und nach 50ms angibt, besitzt es große praktische Bedeutung [22]. Im T-Labor wurden dafür folgende Werte ermittelt:

$C_{50, \text{Leerraum}}$	=	- 4,09 dB	→	ausreichende Verständlichkeit
$C_{50, VPR}$	=	4,99 dB	→	sehr gute Verständlichkeit
$C_{50, BKA}$	=	12,06 dB	→	sehr gute Verständlichkeit
$C_{50, Caruso}$	=	8,24 dB	→	sehr gute Verständlichkeit

Auch hier bestätigen sich die subjektiv ermittelten Werte mit den objektiv gemessenen und berechneten Größenordnungen.

Interessanter und auch wichtiger als die Einteilung der Hörsamkeit des Raumes in schlecht, befriedigend, gut und sehr gut ist die Bestimmung der tatsächlichen Verständlichkeitswerte in Prozentangaben. Inwieweit diese vorausgesagt werden können, sollen die folgenden beiden Kapitel klären.

### 8.3.3. Artikulationsindex AI und Silbenverständlichkeitsfaktor $V_{S/N}$

Mit dem Artikulationsindex AI können Verständlichkeitsverluste berechnet werden, die durch nicht stark fluktuierende Störgeräusche oder durch Frequenzbandbegrenzungen hervorgerufen werden [22]. Damit ist diese Größe ideal, um die Reduktion der Silbenverständlichkeit durch die beiden Störgeräusche Tiefpassrauschen mit  $f_g = 250\text{Hz}$  und breitbandiges Rosa Rauschen zu bestimmen.

Es ergeben sich folgende Größen: Die für den leeren Ruheraum angenommene Verständlichkeit von 100% wird durch die Berechnung bestätigt (99,99%). Dies bedeutet, dass die Ergebnisse, die ohne Störgeräusch erzielt wurden, ohne weitere Korrektur als Endresultat übernommen werden können. Wird jedoch das tieffrequente Störgeräusch zusätzlich zum Nutzsinal abgespielt, sinkt der Artikulationsindex auf 93,2%. Es ergibt sich ein Verständlichkeitsverlust von 6,8% gegenüber der ruhigen Umgebungssituation. Wird anstelle des Tiefpassrauschens das breitbandige Störgeräusch verwendet, sinkt der AI auf nur noch 33,0%, die Verständlichkeit wird um 2/3 reduziert!

Dieser letzte Wert von 67% Verständlichkeitsverlust erscheint sehr hoch. Deshalb wird ein zweiter Korrekturfaktor als Vergleichswert herangezogen: der Silbenverständlichkeitsfaktor  $V_{S/N}$ , der in Abbildung 50 dargestellt ist. Bei diesem Faktor tritt jedoch das Problem auf, dass bei seiner Verwendung überprüft werden muss, inwieweit das Nutzsinal mit dem Störsignal übereinstimmt. Nur wenn sich beide Signale ähnlich sind, kann dieser tabellarische Wert tatsächlich verwendet werden. Im untersuchten Fall muss davon ausgegangen werden, dass der Silbenverständlichkeitsfaktor nur für das breitbandige Störgeräusch verwendet werden kann. Nur hier sind Spektralanteile im für Sprache wichtigen Frequenzbereich von 800Hz bis 2,5kHz vorhanden und nur dieses Rauschen beeinflusst die Verständlichkeit deutlich. Deshalb wird der aus der Tabelle herausgelesene Wert von 65% für das breitbandige Rauschen angewendet (Reduktion um ca. 1/3) und mit den Ergebnissen mit AI-Faktor verglichen.

8.3.4. Objektive Gütemaße:  $v$ ,  $D_{50}$ ,  $1-AI_{\text{cons}}$ , STI und RASTI

Die erste objektiv ermittelte Größe ist die **Objektive Silbenverständlichkeit  $v$** . Da diese sowohl die räumliche Gegebenheit als auch das Störsignal berücksichtigt und in das Endergebnis miteinbezieht, wird hier weder der Artikulationsindex AI noch der Silbenverständlichkeitsfaktor  $V_{S/N}$  benötigt. Bei den anderen vier Definitionen wird für die Ergebnisse mit Störgeräusch der "Ruhe-Wert" mit dem Korrekturfaktor AI bzw.  $V_{S/N}$  bewertet.

Der **Artikulationsverlust  $AI_{\text{cons}}$**  bestimmt dabei, wie durch seinen Namen bereits ausgedrückt, den Verlust, also die Reduktion der Verständlichkeit. Um diese Berechnung trotzdem mit den anderen Größen vergleichen zu können, wird in den folgenden Graphiken anstelle von  $AI_{\text{cons}}$   $1-AI_{\text{cons}}$  aufgetragen, so dass wieder, wie gewünscht, die tatsächlich vorhandene Verständlichkeit dargestellt wird. Die mit dem Messsystem *Monkey Forest* bestimmten **STI**- und **RASTI**-Werte sind einander sehr ähnlich und unterscheiden sich um maximal 4%. Deshalb wird in der folgenden Graphik auf die Darstellung der RASTI-Werte verzichtet, um eine größere Übersichtlichkeit zu erreichen. Alle berechneten Werte können zusätzlich dazu im Anhang aus den entsprechenden Tabellen abgelesen werden.

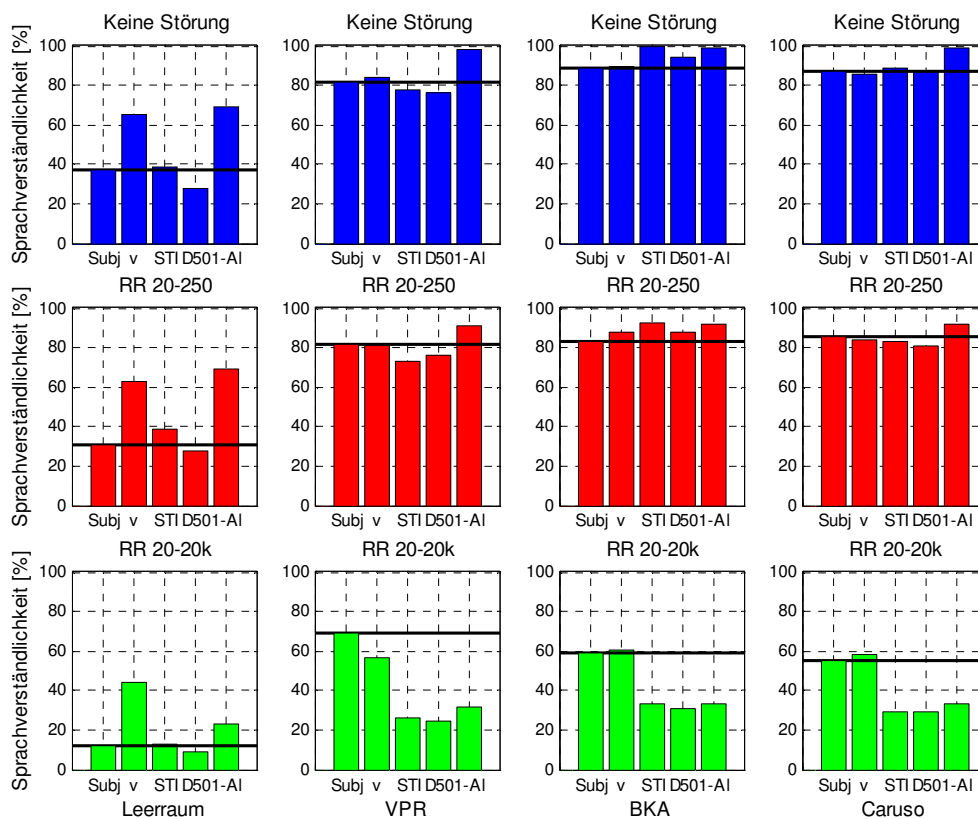


Abbildung 155: Vergleich Subjektive – Objektive Verständlichkeit,  
Korrekturfaktor:  $AI_{250} = 93,2\%$ ,  $AI_{20k} = 33\%$

Wie deutlich zu sehen ist, ergeben sich mit Verwendung des **Artikulationsindex AI als Korrekturfaktor** vor allem **im Leerraum** durch die verschiedenen Berechnungsarten sehr unterschiedliche Verständlichkeitswerte (Abb. 155, 1. Spalte). Dem tatsächlichen Wert am nächsten kommt dabei in allen Fällen (keine Störung, Tiefpassrauschen und breitbandiges Rosa Rauschen) der STI, gefolgt vom Deutlichkeitsgrad  $D_{50}$ . Dabei liegt der STI jeweils knapp über dem subjektiven Ergebnis, der  $D_{50}$  hingegen immer knapp darunter. Die objektive Verständlichkeit  $v$  und der Artikulationsverlust erzielen deutlich zu hohe Werte, die nicht mehr akzeptiert werden können.

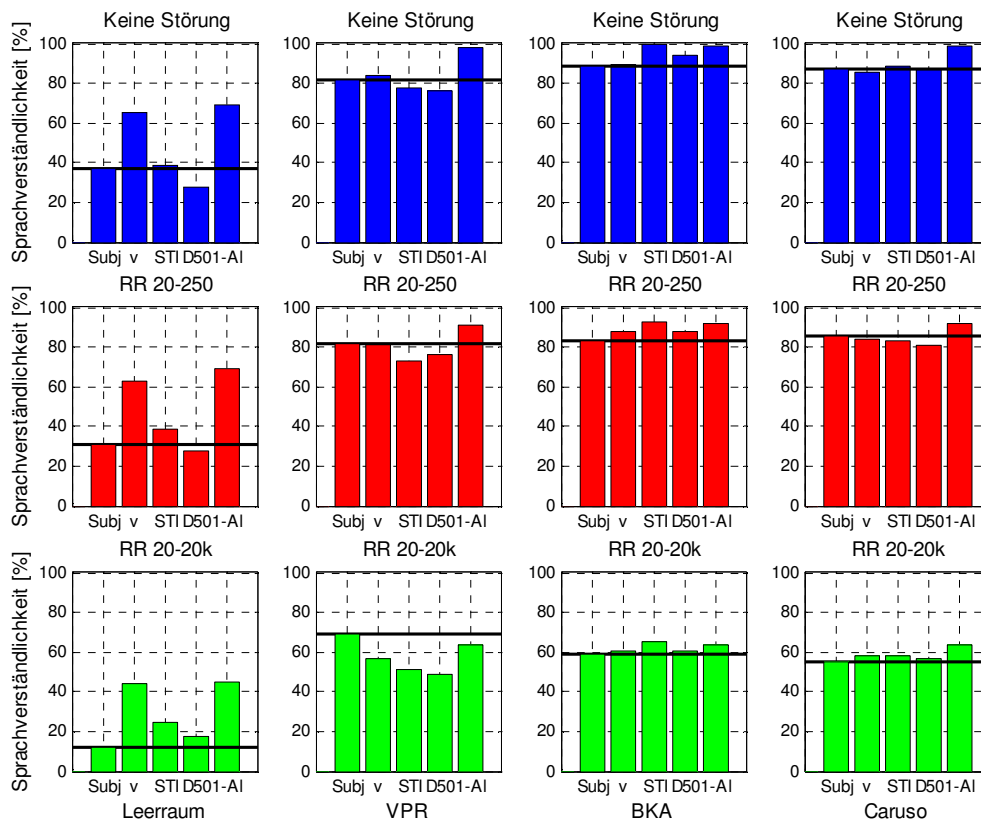
In allen **mit Absorbern ausgekleideten Räumen**, also in den Raumvarianten mit guter Hörsamkeit, erreicht man **im Ruhezustand und mit tieffrequentem Störgeräusch** bei allen objektiv berechneten Größen eine sehr gute Übereinstimmung mit den tatsächlich vorhandenen Werten. Die exaktesten Ergebnisse werden dabei von der Objektiven Verständlichkeit  $v$  erzielt (Abweichungen von maximal 5%), während STI,  $D_{50}$  und  $1-AI_{\text{cons}}$  etwas stärker um das Idealergebnis pendeln. Betrachtet man jedoch die Darstellung der Situation **mit breitbandigem Störgeräusch** (Abb. 155, 3. Zeile), so werden große Unstimmigkeiten sichtbar. Zwar wird mit der Objektiven Verständlichkeit  $v$  noch gute Ähnlichkeit zu den subjektiven Testergebnissen erreicht, die restlichen Werte fallen jedoch deutlich zu klein aus. Dabei fällt auf, dass dieses Problem nur bei denjenigen Ergebnissen auftritt, die mit dem Artikulationsindex  $AI = 33\%$  gewichtet werden. Offensichtlich ist dieser berechnete Korrekturfaktor falsch. Die obige Annahme, dass eine Verständlichkeitsreduktion von 67% zu hoch erscheint, hat sich bestätigt.

Verwendet man statt dessen für die **Korrektur** den **Silbenverständlichkeitsfaktor**  $V_{S/N}$ , ergeben sich die Resultate aus Abbildung 156, 3. Zeile. Die ersten beiden Zeilen (blau und rot) wurden nicht verändert, deshalb wird hier darauf nicht eingegangen.

Im halligen **Leerraum** erzielt lediglich der Deutlichkeitsgrad  $D_{50}$  mit einer Abweichung von 6% eine gute Übereinstimmung zur tatsächlichen Situation. Der STI liegt zwar mit 13% Differenz schon deutlich zu hoch, liefert aber trotzdem noch das Ergebnis einer völlig unzureichenden Verständlichkeit. Dem gegenüber sind die Objektive Verständlichkeit und der Artikulationsverlust  $1-AI_{\text{cons}}$  drastisch zu hoch (32% bzw. 33% Unterschied) und berechnen bei entsprechender Auswertung ein falsches Ergebnis mit schlechter, aber dennoch ausreichender Verständlichkeit.

Die objektiv berechneten Gütemaße erzielen in den Räumen mit **BKA- und Caruso-Auskleidung** durchgängig stimmige Lösungen. Die Objektive Verständlichkeit  $v$ , der STI

und der Deutlichkeitsgrad pendeln dabei mit einer maximalen Abweichung von 6% um den subjektiven Wert. Der Artikulationsverlust ist mit 9% Differenz immer noch akzeptabel.



**Abbildung 156: Vergleich Subjektive – Objektive Verständlichkeit,**  
**Korrekturfaktor:  $AI_{250} = 93,2\%$ ,  $V_{S/N, 20k} = 65\%$**

Der **VPR-Raum** stellt bei dieser Auswertung eine Besonderheit dar. Er liefert in der gesamten Untersuchung den einzigen Fall, bei dem alle berechneten Werte deutlich zu gering ausfallen. Sogar der Artikulationsverlust, der in allen anderen Situationen durchgängig zu hohe Lösungen erzielt, liegt dieses mal unter dem subjektiv ermittelten, realen Ergebnis. Dabei erreichen diese Abweichungen Größenordnungen von 5% ( $1-AI_{cons}$ ) bis zu 20% ( $D_{50}$ ) und bilden das schlechteste Resultat in den ausgekleideten Räumen. Lediglich die Untersuchungen in den Leerräumen brachten noch schlechtere Lösungen.

#### 8.4. Zusammenfassung der Ergebnisse

Auf der Grundlage von Logatomtests wurde in vier verschiedenen Umgebungssituationen die subjektive Silbenverständlichkeit mehrerer Personengruppen ermittelt. Dabei wurde vor allem

auf die Auswirkungen unterschiedlicher Störungen und auf die Nachhallzeiten in den Messumgebungen eingegangen.

Es stellte sich heraus, dass **Störgeräusche** erst dann einen merklichen Einfluss auf die Verständlichkeit gewinnen, wenn sie Frequenzkomponenten im für Sprache wichtigen Bereich besitzen (800Hz bis 2500Hz). Besonders stark war dies bei der älteren Gruppe (Über-50-Jährige) und bei der hörgeschädigten Person zu beobachten. Bei den ausländischen Probanden, lag die Verständlichkeit zwar insgesamt niedriger als bei Testpersonen mit deutscher Muttersprache, die Reduktion durch Störgeräusche lag jedoch in der gleichen Größenordnung wie bei der deutschen Versuchsgruppe der 20- bis 50-Jährigen.

Die **Nachhallzeit** wirkt sich auf die Verständlichkeit von Sprache in mehrfacher Hinsicht aus. Zum einen sinkt die Anzahl der verstandenen Silben mit steigender Nachhallzeit. Zum anderen hat es auch einen Einfluss, welchen Verlauf die Nachhallzeit über der Frequenz aufweist. So werden mit einer zu tiefen Frequenzen hin fallenden Nachhallzeit (VPR) im Durchschnitt um 10% bis 20% mehr Logatome verstanden als bei konstantem oder abfallendem Verlauf. Diese Größenordnung wurde bei allen untersuchten Gruppen bestätigt.

Um die Verständlichkeit, die in einem zu untersuchenden Raum vorhanden ist, nicht mit aufwändigen Logatomtests bestimmen zu müssen, ist es sinnvoll, diese mit geeigneten raumakustischen Gütemaßen objektiv zu berechnen. Dadurch ergeben sich wesentlich kürzere Messzeiten und es werden keine Probanden benötigt.

Dies hat sich jedoch in halligen, unhörsamen Räumen als problematisch herausgestellt. Die berechneten Ergebnisse wichen teilweise stark von den tatsächlichen, subjektiv ermittelten Werten ab und lieferten unbefriedigende Resultate. Wurden jedoch Nachhallzeiten gemessen, die im Durchschnitt unter eineinhalb Sekunden lagen, erzielten die Objektive Silbenverständlichkeit  $v$ , der Deutlichkeitsgrad  $D_{50}$ , der Speech Transmission Index STI und der Artikulationsverlust  $Al_{\text{cons}}$  gute Übereinstimmung mit der tatsächlich vorhandenen Situation. Die Lösungen der Berechnungen pendelten mit nur geringen Differenzen um den Realwert.

## 9. Schlussfolgerung

### 9.1. Hörschwellenversuche

Die Untersuchung der unterschiedlichen Hörschwellen von Slawin und Zwicker, sowie deren Auswirkungen auf die Verständlichkeit von Sprache, lieferten aufschlussreiche Erkenntnisse, die bisherige Missverständnisse aufklären konnten.

Ein lautes Störsignal verursacht gegenüber einem Nutzsignal ein Auftreten von Verdeckungserscheinungen. Diese sind umso deutlicher, je größer der Pegel des Maskierers ist. Entgegen der Darstellung Slawins wurden die Veröffentlichungen von Zwicker bestätigt, dass Verdeckung nur in unmittelbarer Umgebung der Frequenzkomponenten des Störsignals auftritt. Dabei erreicht die Hörschwelle an der Störfrequenz ihr Maximum, das knapp unter dem Schalldruckpegel des Maskierers liegt.

Die Aussagen von Slawin, dass ein tieffrequenter Sinusstörton Verdeckungseffekte bis weit in den kHz-Bereich hinein verursacht, wurden hingegen widerlegt. Es muss davon ausgegangen werden, dass diese Interpretation der Veröffentlichungen auf Grund einer mehrdeutigen Formulierung und damit durch ein Missverständnis zustande gekommen ist. Slawin meint nicht einen Sinuston (Ton im physikalischen Sinn), wenn er seine Ergebnisse präsentiert, sondern einen Ton im musikalischen Sinn, einem Tongemisch. Nur mit dieser Annahme können die dargestellten Kurven nachvollzogen und auch bestätigt werden. Unter diesem Blickwinkel ergänzen sich auch die Ausführungen von Slawin und Zwicker. Das Missverständnis der Interpretation konnte ausgeräumt werden.

Für zukünftige Versuche mit Hörschwellen wurde außerdem die Verwendungsmöglichkeit von Low-noise noise (LNN) untersucht. Diese Signalform besitzt die Eigenschaft, dass die Frequenzen und Amplituden der enthaltenen Spektralkomponenten nahezu frei wählbar sind, während mit einer exakten Vorgabe zur Berechnung der Phasenwinkel der Einzelkomponenten die Amplitudenschwankung des Signals minimiert wird. Dies ist ein bedeutender Vorteil gegenüber herkömmlichen Testsignalen (z.B. Weißes Rauschen), deren Amplitudenverteilung zufällig definiert wird. Der Pegel der Signale musste bis jetzt über längere Zeit gemessen werden, um die Schwankung zu mitteln und einen korrekten Durchschnittswert zu erhalten. Auch ist es nicht möglich, zweimal ein exakt gleiches Signal zu generieren. Die Phasenlagen der Spektralanteile, damit die Amplitudenschwankung und dadurch letztlich die Signalform wird immer unterschiedlich ausfallen.

Diese Schwierigkeiten können mit Low-noise noise umgangen werden, deshalb ist diese Signalform von großer Bedeutung für zukünftige Versuche und Untersuchungen. Es muss lediglich unterschieden werden, ob es sich beim generierten LNN-Signal um ein schmalbandiges Störgeräusch handelt, oder ob es der Definition breitbandig entspricht. Wird anstelle von schmalbandigen Zufallsrauschen ein schmalbandiges LNN-Signal als Maskierer verwendet, sinkt die Mithörschwelle um gut 4dB. Kein Unterschied ergibt sich jedoch, wenn anstelle von breitbandigem Rauschen mit zufälliger Phasenlage breitbandiges LNN verwendet wird. Beide Signalformen erzielen gleiche Effekte.

Werden diese Eigenheiten der Signalform Low-noise noise beachtet, ist es in Zukunft sehr sinnvoll auf diese Möglichkeit der Signalgenerierung zurückzugreifen und LNN als Störsignal zu verwenden.

## 9.2. Untersuchungen der Sprachverständlichkeit mit Logatomtests

Bei der weiteren Untersuchung des Einflusses tieffrequenter Störgeräusche auf die Sprachverständlichkeit wurden interessante Zusammenhänge aufgedeckt, die subjektive Annahmen und Empfindungen bestätigt und wissenschaftlich untermauert.

Es ist nicht nur entscheidend, welches Störgeräusch auf eine Testperson einwirkt, sondern auch, in welcher Messumgebung sich der Proband befindet. Besonders wichtig ist dabei der Verlauf der Nachhallzeit über der Frequenz. Die Effekte sind dabei sowohl objektiv (Logatomtest) als auch subjektiv (Beschreibungen der Probanden) nachweisbar.

Werden zunächst nur die Auswirkungen betrachtet, die durch unterschiedliche Störgeräusche auftreten, ergeben sich folgende Ergebnisse.

Eine Störung entwickelt einen umso größeren Einfluss auf die Verständlichkeit von Sprache, je mehr Frequenzkomponenten in dem für Sprache wichtigen Frequenzbereich (800Hz bis 2,5kHz) liegen. So ist mit einem Tiefpassrauschen bis  $f_g = 250\text{Hz}$  und einem Pegel, der knapp unter dem Sprachsignalpegel liegt, nahezu kein Unterschied zur Situation ohne Störung detektierbar. Wird jedoch ein gleichlautes Rauschen eingespielt, das über den gesamten Hörbereich Frequenzanteile enthält, sinkt die Verständlichkeit stark ab.

Die Größenordnung, um die die Verständlichkeit abnimmt, ist dabei jedoch stark abhängig von der räumlichen Gegebenheit, insbesondere vom Verlauf der Nachhallzeit. Ist die Nachhallzeit über den gesamten Frequenzbereich annähernd konstant (Breitband-Absorber), oder



nimmt sie zu hohen Frequenzen hin ab (herkömmliche Höhenabsorber), so beträgt die Reduktion der Verständlichkeit ca. 25% bis 30%. Dieser Wert wird bei allen untersuchten Personengruppen (normalhörende Personen im Alter zwischen 20 und 50 Jahre, Probanden, die älter sind als 50 Jahre, normalhörende Personen, deren Muttersprache nicht Deutsch ist und eine hörgeschädigte Person) bestätigt. Sinkt die Nachhallzeit hingegen zu tieferen Frequenzen an (Tiefenabsorber), halbiert sich diese Abnahme nahezu auf ca. 15% – ebenfalls bei allen Testgruppen. Der Verlust der Verständlichkeit reduziert sich damit erheblich.

Zusätzlich dazu steigt die persönliche subjektive Behaglichkeit in Räumen mit diesem Profil. Alle Versuchspersonen beschrieben ausnahmslos den Raum mit eingebauten Höhenabsorbern als unangenehm und stuften ihre Behaglichkeit während der Versuchsdauer als schlecht und mangelhaft ein. Auch wurde oft erklärt das Störsignal erscheine lauter als in vorher geprüften Umgebungen, obwohl die Pegel konstant gehalten wurden. Besonders erwähnenswert dabei ist, dass die Verständlichkeit der Logatome von den Probanden durchwegs schlechter gefühlt wurde, als sie tatsächlich ausgewertet wurde. Hingegen wurde die Versuchsanordnung mit Tiefenabsorbern durchwegs gut beschrieben. Die Behaglichkeit der teilnehmenden Probanden war deutlich höher als in allen anderen untersuchten Messumgebungen und die Verständlichkeit wurde subjektiv oft als sehr gut eingeschätzt. Der Raum mit gleichbleibender Nachhallzeit über der Frequenz kann überblicksmäßig als neutral bezeichnet werden. Keine der Versuchspersonen äußerte sich schlecht über die vorhandenen raumakustischen Bedingungen, allerdings bewertete auch nur ein Teilnehmer (insgesamt 15 Probanden) den Raum besser als den mit eingebauten Tiefenabsorbern. Die Verständlichkeit wurde gut eingeschätzt, die Behaglichkeit war ebenfalls gut, aber selten sehr gut.

Der Verlauf der Nachhallzeit hat also – bei konstantem Mittelwert der Nachhallzeit – Einfluss auf die Verständlichkeit von Sprachsignalen. Entsprechende Publikationen wurden bestätigt und oft angezweifelte Aussagen objektiv nachgewiesen: Räume, in denen v.a. tiefe Frequenzen bedämpft werden, weisen sehr gute akustische Verhältnisse auf: das unangenehme "Dröhnen" wird vermieden, und die Sprachverständlichkeit ist sehr gut [2]. Aber auch in Bezug auf Behaglichkeit, Wohlbefinden und subjektiver Einschätzung der Verständlichkeit wird durch die Reduktion tiefer Frequenzen eine wesentliche Voraussetzung für eine entspannte Atmosphäre und eine ungestörte Kommunikation geschaffen.

In Zukunft wird es deshalb nicht mehr genügen, nur die Nachhallzeiten der mittleren und hohen Frequenzen festzulegen [41, 42] und diese mit konventionellen Mitteln wie Akustik-

decken und –putzen, Teppichen oder Vorhängen zu realisieren [25]. Die akustische Gestaltung (auch kleiner Räume) vor allem bei tiefen Frequenzen [43], wird oberste Priorität erhalten.

Abschließend wurde gezeigt, dass mit objektiven Berechnungen raumakustischer Gütemaße meist eine gute Abschätzung der realen Situation erreicht werden kann. Dazu liefern die Schwerpunkzeit TS und das Deutlichkeitsmaß  $C_{50}$  in einem ersten Schritt Anhaltspunkte über die Hörsamkeit des Raumes.

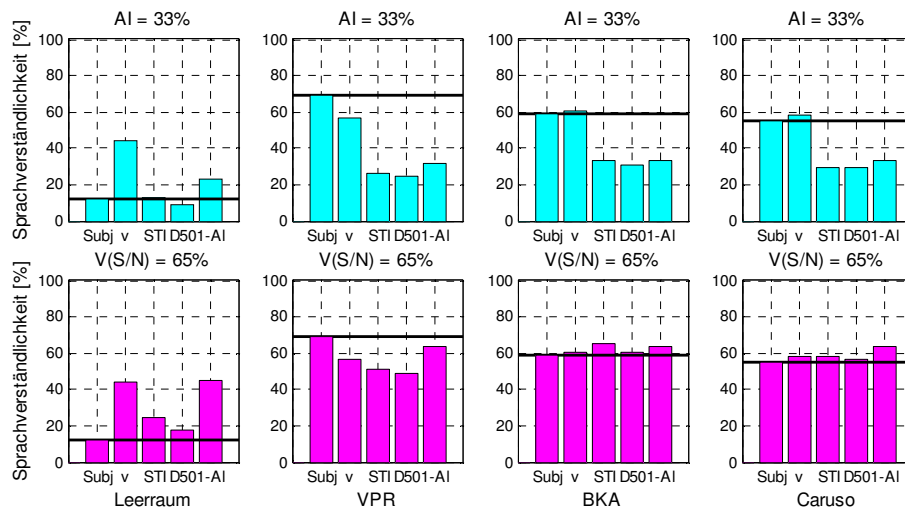
Bewerten diese beiden Größen die Verständlichkeit im gegebenen Raum als schlecht oder maximal als ausreichend, so sollte für weitere Berechnungen der exakten Größen (Prozentangaben) vor allem der Deutlichkeitsgrad  $D_{50}$  oder der STI verwendet werden. Die beiden anderen Abschätzungen sind für diese Raumsituation ungeeignet und liefern falsche Ergebnisse.

Wird die Hörsamkeit des Raums hingegen als gut oder sogar sehr gut eingeschätzt, so muss auch hier noch der Verlauf der Nachhallzeit beachtet werden. Für die Berechnungen ist es von Vorteil, wenn sie einen Verlauf aufweist, der durch herkömmliche Absorber geprägt wird: annähernd geradlinig über den gesamten Frequenzbereich bzw. zu hohen Frequenzen hin abnehmend. Mit diesen Gegebenheiten liefern vor allem die Objektive Verständlichkeit  $v$  und der Deutlichkeitsgrad in jedem der untersuchten Fälle gute Ergebnisse. Der STI pendelt etwas stärker um den Realwert, während der Artikulationsverlust durchgehend zu hohe Werte berechnet.

Einen Sonderfall stellt dabei ein zu tiefen Frequenzen hin fallender Nachhallzeitverlauf dar. Die Besonderheit, dass nur tiefe Frequenzen absorbiert werden, während die höheren Frequenzen ab 500Hz aufwärts unbedämpft bleiben, stellt die objektiven Berechnungen vor schwierige Aufgaben. Sie sind an die konventionellen und momentan weit verbreiteten Anwendungen angepasst, in denen der Raum entweder breitbandig bedämpft wird oder nur höhere Frequenzen absorbiert werden. Für steigende Nachhallzeiten (zu hohen Frequenzen hin) liefern jedoch alle Abschätzungen durchgängig zu geringe Werte. Soll trotzdem eine Prognose aufgestellt werden, ist es sinnvoll den Artikulationsverlust zu wählen, da dieser hier die tatsächliche Situation am besten beschreibt.

Bei allen diesen Auswertungen ist es zusätzlich zu den eben genannten Kriterien entscheidend, welcher Korrekturfaktor für vorhandene Störeinflüsse verwendet wird.

Der Artikulationsindex AI hat in diesem Fall für das tieffrequente Störgeräusch eine hervorragende Lösung berechnet, während das Ergebnis bei breitbandigem Rauschen weit unter den tatsächlichen Werten lag. Eine Verbesserung wurde durch den Silbenverständlichkeitsfaktor  $V_{S/N}$  erreicht (siehe Abb. 157).



**Abbildung 157: Vergleich der Korrekturfaktoren AI und  $V_{S/N}$  beim breitbandigen Störgeräusch Rosa Rauschen 20Hz bis 20kHz**

Hier hat sich herausgestellt, dass der Artikulationsindex AI vor allem für Störgeräusche sinnvoll ist, die sich im Frequenzbereich stark vom Nutzsignal unterscheiden. Zusätzlich dazu dürfen sich diese Signale im Pegel nicht bedeutend verändern. Für zum Nutzsignal ähnliche Signale ist hingegen der Silbenverständlichkeitsfaktor  $V_{S/N}$  zu verwenden.

Für die Zukunft ist es wünschenswert, ein Messverfahren zu entwickeln, das auch für diesen immer wichtiger werdenden Fall, exakte Abschätzungen liefert.

Die in dieser Arbeit gewonnenen Erkenntnisse bieten ideale Ansatzpunkte für weitere Forschungen zu diesem Thema. Diese sollten noch detaillierter auf die bisherigen, durch die begrenzte Zeit noch nicht abgeschlossenen Untersuchungen und deren Ergebnisse eingehen. Zusätzlich müssen die Auswirkungen unterschiedlicher zeitlicher Strukturen der Störgeräusche (z.B. Verhalten von Sprache als Störgeräusch) geklärt werden.

## Anhang

### Zahlenwerte der berechneten raumakustischen Gütemaße

Zusätzlich zu den vorne dargestellten Graphiken sollen hier die ermittelten Zahlenwerte der raumakustischen Gütemaße angegeben werden:

**Tabelle 5: Subjektiv ermittelte Silbenverständlichkeit  $v_s$  [%]**

	Leerraum	VPR	BKA	Caruso
Ruhe	37	82	89	87
RR 20-250	31	82	83	86
RR 20-20k	12	69	59	55

**Tabelle 6: Objektiv ermittelte Silbenverständlichkeit  $v$  [%]**

	Leerraum	VPR	BKA	Caruso
Ruhe	65	84	90	86
RR 20-250	63	82	88	84
RR 20-20k	44	57	61	58

**Tabelle 7: Schwerpunktzeit TS [s]**

	Leerraum	VPR	BKA	Caruso
Ruhe	184	47	22	37
RR 20-250	-	-	-	-
RR 20-20k	-	-	-	-

**Tabelle 8: Deutlichkeitsmass  $C_{50}$  [dB]**

	Leerraum	VPR	BKA	Caruso
Ruhe	-4,09	4,99	12,06	8,24
RR 20-250	-	-	-	-
RR 20-20k	-	-	-	-

**Tabelle 9: Artikulationsindex AI [%]**

	Leerraum	VPR	BKA	Caruso
Ruhe	100	-	-	-
RR 20-250	93,2	-	-	-
RR 20-20k	33,0	-	-	-

**Tabelle 10: Silbenverständlichkeitsfaktor  $V_{S/N}$  [%] aus Tabelle 49**

	Leerraum	VPR	BKA	Caruso
Ruhe	100	-	-	-
RR 20-250	-	-	-	-
RR 20-20k	65	-	-	-

**Tabelle 11: Deutlichkeitsgrad  $D_{50}$  [%], Werte bei Störsignal mit AI berechnet**

	Leerraum	VPR	BKA	Caruso
Ruhe	28	76	94	87
RR 20-250	26	71	88	81
RR 20-20k	9	25	31	29

**Tabelle 12: Deutlichkeitsgrad  $D_{50}$  [%], Werte 20-20k mit Wert aus Tabelle berechnet**

	Leerraum	VPR	BKA	Caruso
Ruhe	28	76	94	87
RR 20-250	26	71	88	81
RR 20-20k	<b>18</b>	<b>49</b>	<b>61</b>	<b>57</b>

**Tabelle 13: 1-Artikulationsverlust  $1-AI_{\text{cons}}$  [%],  
Werte bei Störsignal mit AI berechnet**

	Leerraum	VPR	BKA	Caruso
Ruhe	69	98	99	99
RR 20-250	64	91	92	92
RR 20-20k	23	32	33	33

**Tabelle 14: 1-Artikulationsverlust  $1-AI_{\text{cons}}$  [%],  
Werte 20-20k mit Wert aus Tabelle berechnet**

	Leerraum	VPR	BKA	Caruso
Ruhe	69	98	99	99
RR 20-250	64	91	92	92
RR 20-20k	<b>45</b>	<b>64</b>	<b>64</b>	<b>64</b>

**Tabelle 15: Speech Transmission Index STI [%],  
Werte bei Störsignal mit AI berechnet**

	Leerraum	VPR	BKA	Caruso
Ruhe	39	78	100	89
RR 20-250	36	73	93	83
RR 20-20k	13	26	33	29

**Tabelle 16: Speech Transmission Index STI [%],  
Werte 20-20k mit Wert aus Tabelle berechnet**

	Leerraum	VPR	BKA	Caruso
Ruhe	39	78	100	89
RR 20-250	36	73	93	83
RR 20-20k	<b>25</b>	<b>51</b>	<b>65</b>	<b>58</b>

**Tabelle 17: Rapid STI RASTI [%],  
Werte bei Störsignal mit AI berechnet**

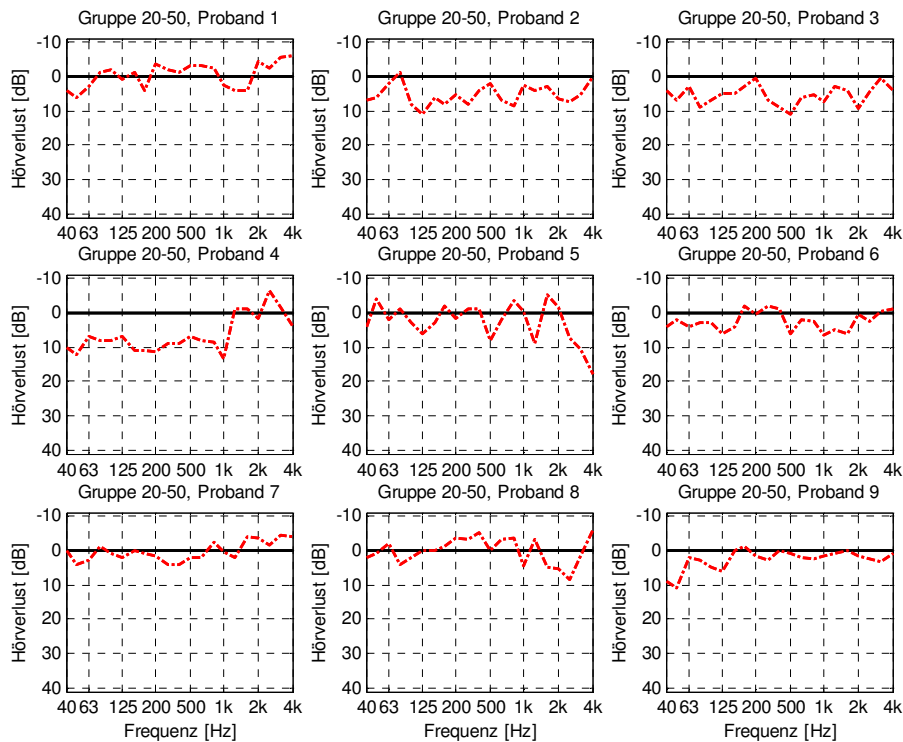
	Leerraum	VPR	BKA	Caruso
Ruhe	38	76	100	94
RR 20-250	35	71	93	88
RR 20-20k	13	25	33	31

**Tabelle 18: Rapid STI RASTI [%],  
Werte 20-20k mit Wert aus Tabelle berechnet**

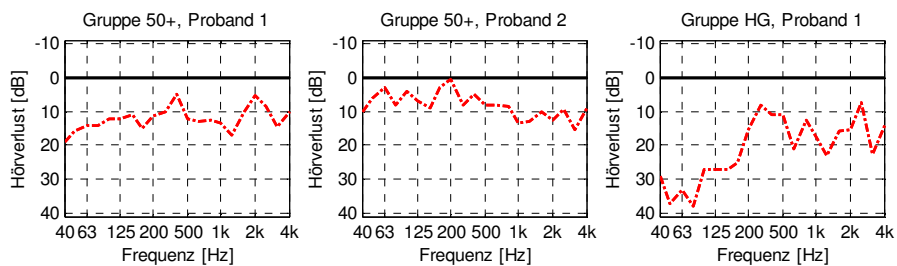
	Leerraum	VPR	BKA	Caruso
Ruhe	38	76	100	94
RR 20-250	35	71	93	88
RR 20-20k	<b>25</b>	<b>50</b>	<b>65</b>	<b>60</b>

## Hörschwellen der Testpersonen

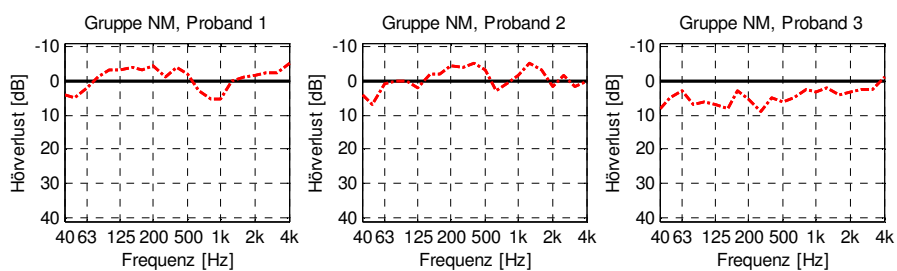
In den folgenden Abbildungen werden die Hörverluste der an den Versuchen teilnehmenden Testpersonen dargestellt.



**Abbildung 158: Hörschwellen, Gruppe 20- bis 50-Jährige**



**Abbildung 159: Hörschwellen, Gruppe Über-50-Jährige und Hörgeschädigte**



**Abbildung 160: Hörschwellen, Normalhörende Personen, deren Muttersprache nicht deutsch ist**

## Logatomlisten

Für die Ermittlung der Sprachverständlichkeit wurden folgende Logatome (12 Listen mit je 50 sinnlosen Einsilbern) verwendet:

**Liste 1/1 (Deutsch)**

lid	hecht	kroch	banz	spusch
lan	wint	maust	spres	sön
schles	schrolb	heir	blek	fetz
sen	falt	brüm	driir	ping
grirk	frorch	garn	wet	teus
gold	keg	dat	sef	glund
trül	nand	där	ruch	dib
zwet	wein	stird	weink	zirt
bar	degt	zend	neft	klils
jul	gautz	duns	mem	schen

**Liste 1/2 (Deutsch)**

hes	treunt	tot	gurch	gren
bitz	detz	lir	kar	lern
dert	jöng	dol	glir	den
rils	sin	hen	sis	spaund
schogt	schrör	mend	brüt	becht
fim	weft	kräm	zil	nauk
blest	wesch	wab	kluch	geinz
drach	fen	wunk	sand	schlard
zirk	nuf	meg	spralt	freit
stald	zwed	güns	peis	dulb

**Liste 2/1 (Deutsch)**

mer	heing	got	län	ker
nark	fut	hogt	kril	drauls
fard	zecht	silk	jas	wes
dauch	schink	war	spir	parch
seit	girn	tint	gresch	stenz
nust	dalt	zwem	bech	send
mun	glelb	deuf	zus	weft
wol	schlein	detz	sprig	rod
schrützlert	klöns	bran	gem	
bend	blid	frün	duld	trib

**Liste 2/2 (Deutsch)**

schicht	dügt	gisch	sard	nul
sar	hes	bend	tat	zeich
ren	wes	list	wom	schlern
schrufft	glilt	meld	gruk	wetz
keim	derk	frar	bläng	nen
trub	zwulb	gin	der	drenk
seis	zaul	wof	dich	pals
störch	breg	leut	fetz	sprond
banz	fins	hort	mint	spin
jat	ger	düd	klend	kraun

**Liste 3/1 (Deutsch)**

wur	senk	heint	fror	tün
spugt	wis	kenz	wind	bran
sit	lächt	stib	goch	selt
furn	jens	schreil	hed	ferk
ret	gald	ditz	treft	deit
lirt	pech	drem	sprar	band
zerd	böls	daulb	dast	blak
mun	schlün	schor	nerch	krind
nus	zweung		ditz	klauf
grin	mal	gesch	zem	glog

**Liste 3/2 (Deutsch)**

zwest	drok	trech	düst	sprelt
grend	schleg	pun	tuf	sort
wing	bech	freur	dift	lin
krel	degt	neit	want	was
hil	gleirn	gerd	mit	klesch
fitz	schos	jaur	zerch	gand
dim	naud	rur	ser	feicht
bink	stet	bröns	huns	kam
blatz	lünd	gols	zuld	manz
schran	delb	sirk	speb	wän

**Liste 4/1 (Deutsch)**

kraud	schran	wis	gres	spen
dift	schlon	degt	drinz	rif
sarn	ler	geuch	gend	wer
det	wön	nult	fot	batz
blent	hink	sching	pesch	süld
but	stir	trün	gicht	sek
meim	klerk	sprutz	zerch	zar
lul	mab	jeil	kilb	wens
zweich	tond	hauls	das	fort
bräst	glend	nurd	dam	freg

**Liste 4/2 (Deutsch)**

zugt	geicht	wült	klar	puff
bris	tralb	lerch	scheud	mart
glis	schron	nunk	daut	gerk
dam	süls	grein	fron	mich
wet	zwig	spund	wir	bang
hird	sent	schlen	fens	blen
henz	jek	stutz	resch	sprot
ker	nest	tarch	leld	draub
def	fern	krönd	wis	dir
seitz	gand	zäl	dim	bol

**Liste 5/1 (Deutsch)**

dun	beun	schrint	glerd	spols
zat	rült	frir	must	blen
him	kert	zward	seb	lof
serk	heg	fes	schles	sprän
wigt	braul	det	tem	wosch
greil	hinz	detz	lik	der
meift	klüt	krad	fuch	wis
zocht	schir	beld	drer	stand
datz	palb	treind	guns	jun
nech	wirn	gaung	sönk	garch

**Liste 5/2 (Deutsch)**

zend	rinz	sird	schrog	wer
frel	zwauns	klim	dad	drun
grat	meng	desch	satz	heils
nink	trus	spend	pift	zirn
brak	lert	tigt	düs	bleil
bas	jir	spren	gär	lont
steult	schlerk	schurch	seind	kör
hon	waul	duf	glen	ducht
gen	nütz	bet	gib	wach
fet	kram	west	milb	fech

**Liste 6/1 (Deutsch)**

lid	glisch	jes	lend	zin
wort	sans	wost	spetz	burch
klels	munk	wenz	steigt	nar
düch	spret	focht	frend	traun
dil	duk	galb	ber	galt
scher	bris	pat	sör	blaul
tren	zeb	nas	gif	hen
schret	kild	durd	kruch	süg
deim	droft	mirn	zwark	grem
hätz	rint	feind	weng	schleun

**Liste 6/2 (Deutsch)**

deing	sab	hilb	bend	drüm
nind	len	brift	faus	sprind
gar	mos	zald	wof	zwök
wet	spens	den	kralt	her
blust	ser	stärk	metz	leich
glin	gud	det	bigt	peutz
grenk	gunz	ses	wel	schlucht
norch	trich	resch	kaunt	früm
scherd	klag	schran	wils	fur
deit	jort	tel	din	zarn



---

## Literaturverzeichnis

- [1] **Fuchs, H. V.**, Zur Bedeutung tiefer Frequenzen für die Sprachverständlichkeit, Fraunhofer-Institut für Bauphysik (IBP) Stuttgart, 2003
- [2] **Fuchs, H. V.**, Bessere Kommunikation durch "transparente" Raum-Akustik, Gesundheitsingenieur 120 (1999), Heft 4, S. 159-168
- [3] **DIN ISO 8253-2**, Audiometrische Prüfverfahren, Teil 2: Schallfeld-Audiometrie mit reinen Tönen und schmalbandigen Prüfsignalen, 1994
- [4] **Pflüger M.**, Psychoakustik, Skriptum der Österreichischen Hochschülerschaft TU Graz, 2000
- [5] **Hönig H.**, Akustik der Musikinstrumente, Skriptum der Österreichischen Hochschülerschaft KUG Graz, 1994
- [6] **Birbaumer, N., Schmidt, R.**, Biologische Psychologie (2. Auflage), Springer Verlag, 1991, S. 23-51
- [7] **Dickreiter M.**, Handbuch der Tonstudioteknik, Band I (6. verbesserte Auflage), K. G. Saur Verlag München, 1997, S. 1-11, 25-38, 61-68, 107-117
- [8] **Rieländer, M.**, Reallexikon der Akustik, Verlag Erwin Bochinsky Frankfurt, 1982, S. 19, 62, 104, 138, 203, 204, 292, 356, 364-367
- [9] **Blutner, F., Weißing H.**, Schallwahrnehmung, aus Fasold, W., Kraak, W., Schirmer, W., Taschenbuch Akustik, Teil 1, VEB Verlag Technik Berlin, 1984, S. 203-227, 235-244 und 249-251
- [10] **Bundesministerium für Umwelt, Jugend und Familie** unter Mitwirkung des Österreichischen Arbeitsringes für Lärmbekämpfung, Psychoakustik - Gehörbezogene Lärmbewertung (2. unveränderte Auflage), Signale Ges.m.b.H. Pinkafeld, 1994, S. 61-78
- [11] **Zwicker, E.**, Psychoakustik, Springer-Verlag, 1982, S. 15-53, 93-101
- [12] **Zwicker, E., Feldtkeller, R.**, Das Ohr als Nachrichtenempfänger (2. Auflage), S.Hirzel Verlag Stuttgart, 1967, S. 1-26, 31-78, 159-167, 179-184, 204-214
- [13] **Pumplin, J.**, Low-noise noise, Journal of the Acoustical Society of America, Vol. 78, No. 1, 1985
- [14] **Hartmann, W., Pumplin, J.**, Periodic signals with minimal power fluctuations, Journal of the Acoustical Society of America, Vol. 90, No. 4, 1991, S. 1986-1999
- [15] **Hartmann, W., Pumplin, J.**, Expanded tables for Periodic signals with minimal power fluctuations (MPFET), Journal of the Acoustical Society of America, Vol. 90, No. 15, 1986
- [16] **Hartmann, W., Pumplin, J.**, Noise power fluctuations and the masking of sine signals, Journal of the Acoustical Society of America, Vol. 83, No. 6, 1988, S. 2277-2289
- [17] **Lazarus, H., Lazarus-Mainka, G., Schubeius, M.**, Sprachliche Kommunikation unter Lärm, Friedrich Kiehl Verlag GmbH Ludwigshafen, 1985, S. 41-86, 96-146, 209-230, 299-310

- 
- [18] **Kollmeier, B.**, Moderne Verfahren der Sprachaudiometrie, Band 1 der Buchreihe "Audiologische Akustik", median-verlag von Killisch-Horn GmbH, 1992, S. 137-169
- [19] **Slawin, I.**, Industrielärm und seine Bekämpfung, VEB Verlag Technikum Berlin, 1960, S. 28-54
- [20] **Hoffmeier, J.**, Untersuchungen zum Einfluss von Raumklangfärbungen auf die Deutlichkeit von Sprache, Diplomarbeit, Technische Universität Dresden, Fakultät Elektrotechnik, 1996
- [21] **Schmidt W., Weichardt, W.**, Raumakustische Kriterien für Sprache und Musik, aus Fasold, W., Kraak, W., Schirmer, W., Taschenbuch Akustik, Teil 2, VEB Verlag Technik Berlin, 1984, S. 1188-1203
- [22] **Graber G., Weselak W.**, Skriptum Raumakustik (Version 1.0), Institut für Nachrichtentechnik und Wellenausbreitung TU Graz, 2002
- [23] **Zollner, M., Zwicker, E.**, Elektroakustik (3. Auflage), Springer-Verlag, 1993
- [24] **Kuttruff, H.**, Raumakustik, aus Heckl, M., Müller, H., Taschenbuch der Technischen Akustik (2. Auflage, Korrigierter Nachdruck), Springer Verlag, 1995, S. 596-624
- [25] **Fuchs, H. V.**, Gute Akustik in Büros, Teil 1 und 2, Trockenbau AKUSTIK, Heft 05/03 und 06/03, 2003, S. 34-38 und 26-29
- [26] **Zha, X., Fuchs, H. V., Eckold, D., Babuke, G.**, Alternative reflexionsarme Raumauskleidung, Zeitschrift für Lärmbekämpfung 45, 1998, S. 32-35
- [27] **DIN ISO 4869-1**, Gehörschützer, Subjektive Methode zur Messung der Schalldämmung, 1991
- [28] **musikelectronic geithain**, Datenblatt ME100, High-End Koaxial-Lautsprecher, 2003
- [29] **musikelectronic geithain**, Datenblatt ME100, High-End Koaxial-Lautsprecher, 2003
- [30] **Kammeyer, K., Kroschel, K.**, Digitale Signalverarbeitung, Filterung und Spektralanalyse, 3. Auflage, B.G. Teubner Verlag Stuttgart, 1996, S. 77-92
- [31] **Zha, X., Späh, M., Nocke, C., Jiang, G.**, Neuer Laborraum für Untersuchungen an Schallabsorbern bei tiefen Frequenzen, Fraunhofer-Institut für Bauphysik (IBP), Mitteilung 317, 1997
- [32] **FhG – IBP**, Messung der Schallabsorption von Verbund-Platten-Resonator (VPR) im Hallraum nach DIN EN 20354, 2003
- [33] **FhG – IBP**, Messung der Schallabsorption von Breitband-Kompakt-Absorber (BKA) im Hallraum nach DIN EN 20354, 2002
- [34] **FhG – IBP**, Messung der Schallabsorption von CARUSO-ISO-BOND (CIB) im Impedanzrohr nach ISO 10534-2, 2003
- [35] **DIN 45 630, Blatt 2**, Grundlagen der Schallmessung, Kurven gleicher Lautstärkepegel, 1967
- [36] **Zha, X., Fuchs, H. V., Späh, M.**, Ein neues Konzept für akustische Freifeldräume, Sonderdruck aus den rundfunktechnischen Mitteilungen, Heft 3, 42. Jahrgang, 1998, S. 81-91
- [37] **FhG – IBP**, Technische Daten des akustischen Freifeldraums

- [38] **Fuchs, H. V., Zha, X., Zhou, X., Castor. F.**, Raumakustik mit System, Glas-Verarbeitung 3, 2001, S. 59-64
- [39] **Fuchs, H. V.**, Planer, hört die Signale, Mensch & Büro International, Heft 13, 1999, S. 56-57
- [40] **HEAD acoustics GmbH**, Auswertungen zu den raumakustischen Untersuchungen, Messprotokoll, Fraunhofer-Institut für Bauphysik (IBP), 2002
- [41] **DIN 18 041**, Hörsamkeit in kleinen bis mittelgroßen Räumen, 1968
- [42] **Tennhardt, H.-P.**, Die neue DIN 18041 – Anwendungen, Veröffentlichungen der Hochschule für Technik Stuttgart, Band 64, 2003, S. 21-32
- [43] **Fuchs, H. V., Späh, M., Pommerer, M., Schneider, W., Roller, M.**, Akustische Gestaltung kleiner Räume bei tiefen Frequenzen, Bauphysik, Sonderdruck aus: 20 (1998), Heft 6, S. 181-190
- [44] **Zickmantel, F.**, Reflexionsarme Kabine für audiometrische Messungen, Diplomarbeit, Hochschule Mittweida (FH), Fachbereich Medien & Elektrotechnik, 2000