

Technische Universität Berlin
Studiengang Kommunikationswissenschaft

Perzeptive Bewertung
von kammfilterartigen linearen Verzerrungen in
der Audiotechnik

Magisterarbeit

Vorgelegt von:

Stefan Brunner

Immatrikulationsnummer 190634

Betreuender Erstgutachter: Prof. Dr. Stefan Weinzierl

Zweitgutachterin: Prof. Dr. Brigitte Schulte – Fortkamp

Berlin, den 19. Juni 2006

Die selbständige Anfertigung dieser Arbeit versichere ich an Eides statt.

Berlin, den

Stefan Brunner

Inhaltsverzeichnis

Abbildungsverzeichnis	VI
Tabellenverzeichnis	VII
1 Einleitung	1
1.1 Vorwort	1
1.2 Ziel der Arbeit	2
1.3 Struktur der Arbeit.....	3
2 Grundlagen.....	4
2.1 <i>Entstehung und spektrale Merkmale kammfilterartiger linearer Verzerrungen.....</i>	<i>4</i>
2.1.1 Konstruktive und destruktive Interferenz	5
2.1.2 Zusammenhang von Interferenz und <i>gain</i> des überlagernden Signals	6
2.2 <i>Kammfiltereffekte in der Audiotechnik</i>	<i>7</i>
2.2.1 Laufzeitstereofonie mit mehr als zwei Mikrofonen.....	8
2.2.2 Haupt- und Stützmikrofonie	8
2.2.3 Reflexion an schallharten Begrenzungsflächen.....	9
2.2.4 Mono-Kompatibilität.....	10
2.3 <i>Hörbarkeit von Kammfiltereffekten</i>	<i>10</i>
2.4 <i>Klangverfärbungen durch kammfilterartige Verzerrungen</i>	<i>14</i>
2.4.1 Kammfiltereffekt bei Rauschsignalen	15
2.4.2 Kammfiltereffekt bei Sprachsignalen	16
2.4.3 Kammfiltereffekt bei Musik	17
2.4.4 Klangverfärbungen bei veränderlichem ΔT	17
2.4.5 Erwünschte Klangverfärbungen – Flanging und Phasing	17
3 Methodik.....	19
3.1 <i>Überblick über psychophysische Messmethoden.....</i>	<i>19</i>

3.2	<i>Adaptive Testverfahren</i>	21
3.2.1	PEST-Verfahren	22
3.2.2	Maximum Likelihood Verfahren	22
3.2.3	<i>Staircase</i> Verfahren.....	22
4	Untersuchung	27
4.1	<i>Quellsignale</i>	27
4.1.1	Auswahl und Vorbereitung	27
4.1.2	Spektrale Struktur.....	28
4.2	<i>Generierung der Stimuli</i>	30
4.3	<i>Anpassung der Lautheiten</i>	31
4.4	<i>Design des Experimentes</i>	32
4.4.1	Grundlegendes Design.....	32
4.4.2	Weiterführendes Design	33
4.4.3	Software	35
4.5	<i>Vorversuche</i>	35
4.5.1	Erster Vorversuch.....	36
4.5.2	Zweiter Vorversuch.....	39
4.6	<i>Hauptversuch</i>	41
4.6.1	Stimuli.....	41
4.6.2	Versuchspersonen.....	42
4.6.3	Versuchsaufbau.....	42
4.6.4	Versuchsablauf.....	42
4.6.5	Ergebnisse	43
5	Auswertung	51
5.1	<i>Interpretation der Ergebnisse</i>	51
5.1.1	Piano	51
5.1.2	Snare.....	52
5.1.3	Sprache	52
5.2	<i>Diskussion</i>	53
5.3	<i>Kritik</i>	54

6	Zusammenfassung	57
7	Literatur	59
	Anhang A – Versuchsergebnisse	63
	Anhang B – Instruktionen für den Hauptversuch	66
	Anhang C – Fragebogen	67
	Anhang D – Inhalt der CD und Bedienungsanleitung	68
	Anhang E – Quellcodes	70

Abbildungsverzeichnis

Abbildung 1: Blockschaltbild für ein delay- and- add Filter.	4
Abbildung 2: Darstellung einer typischen Kammfilterkurve.	4
Abbildung 3: Drei Beispiele für den Kammfiltereffekt; links: $\Delta T= 0.1$ ms; Mitte: $\Delta T=0.5$ ms; rechts: $\Delta T= 1$ ms.	6
Abbildung 4: Zusammenhang zwischen der Pegeldifferenz der beteiligten Signale und Pegelanhebung an den <i>peak</i> -Frequenzen des Ausgangssignals.	7
Abbildung 5: Blockschaltbild des Routings bei einer ABC- Mikrofonanordnung (nach Rode (2006), S. 10).	8
Abbildung 6: Reflexion an schallharter Begrenzungsfläche; links: ungünstige Mikrofonposition; rechts: Abhilfe durch Absenkung des Mikrofons (aus: Dickreiter, 2003, S.189).	9
Abbildung 7: Typische psychometrische Funktion (nach: Leek (2001), S. 1280).	20
Abbildung 8: Simple Up- Down Staircase (aus: Leek (2001), S.1285).	24
Abbildung 9: 3-Up-1-Down Staircase (aus: Leek 2001, S.1285).	26
Abbildung 10: Leistungsdichtespektrum des unverzerrten Pianosignals.	28
Abbildung 11: Leistungsdichtespektrum des unverzerrten Snaresignals.	29
Abbildung 12: Leistungsdichtespektrum des unverzerrten Sprachsignals.	29
Abbildung 13: Schema des Ablaufs der MATLAB- Routine. Die Bezeichnungen der Blöcke entsprechen den Funktions- und Scriptnamen.	35
Abbildung 14: Screenshot während einer Versuchsreihe aus dem Hauptversuch. Gerade wird das erste Sample abgespielt.	37
Abbildung 15: Ergebnisse des ersten Vorversuchs. Dargestellt sind die <i>peak</i> – Überhöhungen an der für die jeweilige Verzögerungszeit ermittelten Hörschwelle.	38
Abbildung 16: Ergebnisse des zweiten Vorversuches; links: Die Hörschwelle bei 0.8 ms Verzögerung; rechts: Die Hörschwelle bei 1.2 ms Verzögerung.	41
Abbildung 17: Diagramm der ermittelten Hörschwelle für die Kategorie Piano. Die Standardabweichung ist mittels der Dreiecke dargestellt.	44
Abbildung 18: Boxplot der Werte an der Hörschwelle für die Kategorie Piano.	45
Abbildung 19: Diagramm der Pegeldifferenzen an der ermittelten Hörschwelle für die Kategorie Piano.	46
Abbildung 20: Diagramm der ermittelten Hörschwelle für die Kategorie Snare. Die Standardabweichung ist mittels der Dreiecke dargestellt.	47
Abbildung 21: Boxplot der Werte an der Hörschwelle für die Kategorie Snare.	47
Abbildung 22: Diagramm der Pegeldifferenzen an der ermittelten Hörschwelle für die Kategorie Snare.	48
Abbildung 23: Diagramm der ermittelten Hörschwelle für die Kategorie Sprache. Die Standardabweichung ist mittels der Dreiecke dargestellt.	49
Abbildung 24: Boxplot der Werte an der Hörschwelle für die Kategorie Sprache.	49
Abbildung 25: Diagramm der Pegeldifferenzen an der ermittelten Hörschwelle für die Kategorie Sprache.	50
Abbildung 26: Pegeldifferenzen der überlagernden Signale an der ermittelten Hörschwelle für alle drei Kategorien (blau: Piano, orange: Sprache, grün: Snare).	50

Tabellenverzeichnis

Tabelle 1: Beschreibungen des Kammfiltereffektes.	14
Tabelle 3: Zusammenhang zwischen Verzögerungszeiten und Schallwegdifferenz der überlagerten Signale.	30
Tabelle 4: Ermittelte Hörschwelle in der Kategorie Piano.....	44
Tabelle 5: Pegeldifferenzen an der ermittelten Hörschwelle für die Kategorie Piano.	46
Tabelle 6: Ermittelte Hörschwelle in der Kategorie Snare.	46
Tabelle 7: Pegeldifferenzen an der ermittelten Hörschwelle für die Kategorie Snare.	48
Tabelle 8: Ermittelte Hörschwelle in der Kategorie Sprache.	48
Tabelle 9: Pegeldifferenzen an der ermittelten Hörschwelle für die Kategorie Sprache.	50
Tabelle 10: Erster Vorversuch. Überhöhungen an den <i>peak</i> - Frequenzen an der ermittelten Hörschwelle [dB].	63
Tabelle 11: Zweiter Vorversuch. Überhöhungen an den <i>peak</i> - Frequenzen an der ermittelten Hörschwelle [dB].	63
Tabelle 12: Ergebnisse des Hauptversuchs in der Kategorie Piano. Angegeben sind die <i>peak</i> - Überhöhungen [dB] an der Hörschwelle.	64
Tabelle 13: Ergebnisse des Hauptversuchs in der Kategorie Snare. Angegeben sind die <i>peak</i> - Überhöhungen [dB] an der Hörschwelle.	64
Tabelle 14: Ergebnisse des Hauptversuchs in der Kategorie Sprache. Angegeben sind die <i>peak</i> - Überhöhungen [dB] an der Hörschwelle.	65

1 Einleitung

1.1 Vorwort

On ne savait pas d'où venait ce son ou en découvrait des causes peu vraisemblables ce qui me donna envie d'en chercher une meilleure. Je trouvai bientôt qu'il procédait de la réflexion du bruit de la fontaine contre les pierres du degré¹.

Christiaan Huygens²

Als der holländische Physiker und Astronom Christiaan Huygens im Jahre 1693 im Garten des Schlosses von Chantilly de la Cour in Frankreich spazieren ging, vernahm er auf einmal das Rauschen eines hinter ihm befindlichen Springbrunnens, jedoch mit einer merkwürdigen Klangfarbe. Ein deutlicher tonaler Anteil war erkennbar, wenn Huygens die von einer steinernen Treppe vor ihm herrührenden Reflexionen zugleich mit dem Direktschall des Brunnens hörte.³ Diese Entdeckung gilt als einer der ersten Belege für die bewusste Wahrnehmung einer kammfilterartigen Verzerrung.

Während dieses Phänomen zu jener Zeit wohl eher als Kuriosität Aufmerksamkeit erregte, werden Klangverfärbungen, welche durch kammfilterartige Verzerrungen entstehen, heute als schwerwiegendes Defizit einer Tonaufnahme gewertet. Zur Vermeidung dieser störenden Verfärbungen haben sich unter Tonmeistern und Toningenieurern gewisse Faustregeln, unter denen die 3:1-Regel nach Burroughs wohl die prominenteste ist, etabliert. Ungeachtet der Tatsache, dass diese Faustregeln in der Praxis häufig tatsächlich funktionieren, ist die allgemeine Anwendbarkeit derselben ohne wissenschaftlich fundierten Hintergrund in Frage zu stellen. Unter den wissenschaftlichen Arbeiten, welche sich in der

¹ „Erst war nicht festzustellen, woher dieser Ton kam; dann aber machte die gefundene, wenig plausible Ursache Lust, eine bessere zu finden. Schlussendlich fand ich heraus, dass der Ton ein Produkt der Reflexionen des Geräusches des Springbrunnens an den Treppenstufen war.“

² Huygens (1696), S. 570, zitiert nach: Bilsen & Ritsma (1969), S. 64.

³ Vgl. Huygens (1696), S. 570, zitiert nach: Bilsen & Ritsma (1969), S. 64f.

Vergangenheit mit der Wahrnehmbarkeit von Kammfiltereffekten beschäftigten, findet sich jedoch keine durchgehend dokumentierte wissenschaftliche Untersuchung, welche die Ermittlung einer Hörschwelle für kammfilterartige Verzerrungen zum Gegenstand hat. So ist eine Einschätzung der Wahrnehmbarkeit dieser Art von Verzerrungen und ein darauf aufbauendes Konzept zur Vermeidung störender Verfärbungen mithin ein Konstrukt aus tradierten oder eigenen Erfahrungen, welches lediglich unter weitgehend invarianten Bedingungen funktioniert.

1.2 Ziel der Arbeit

Ein exakter Hörschwellenverlauf für kammfilterartige Verzerrungen ist ein Werkzeug, welches den in der Praxis der Tonaufnahme kursierenden Faustregeln einen theoretischen Unterbau verleihen bzw. dazu anregen könnte, diese Regeln kritisch zu hinterfragen.

Ziel dieser Arbeit ist somit die Ermittlung einer Hörschwelle für kammfilterartige lineare Verzerrungen durch die Herstellung eines Zusammenhangs zwischen den wichtigsten bei der Superposition von Schallwellen Einfluss nehmenden physikalischen Variablen und der Wahrnehmbarkeit einer durch diese Superposition entstehenden klanglichen Veränderung. Diese Aufgabe soll durch eine Reihe von psychoakustischen Experimenten erfüllt werden. Zur Durchführung der Hörversuche wird eine Software erstellt, welche den automatischen und damit störungsfreien Ablauf der Experimente garantieren soll.

Da es sich beim Untersuchungsgegenstand um eine spektrale Manipulation handelt, ist anzunehmen, dass die spektrale Gestalt der Quellsignale einen wesentlichen Einfluss auf die Versuchsergebnisse hat. Aus diesem Grund soll die Auswahl der Quellsignale eine möglichst gute Approximation der tatsächlich in der Praxis zu erwartenden Signale sein.

Diese Art von Verzerrung soll von einer Personengruppe gehört und vermieden werden, deren Vertreter auf das Erkennen klanglicher Veränderungen spezialisiert sind. Deshalb ist es ratsam, auch die Probanden für den durchzuführenden Versuch aus dieser Personengruppe zu rekrutieren. Damit wird zwar die Möglichkeit zur Verallgemeinerung der

Versuchsergebnisse beschnitten, jedoch die Anwendbarkeit der zu erwartenden Ergebnisse gewährleistet.

1.3 Struktur der Arbeit

Um den Leser mit der behandelten Problematik vertraut zu machen, wird im Grundlagenteil zunächst die Theorie des Zustandekommens einer kammfilterartigen linearen Verzerrung erläutert. Im Folgenden wird das Problem des Auftretens von Kammfiltereffekten in der Audiotechnik anhand von Beispielen verdeutlicht, sowie eine Bestandsaufnahme der bisher geleisteten Forschung zur Frage der Wahrnehmbarkeit kammfilterartiger Verzerrungen versucht. Des Weiteren wird auf die für die vorliegende Arbeit wesentlichsten Aspekte der Psychometrik eingegangen und im Zusammenhang damit die Auswahl der anzuwendenden Verfahren motiviert.

Im Untersuchungsteil werden alle Phasen der Vorversuche und des Hauptversuchs detailliert dargestellt. Dabei wird auch auf die Gestaltung der verwendeten Software eingegangen. Es folgt die tabellarische sowie grafische Präsentation der Versuchsergebnisse.

Wichtigster Bestandteil des Kapitels „Auswertung“ ist die interpretierende Auseinandersetzung mit den Versuchsergebnissen. Weiterhin wird hier die gesamte Arbeit einer kritischen Beurteilung unterzogen.

2 Grundlagen

2.1 Entstehung und spektrale Merkmale kammfilterartiger linearer Verzerrungen

Die Manipulation eines Signals, die sich in der Änderung der spektralen Gestalt durch zeitliche und/oder Amplitudenvariation von bereits im Originalsignal enthaltenen Frequenzkomponenten bemerkbar macht, wird als *lineare Verzerrung* bezeichnet. Im Gegensatz dazu nennt man Verzerrungen eines Signals, bei denen neue Frequenzkomponenten entstehen, *nicht-lineare Verzerrungen*.⁴

Werden zwei spektral identische Signale mit einer zeitlichen Verzögerung Δt summiert (*delay-and-add*), resultiert daraus ein neues Signal, welches durch Anhebungen und Einbrüche im Spektrum, die an die Zähne eines Kammes erinnern, gekennzeichnet ist.⁵

(Abbildungen 1 & 2)

Man spricht hier folgerichtig vom Zustandekommen eines Kammfiltereffektes. Die erstmalige Beschreibung dieses Phänomens wird dem holländischen Physiker und Astronomen Christiaan Huygens (1629 – 1695) zugeschrieben.⁶

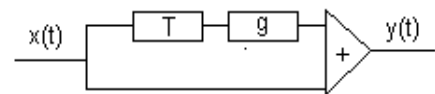


Abbildung 1: Blockschaltbild für ein delay-and-add Filter.

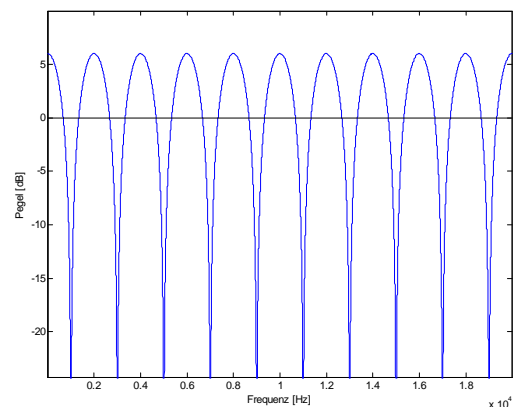


Abbildung 2: Darstellung einer typischen Kammfilterkurve.

⁴ Vgl. Preis (1976), S. 346.

⁵ Am deutlichsten ist die charakteristische Form bei gleichen Amplituden beider Signale und einem kontinuierlichen Spektrum der Quellen (idealer Weise Rauschsignale) erkennbar.

⁶ Vgl. Ritter (2005), S. 4.

2.1.1 Konstruktive und destruktive Interferenz

Wird eine Schallwelle einer anderen überlagert, kann es zu Interferenzen kommen. Je nach Phasenlage der sich überlagernden Schallwellen treten bei der Superposition konstruktive oder destruktive Interferenzen auf. Nach dem Blockschaltbild in Abbildung 1 lässt sich das Ausgangssignal dieser Superposition wie folgt berechnen:

$$y(t) = x(t) + gx(t - T) \tag{1}$$

g = gain des verzögerten Signals
T = Verzögerungszeit

Der Extremfall der vollständigen destruktiven Interferenz tritt dann auf, wenn gegenphasige Frequenzanteile (d. h. Anteile mit einer Phasenlage von 180°) der überlagerten Signale bei der Superposition summiert werden. Im mathematischen Idealfall kommt es hier zur vollständigen Auslöschung des betreffenden spektralen Anteils. Diese *notch*⁷-Frequenzen treten an den ungeradzahigen Vielfachen von $1/2T$ auf. Der entgegengesetzte Fall ist der der konstruktiven Interferenz. Hier werden gleichphasige Frequenzanteile (dies entspricht einer Phasenlage von 0°) der überlagerten Signale dergestalt addiert, dass sich im Falle der ungedämpften Superposition eine Pegelanhebung von 6 dB bei Frequenzen, die den ganzzahligen Vielfachen von $1/T$ entsprechen⁸, ergibt.⁹ Die *peaks* liegen also in einem harmonischen, die *notches* in einem disharmonischen Verhältnis zueinander. Der nominelle Abstand zwischen zwei aufeinanderfolgenden *peaks* und *notches* beträgt $1/T$. Bei der Überlagerung zweier spektral identischer Signale ergibt sich nach Hartmann¹⁰ der Betrag der spektralen Amplitudenübertragungsfunktion $|H(f)|$ zu

⁷ Zur Erhöhung der Lesbarkeit des Textes werden englische Fachbegriffe, die Eingang in die wissenschaftliche Sprache gefunden haben, nicht übersetzt, jedoch zur besseren Abgrenzung vom übrigen Text *kursiv* geschrieben.

⁸ Diese Frequenzen werden im Folgenden als *peak*-Frequenzen bezeichnet.

⁹ Vgl. Sengpiel (2000), S. 1.

¹⁰ Vgl. Hartmann (1998), S. 363.

$$|H(f)| = \sqrt{1 + g^2 + 2g \cos 2\pi fT} \quad (2)$$

sowie der Phasengang zu

$$\varphi(f) = \arctan \frac{-g \sin 2\pi fT}{1 + g \cos 2\pi fT} \quad (3)$$

Das Ergebnis der Superposition ist in Abbildung 3 anhand von drei Beispielen dargestellt.

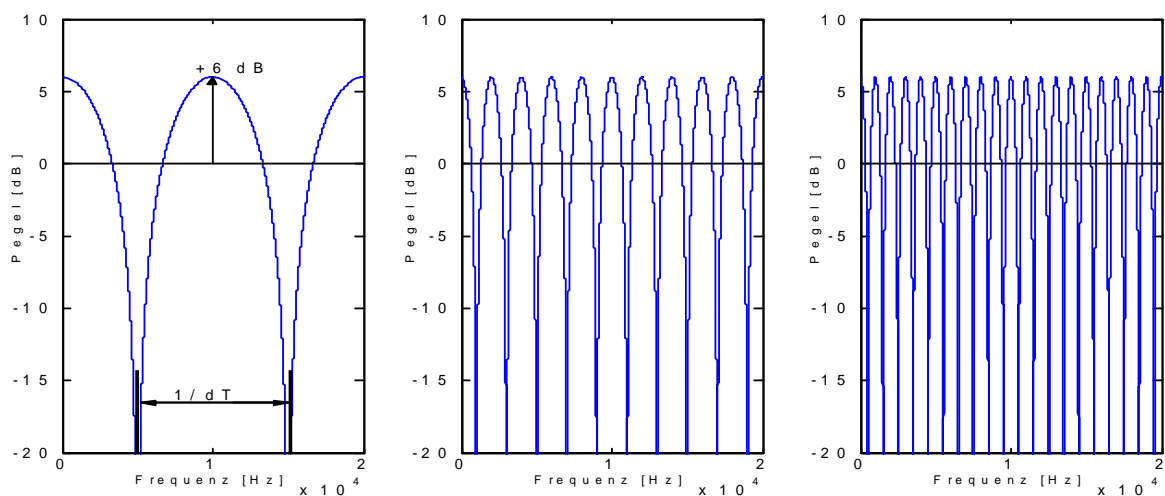


Abbildung 3: Drei Beispiele für den Kammfiltereffekt; links: $T = 0.1$ ms; Mitte: $T = 0.5$ ms; rechts: $T = 1$ ms.

2.1.2 Zusammenhang von Interferenz und *gain* des überlagernden Signals

Der Grad der Auslöschungen und Anhebungen im Spektrum des durch Superposition entstandenen Signals ist stark vom Pegelverhältnis der beiden summierten Signale abhängig. Während eine Summation der ungedämpften Signale eine Pegelanhebung an den *peak*-Frequenzen von maximal 6 dB verursachen kann, resultiert eine Überlagerung zweier Signale mit einem Pegelunterschied von 18,2 dB in einer *peak*-Überhöhung von gerade 1 dB. Der Zusammenhang zwischen der Pegeldifferenz der beteiligten Signale (L_V -

L_O)¹¹ und dem Grad der Pegelanhebung an den *peak*-Frequenzen A ist in Abbildung 4 dargestellt und lässt sich wie folgt berechnen:

$$L_V - L_O = 20 \cdot \log(10^{\frac{A}{20}} - 1) . \quad (4)$$

Durch Umformung ergibt sich

$$A = 20 \cdot \log(10^{\frac{L_V - L_O}{20}} + 1) . \quad (5)$$

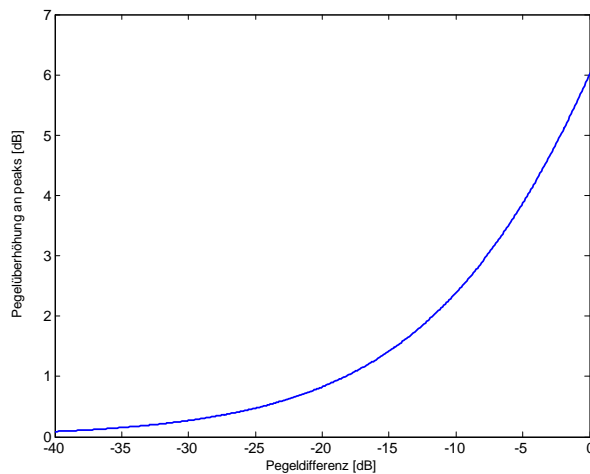


Abbildung 4: Zusammenhang zwischen der Pegeldifferenz der beteiligten Signale und Pegelanhebung an den *peak*-Frequenzen des Ausgangssignals.

2.2 Kammfiltereffekte in der Audiotechnik

In der Aufnahmekette eines Audiorecordingsystems tritt diese Superposition mit verzögerten und mehr oder weniger gedämpften Kopien vor allem in den folgenden Aufnahmesituationen auf: bei der Laufzeitstereofonie mit mehr als zwei Mikrofonen, bei der Schallaufnahme in der Nähe schallharter Begrenzungsflächen und bei der Anwendung von Stützmikrofonen zusätzlich zum Hauptmikrofonsystem. In jedem Fall liegt der Laufzeitverzögerung eine Schallwegdifferenz der sich überlagernden Signale zugrunde. Die Laufzeitdifferenz Δt berechnet sich nach:

¹¹ L_V = Pegel des verzögerten Signals, L_O = Pegel des Originalsignals.

$$\Delta t = \frac{\Delta s}{c} \quad (6)$$

c = Schallgeschwindigkeit, 343 m/s bei 20°C in Luft
 s = Schallwegdifferenz

2.2.1 Laufzeitstereofonie mit mehr als zwei Mikrofonen

Von Laufzeitstereofonie spricht man, wenn bei der Schallaufnahme Laufzeitdifferenzen zwischen Schallsignalen ausgenutzt werden, um später ein stereofones Abbild auf der Wiedergabeseite erzeugen zu können.

Während bei der Verwendung von nur zwei Mikrofonen bei der Laufzeitstereofonie die Ausgangssignale auf zwei diskrete Kanäle geroutet werden, wird das Signal eines dritten Mikrofons¹² auf den linken und rechten Kanal aufgeteilt. An dieser Stelle ist die Gefahr des Auftretens von Kammfiltereffekten gegeben, da ein Signal einem spektral nahezu identischen Signal elektrisch aufsummiert wird (siehe Abbildung 5).

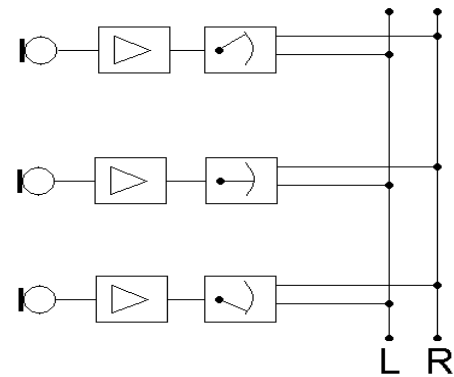


Abbildung 5: Blockschaltbild des Routings bei einer ABC-Mikrofonanordnung (nach Rode (2005), S. 10).

2.2.2 Haupt- und Stützmikrofonie

Bei der Mikrofonierung von ausgedehnten und/ oder komplexen Klangkörpern werden zusätzlich zum Hauptmikrofon oder Hauptmikrofonsystem häufig Stützmikrofone eingesetzt, um beispielsweise Solisten eines Orchesters bei der Live- oder Studiomischung extra hervorheben zu können. Werden die Signale des Haupt- und des Stützmikrofons anteilig oder vollständig auf denselben Kanal geroutet, kann es zu kammfilterartigen Verzerrungen kommen, da zeitlich verschobene

¹² Z. B. in einer ABC- Anordnung

kohärente Signale summiert werden. Die Verteilung der *peak*- und *notch*-Frequenzen richtet sich nach der Laufzeitdifferenz der Signale am Hauptmikrofonsystem, bezogen auf das jeweilige Stützmikrofon. Die Ausprägung des Kammfiltereffektes ist davon abhängig, welcher Anteil des Stützmikrofonsignals auf welchen Kanal geroutet wird.

2.2.3 Reflexion an schallharten Begrenzungsflächen

Eine weitere Ursache für das Entstehen von Kammfiltereffekten ist in der Positionierung des Mikrofons bei einer Schallaufnahme in der Nähe schallharter Begrenzungsflächen zu sehen. In dieser Aufnahmesituation wird zusätzlich zum auf das Mikrofon auftreffenden Direktschall von der Schallquelle auch der von einer Begrenzungsfläche mit geringem Absorptionsgrad reflektierte Schall auf denselben Kanal übertragen (siehe Abbildung 6). Besondere Aufmerksamkeit verdient hier der Sprechertisch in Produktionsstudios von Radioprogrammen. Völker untersuchte die Ursachen für störende Klangfarbenänderungen in Sprecherräumen und fand, dass "Sprechertisch, Typ und Anordnung des Mikrophons (...) den Ausschlag für eine mehr oder weniger starke Klangverfärbung"¹³ geben. Zur Ausschaltung der Tischreflexionen mit etwa 1 ms Verzögerung schlägt er vor, die Mikrofone in den Tisch zu versenken (siehe Abbildung 6) oder die Tischoberfläche mit absorbierendem Material zu verkleiden.¹⁴

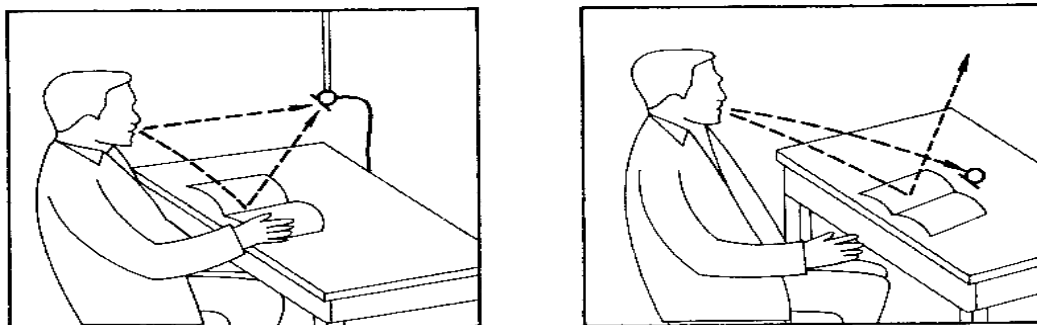


Abbildung 6: Reflexion an schallharter Begrenzungsfläche; links: ungünstige Mikrofonposition; rechts: Abhilfe durch Absenkung des Mikrofons (aus: Dickreiter , 2003, S.189).

¹³ Völker (1974), S. 224.

¹⁴ Vgl. Völker (1974), S. 224.

2.2.4 Mono-Kompatibilität

Ein auf Grund der technischen Entwicklung immer mehr in den Hintergrund geratenes Problem ist das der Wandlung stereofoner Audiosignale für die monofone Wiedergabe. Da im Consumerbereich jedoch noch immer einfache Geräte mit einkanaliger Wiedergabe angeboten und verkauft werden, soll dieser Aspekt nicht unerwähnt bleiben. Während bei der stereofonen Wiedergabe eines mittels AB-Laufzeitstereofonie aufgenommenen Signals auf Grund der zwei separaten Wiedergabekanäle keine Klangverfärbungen durch Kammfiltereffekte auftreten, kann es bei der Reduktion der beiden Kanäle auf einen einzigen Kanal durch die Aufsummierung der in diesem Fall hinreichend kohärenten Signale zu hörbaren klanglichen Verfärbungen kommen¹⁵.

2.3 Hörbarkeit von Kammfiltereffekten

Im Folgenden werden die wesentlichen bisher zur Hörbarkeit linearer kammfilterartiger Verzerrungen veröffentlichten Untersuchungen vorgestellt.

Eine Untersuchung von Müller hat die Wahrnehmbarkeit von Klangverfärbungen durch Reflexionen anhand von Sprache, Musik und weißem Rauschen zum Gegenstand. Allerdings bleibt Müller sowohl in der Beschreibung des Experiments als auch in der Auswertung desselben sehr vage. Als einzige konkrete Zahlenwerte bezüglich der Hörschwelle nennt Müller -17 dB (beim weißen Rauschen), um die das verzögerte bzw. reflektierte Signal an der Wahrnehmbarkeitsschwelle schwächer sei als der Direktschall, sowie -12 dB beim Sprachsignal.¹⁶ Da jedoch über die Art der Signalerzeugung sowie den Versuchsaufbau und die veränderlichen Parameter nichts bekannt ist, ist es fraglich, ob ein Vergleich mit den im

¹⁵ Vgl. Dickreiter (1997), S. 280 f.

¹⁶ Vgl. Müller (1968), S. E61 ff.

vorliegenden Versuch ermittelten Daten sinnvoll ist. Allgemeine Aussagen Müllers zur Klangverfärbung werden in Abschnitt 2.4 aufgegriffen.

Von Kuhl durchgeführte Hörversuche zur Ermittlung einer Hörschwelle für Kammfiltereffekte führten auf eine Pegeldifferenz zwischen Direktschall und einer diesem überlagerten, künstlich erzeugten Reflexion von ca. -17 dB (bei 2 ms zeitlicher Verzögerung zwischen den beiden Signalen) und ca. -13 dB (bei 5 bis 15 ms zeitlicher Verzögerung) an der Hörschwelle. Als Testsignal wurde rosa Rauschen verwendet.¹⁷ Weitere Einzelheiten zur Methode sind leider wiederum nicht verfügbar.

Anazawa et al. führten Hörversuche mit fünf Versuchspersonen durch, bei denen die Detektierbarkeit von Kammfiltereffekten durch das Zumischen von zeitlich verzögerten Signalen diverser Instrumente eines Orchesters und von weißem Rauschen zu einem Referenzsignal untersucht wurden. Dabei ergaben sich Schwellwerte für weißes Rauschen von 15 bis 16 dB Pegeldifferenz bei Verzögerungszeiten von 0,1 ms bis 6 ms, bei längeren Verzögerungszeiten stieg die Detektierbarkeit stark an. Die Verzerrung der Stimme wurde annähernd konstant bei einer Pegeldifferenz von 13 bis 15 dB erkannt. Kammfilterartige Verzerrungen von Trompete und Chor wurden am besten erkannt. Bei 0,7 ms Verzögerung waren selbst Verzerrungen nach dem Zumischen von Signalen mit mehr als 20 dB Pegeldifferenz detektierbar. Zwar ist für diese Untersuchung die Generierung der Stimuli dokumentiert, allerdings fehlt eine Beschreibung des Testverfahrens.¹⁸

Andere bereits erschienene Abhandlungen beschäftigen sich mit verschiedenen Detailaspekten zur Hörbarkeit kammfilterartiger linearer Verzerrungen.

Bücklein führte Versuche zur Wahrnehmbarkeit einzelner *peaks* und *dips* im Spektrum von Sprache, Musik und weißem Rauschen durch. Er fand heraus,

¹⁷ Vgl. Kuhl (1969), S. 206.

¹⁸ Anazawa et al. (1984), S. 2 ff.

dass *peaks* im Vergleich zu *dips* mit gleicher Amplitude und Güte wesentlich besser gehört werden. Die Erkennungsraten sowohl für *peaks* als auch *dips* zeigten Maxima bei 500 Hz und 1300 Hz.¹⁹

Burroughs²⁰ beschäftigte sich mit der Wahrnehmbarkeit von Kammfiltereffekten bei der Haupt- und Stützmikrofonie. Die aus seinen Untersuchungen resultierende und in der Praxis häufig angewendete 3:1-Regel besagt, dass Kammfiltereffekte, welche durch die Mikrofonanstellung bedingt sind, unhörbar werden, wenn der Abstand des Hauptmikrofons zum Stützmikrofon mindestens das Dreifache des Abstandes des Stützmikrofons zur Schallquelle beträgt. Bereits an dieser Stelle sei die Frage erlaubt, ob diese Regel auch (theoretisch) für ein Verhältnis *60cm/ 20cm* gelten würde und wo die Grenze verläuft. In Abschnitt 5.2 wird die 3:1-Regel in Bezug auf die Versuchsergebnisse der vorliegenden Untersuchung diskutiert.

Ahnert untersuchte die Hörbarkeit des Kammfiltereffektes bei Wiedergabesystemen in Abhängigkeit von der Verzögerungszeit bei zwei oder mehr Reflexionen. Diese Versuche offenbarten einen direkten Zusammenhang der Frequenzlage störender Periodizitäten zum größten gemeinsamen Teiler der Verzögerungszeiten bei zwei und mehr ankommenden Reflexionen des Lautsprechersignals. Im Rahmen des hier vorgestellten Hörversuches wird der Kammfiltereffekt allerdings mit nur einer simulierten Reflexion erzeugt. Ahnert folgert auf der anderen Seite aus Bückleins Untersuchungen, dass der Kammfiltereffekt bei einer Pegeldifferenz > 5 dB zwischen Direktschall und erster Reflexion praktisch unhörbar würde.²¹

Dickreiter stellt fest, dass es „wichtig (ist), dass (...) eine Pegeldifferenz von 10 dB, mindestens 6 dB, eingehalten wird, damit der Effekt der Welligkeit des

¹⁹ Vgl. Bücklein (1981), S. 126 ff.

²⁰ Vgl. Burroughs (1974), S.109 ff. zitiert nach: Rode (2005), S. 22 ff.

²¹ Vgl. Ahnert (1988), S. 4 f.

Frequenzgangs bzw. der Kleinräumigkeit unbedeutend bleibt.“²² Leider bleibt unklar, woher diese Zahlen stammen. Ebenfalls von Dickreiter stammt die Aussage, dass Klangverfärbungen, die durch die Monowandlung eines mittels Laufzeitstereofonie aufgenommenen Audiosignals „praktisch unhörbar“ seien, wenn der Mikrofonabstand des Aufnahmesystems kleiner als der Ohrabstand von 17,5 cm gewählt wird²³. Der vorliegende Hörversuch soll u. a. klären, ob eine kammfilterartige Verzerrung, welche unter den genannten Bedingungen entsteht, tatsächlich unhörbar ist.

Meyer ermittelte in einer Versuchsreihe, welche die Hörbarkeit linearer Verzerrungen mittels einer Variation von spektralen Einbrüchen bei natürlichen Klängen untersuchte, dass „in den mittleren und hohen Oktaven (...) mittlere Frequenzabsenkungen von 4 dB noch wahrnehmbar sind“.²⁴

Eine bisher nicht publizierte Untersuchung, welche im Rahmen des Seminars „Empirische Methoden und subjektive Evaluation“ im Jahr 2005 unter der Leitung von Stefan Weinzierl an der TU Berlin von Andreas Rotter und Robert Damrau durchgeführt wurde, beschäftigte sich mit der Wahrnehmbarkeit von Kammfiltereffekten bei einer acht Sekunden langen Popmusikaufnahme mit mehreren Instrumenten²⁵ und Gesang. Der Kammfiltereffekt wurde über ein MATLAB-Script erzeugt. Der Hörversuch wurde weitestgehend nach der Norm ITU-R BS.1116-1²⁶ durchgeführt. Das Untersuchungsergebnis zeigte eine besondere Sensibilität für Laufzeitverzögerungen um 1 ms auf. Eine Abhängigkeit der Hörbarkeit von der Tiefe der Einbrüche im Spektrum wurde nachgewiesen.

²² Dickreiter (2003), S. 157.

²³ Vgl. Dickreiter (1997), S. 280.

²⁴ Meyer (1954), S. 6.

²⁵ Schlagzeug, Gitarre und Bass

²⁶ Dies ist eine Norm für Untersuchungen zur Beurteilung von kleinen Unterschieden im Audiomaterial. Ein Hörversuch soll als doppelblinder ABC/hidden reference-Test durchgeführt werden. Der Proband bewertet den gehörten Unterschied auf einer fünfstufigen *annoyance*-Skala.

Rode untersuchte den Zusammenhang zwischen dem Diffusfeldanteil eines Signals und der Wahrnehmbarkeit von kammfilterartigen Verzerrungen, konnte jedoch eine direkte Abhängigkeit der beiden Variablen nicht nachweisen.²⁷

2.4 Klangverfärbungen durch kammfilterartige Verzerrungen

Die American Standards Association definiert Klangfarbe (*timbre*) wie folgt: “subjective impression that allows one to distinguish between two sounds of of the same fundamental loudness and pitch, presented in the same way.”²⁸

Aus dieser Definition heraus kann angenommen werden, dass es sich bei der Wahrnehmung kammfilterartiger Verzerrungen um die Wahrnehmung einer klangfarblichen Änderung handelt. Theile definiert den Begriff „Wahrnehmung“ folgendermaßen: „Wahrnehmungen sind die durch Reize ausgelösten Sinnesereignisse, die infolge der Sinneserfahrung der Außenwelt zugeordnet auftreten, und die deshalb durch Lernvorgänge und bewusste oder unbewusste Interpretationen beeinflusst werden können; sie kommen durch Assoziationen zustande.“²⁹ Diese Einflüsse sollten sich auch in der Beschreibung der Wahrnehmung des Kammfiltereffektes widerspiegeln. Da keine allgemein anerkannte Sprachregelung für die

Beschreibung durch Versuchspersonen	Anzahl der Nennungen
Tonhöhencharakter	6
nasaler Charakter	1
schlecht abgestimmte Mehrweglautsprecher	1
maschinenhaft	3
qualitätsreduziert	1
Plastikbox	1
Schwebung in den Obertönen/ instabil	5
schlechte Lokalisationsschärfe	1
wie durch Rohr	4
Halleffekt	1
flach	2
giftig	1
Höhen fehlen	2
verfremdet	1
metallisch	1
pfeifend	1

Tabelle 1: Beschreibungen des Kammfiltereffektes.

²⁷ Vgl. Rode (2005), S. 76.

²⁸ Buser & Imbert (1992), S. 94.

²⁹ Theile (1980), S. 31.

Beschreibung des Kammfiltereffektes existiert, wurde in den Fragebogen des im Rahmen der vorliegenden Magisterarbeit durchgeführten Hörversuches auch die Bitte an die Versuchspersonen integriert, vor der Durchführung des Experimentes kurz ihre Ansicht darzulegen, wie sich eine kammfilterartige Verzerrung anhört.³⁰ In Tabelle 1, S. 14, sind die nach Meinung des Autors interessantesten Äußerungen sowie die jeweilige Häufigkeit der Nennungen zusammengefasst. Dabei wurden aus ähnlichen Äußerungen Cluster gebildet. So sind in der Kategorie „Tonhöhencharakter“ auch Antworten wie „Tonhöhe entsteht“ oder „Bildung tonaler Elemente“ enthalten.

Hinsichtlich der Beschreibung der durch kammfilterartige Verzerrungen hervorgerufenen klangfarblichen Veränderungen wird in der Literatur stark nach der Beschaffenheit des Ausgangssignals differenziert. Es wird vor allem zwischen Auswirkungen bei Rauschsignalen, Sprache und Musik unterschieden.

2.4.1 Kammfiltereffekt bei Rauschsignalen

Laut Kuhl ist man „bei Rauschen sehr empfindlich gegen Klangverfärbung“.³¹ Der Kammfiltereffekt wird hier „im Charakter von Tönen hörbar, die dem Rauschen überlagert sind.“³² Diese tonalen Anteile werden als *repetition pitch*, also Wiederholtonhöhe bezeichnet.³³ Die Frequenz dieser Tonhöhe liegt bei $f=1/\Delta T$, wobei mit ΔT die Verzögerungszeit zwischen den überlagernden Signalen bezeichnet wird.³⁴ Wahrgenommen wird laut Müller³⁵ und Bilsen & Ritsma³⁶ nicht etwa der erste *peak* im Spektrum des gefilterten Signals, sondern ein Residualton, welcher aus „möglichst hohen Schalldruckwerten in Spitzennähe (im Spektrum) gebildet wird, die im

³⁰ Der komplette Fragebogen findet sich in Anhang C, S. 67.

³¹ Kuhl (1965), S. 173.

³² Müller (1968), S. E61.

³³ Vgl. Bilsen & Ritsma, (1969), S. 72.

³⁴ Vgl. Müller (1968), S. E61.

³⁵ Ebd., S. E62 ff.

³⁶ Bilsen & Ritsma (1969), S. 63.

Naturtonverhältnis (1:2:3:4 usw.) zueinander liegen“.³⁷ Das hier wirkende Prinzip entspricht dem Phänomen des Hörens der sog. *missing fundamental*“³⁸. Laut Hartmann³⁹ wird die Verfärbung besonders gut hörbar, wenn das gefilterte Signal an sich keine Tonhöhe aufweist. Die Wiederholtonhöhe wird auch als „*reflection tone*“ oder „*time separation pitch*“ bezeichnet.⁴⁰

2.4.2 Kammfiltereffekt bei Sprachsignalen

Die rauschhaften Anteile im Sprachsignal bewirken ebenfalls eine Klangverfärbung durch den *repetition pitch*.⁴¹ Bei Verzögerungszeiten von 0,2 bis 0,8 ms wurden bei Müllers Hörversuchen die Frikative auf Grund des „residualen Grundtones zwischen 1250 und 5000 Hz besonders hart und zischend“ gehört.⁴² Bei Verzögerungszeiten von um die 8 bis 10 ms kam es durch Verschiebung des Residualtones zu den tiefen Frequenzen hin erst zu einer Bassbetonung und schließlich zur Wahrnehmung von einem „Dröhnen“ oder einem „topfigen Klang“.⁴³ Müller spricht ab einer Verzögerungszeit von 10 ms von einer „Empfindung der Raumhalligkeit“.

Besondere Aufmerksamkeit verdienen Müllers Ausführungen hinsichtlich der Veränderungen der Formanten beim Auftreten von Verzögerungen von 1 – 2 ms. Am Beispiel der Vokale „u“ und „o“ beschreibt er, dass, „wenn Auslöschungen durch Interferenz mit entscheidenden Formantfrequenzen zusammenfallen“, sich die Formantstruktur so weit ändern kann, dass Laute, wie im genannten Beispiel, „nicht mehr unterschieden werden können“.⁴⁴

³⁷ Müller (1968), S. E61.

³⁸ Vgl. u.a. Buser & Imbert (1992), S. 93 f.

³⁹ Hartmann (1998), S. 364.

⁴⁰ Bilsen & Ritsma, (1969), S. 72.

⁴¹ Vgl. Müller (1968), S. E63.

⁴² Ebd., S. E63.

⁴³ Ebd., S. E63.

⁴⁴ Ebd., S. E63.

2.4.3 Kammfiltereffekt bei Musik

Eine Aussage in Bezug auf die Effekte kammfilterartiger Verzerrungen auf musikalische Darbietungen kann auf Grund der unüberschaubaren Fülle von möglichen Darbietungsformen im Rahmen dieser Arbeit nur sehr allgemein gehalten sein. Zu den prekärsten Folgen zählen sicherlich harmonische, also inhaltliche Entstellungen. Während durch Auslöschungen Melodietöne im Extremfall komplett entfernt werden können, ist es auch möglich, dass durch zusätzliche *peaks* im Spektrum disharmonische Klänge in Folge von Residualtönen entstehen können.⁴⁵ Allerdings ist dies stark von der Rauschhaftigkeit des Spektrums des Originalsignals abhängig, da es sich um eine lineare Verzerrung handelt und also keine neuen Frequenzanteile erzeugt, sondern lediglich vorhandene verstärkt werden können.⁴⁶

2.4.4 Klangverfärbungen bei veränderlichem T

Ändert sich die Verzögerungszeit ΔT kontinuierlich – beispielsweise durch Bewegung der Schallquelle oder des Mikrofons – so verschieben sich auch die *peaks* und *dips* im Spektrum. Die Veränderung eines bei rauschhaften Signalen auftretenden Tonhöhencharakters beschreibt Dickreiter mit: „etwa wie ‚üü‘ bzw. ‚üü‘“. ⁴⁷ Müller bezeichnet den Effekt gleitender Verzögerungszeiten bei Rauschsignalen als „eindrucksvoll“, bei Sprache als „eklatante Klangfarbenänderung“ und bei Musik als Grund für „fatale melodische und harmonische Änderungen“. ⁴⁸

2.4.5 Erwünschte Klangverfärbungen – Flanging und Phasing

Die im vorangegangenen Absatz erwähnten „fatalen melodischen und harmonischen Änderungen“ können jedoch durchaus erwünscht sein und werden teilweise bewusst als künstlerisches Mittel eingesetzt.

⁴⁵ Vgl. ebd., S. E63.

⁴⁶ Siehe Abschnitt 2.4.1.

⁴⁷ Dickreiter (2003), S. 17.

⁴⁸ Müller (1968), S. E63 f.

Prominentestes Beispiel ist hier das Flanging. Der Name ist bereits ein Hinweis auf die Art und Weise, wie dieser Effekt, dessen Klang meist mit Attributen wie „sphärisch“ oder „ätherisch“ beschrieben wird, ursprünglich erzeugt wurde. Das in zweifacher Ausfertigung vorliegende Audiomaterial wurde auf zwei Tonbandgeräten simultan abgespielt und die Summe der Signale aufgenommen. Nachdem es bereits in dieser Situation durch die geringen Phasendifferenzen, bedingt durch minimal unterschiedliche Bandgeschwindigkeiten, zu Klangverfärbungen kam, wurde durch Ausübung mechanischen Drucks auf die Tonbandspule (engl.: *tape reel flange*) die Bewegung derselben verlangsamt. Die auftretenden gleitenden Laufzeitdifferenzen zwischen den Signalen der beiden Bandgeräte bedingten ständig wechselnde Einbrüche bzw. Überhöhungen im Spektrum des aufgenommenen Summensignals.⁴⁹ Heutzutage wird diese Laufzeitverzögerung elektronisch oder digital erzeugt.

Während beim Flanging echte Laufzeitverzögerungen der Grund für die hörbaren Verfärbungen sind, werden die *notches* im Spektrum beim Phasing durch elektronische Phasenverschiebung erzeugt. Dieses *phase shifting* kann beispielsweise durch eine Reihenschaltung verschiedener Allpass-Netzwerke erreicht werden. Dabei gilt: je mehr Netzwerke man verknüpft, desto mehr *notches* weist das Spektrum auf. Obwohl Phasing- und Flangingeffekte ähnlich klingen, unterscheiden sie sich nicht nur in der Art und Weise ihrer Erzeugung stark. So ist das Frequenzverhältnis der *notches* beim Phasing nicht konstant und die Form der Einbrüche kann Nichtlinearitäten aufweisen.⁵⁰

⁴⁹ Vgl. Hartmann (1978), S. 439 ff.

⁵⁰ Vgl. Webers (1999), S. 376 f.

3 Methodik

3.1 Überblick über psychophysische Messmethoden

Die Psychophysik untersucht die Verbindung zwischen physikalisch messbaren Stimuli und deren psychologischen Auswirkungen.⁵¹ Da es nicht möglich ist, auf direktem Wege zu messen, wie der Mensch seine Umwelt wahrnimmt, ist der Umweg über psychoakustische Experimente zur Ermittlung von akustischen Wahrnehmungsgrößen unumgänglich. Der Prozess der Wahrnehmung eines Reizes inklusive aller Verarbeitungsstufen wird dabei als *black box* angesehen.⁵² Im Jahre 1860 beschrieb Gustav Theodor Fechner erstmals Methoden, welche erlaubten, die innere Wahrnehmung von Versuchspersonen mittels Messungen der Reaktionen dieser Versuchspersonen auf Reize in Experimenten approximativ zu quantifizieren.⁵³ Aus diesem Zusammenhang heraus lässt sich das psychophysische Gesetz, welches die Wahrnehmung W als eine Funktion des Reizes R beschreibt, formulieren.

Die psychometrische Funktion stellt eine Beziehung zwischen der physikalischen Ausprägung eines Merkmals⁵⁴ zu den Reaktionen der Versuchsperson auf ein Experiment mit binärem Ergebnis⁵⁵ her.⁵⁶ Diese Reaktionen werden dabei als Ergebnis einer Detektions- oder Diskriminationsaufgabe in „Prozent korrekt“ auf der Ordinate repräsentiert. Abbildung 7, S. 20, zeigt eine typische psychometrische Funktion.

⁵¹ Vgl. Luce (1993), S. 111.

⁵² Vgl. Jennings (2005), S. 88.

⁵³ Nach: Treutwein (1995), S. 2503.

⁵⁴ Z. B. dem Schalldruckpegel eines Stimulus

⁵⁵ beispielsweise die korrekte oder nicht korrekte Erkennung eines Stimulus

⁵⁶ Vgl. Treutwein (1995), S. 2504.

Im Laufe der Zeit haben sich diverse Verfahren zur Bestimmung von Wahrnehmungsgrößen etabliert, von denen hier die klassischen Vertreter der Vollständigkeit halber kurz erläutert werden sollen, auch wenn sie im hier vorgestellten Experiment keine Anwendung finden.

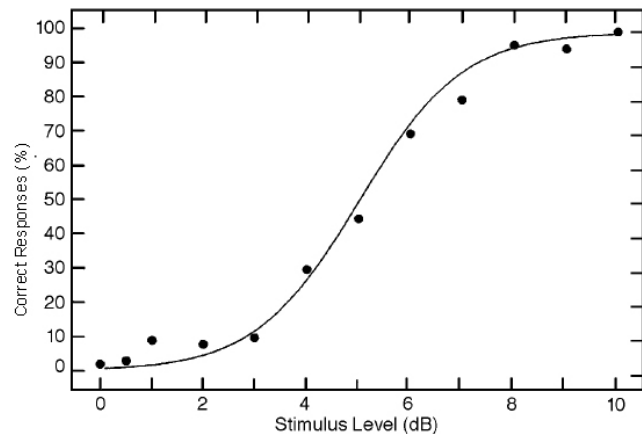


Abbildung 7: Typische psychometrische Funktion (aus: Leek (2001), S. 1280).

Das **Konstanzverfahren** (*method of constant stimuli*) beinhaltet die Darbietung einer großen Anzahl von Reizen, bei denen das zu untersuchende Merkmal in Ausprägungen von *sehr schwach* bis *sehr stark* vorkommt. Die Stimuli werden randomisiert und mehrfach präsentiert. Durch die Zuordnung der positiv korrekten Antworten zu den präsentierten Stimuli kann nun mit einem geeigneten Verfahren⁵⁷ die zu Grunde liegende psychometrische Funktion rekonstruiert werden.⁵⁸

Beim **Herstellungsverfahren** (*method of adjustment*) kontrolliert die Versuchsperson den Verlauf des Experimentes. Die Versuchsperson stellt mittels einer Kontrolle⁵⁹ eine Reizgröße⁶⁰ so ein, bis der Reiz ihrer Meinung nach einem Testton entspricht. Auf diese Weise lassen sich beispielsweise Kurven gleicher Lautheit generieren.⁶¹

⁵⁷ Klein (2001) fasst die gängigen Verfahren zur Erfassung der psychometrischen Funktion kompakt zusammen und kommentiert die jeweiligen Vor- und Nachteile.

⁵⁸ Vgl. Zwicker (1999), S. 9.

⁵⁹ Diese ist oft als Drehregler ausgeführt.

⁶⁰ Z. B. Frequenz oder Amplitude

⁶¹ Vgl. Zwicker (1999), S. 9.

Das **Grenzverfahren** (*method of limits*) wird vom Versuchsleiter kontrolliert. Die Versuchsperson hört jeweils einen Stimulus aus einer Serie mit auf- oder absteigender Merkmalsausprägung. Dabei antwortet die Versuchsperson nach jeder Präsentation, ob sie den Stimulus gehört hat oder nicht. Bei absteigenden Serien endet die Versuchsreihe mit dem ersten nicht gehörten Stimulus. Als Wahrnehmungsschwelle wird der Mittelwert des kleinsten hörbaren und des größten nicht mehr hörbaren Wertes angenommen. Äquivalent verhält es sich bei aufsteigenden Serien, die mit dem ersten gehörten Stimulus enden.⁶²

Da in der vorliegenden Arbeit ein adaptives Testverfahren angewendet werden soll, wird im folgenden Abschnitt diese Methodengruppe näher beleuchtet.

3.2 Adaptive Testverfahren

Der Vorteil von adaptiven Verfahren hinsichtlich der Verwendbarkeit im hier vorliegenden Experiment gegenüber den im voran gegangenen Abschnitt geschilderten Methoden ist die Konzentration der zu präsentierenden Stimuli auf den jeweils relevanten Bereich der psychometrischen Funktion.⁶³ Adaptive Verfahren sind zudem deutlich flexibler als beispielsweise Konstanzverfahren.⁶⁴ So werden bei Hörschwellenuntersuchungen, zu denen sich auch das hier durchgeführte Experiment zuordnen lässt, lediglich die Reaktionen der Versuchsperson auf Stimuli getestet, die in unmittelbarer Nähe der zu erwartenden Hörschwelle liegen. Die Versuchsperson bestimmt demnach durch ihr Antwortverhalten den Verlauf des Experimentes. Die heutzutage in der Praxis verwendeten adaptiven Verfahren lassen sich in drei Kategorien einteilen: die PEST-Verfahren, die *maximum likelihood* Verfahren und die *staircase* Verfahren.

⁶² Vgl. Zwicker (1999), S. 9.

⁶³ Vgl. u.a. Klein (2001), S. 1427.

⁶⁴ Vgl. Dai (1995), S. 3135.

3.2.1 PEST-Verfahren

Das *Parameter Estimation by Sequential Testing* (PEST-) Verfahren prüft nach jedem einzelnen *trial*⁶⁵ mittels eines statistischen Verfahrens, ob die Faktorstufe des Stimulus verändert werden sollte oder nicht. Bei einer dahingehend positiven Entscheidung wird die Faktorstufe in der entsprechenden Richtung um die *step size*, welche ebenfalls nach festgelegten Regeln variiert wird, verändert.⁶⁶ Der letztendliche Schwellwert ist mit der Faktorstufe, welche nach der Auswertung des letzten *trials* folgen würde, identisch. Das PEST- Verfahren wurde erstmals 1967 von Taylor und Creelman⁶⁷ beschrieben.

3.2.2 Maximum Likelihood Verfahren

Bei der Verwendung einer *maximum likelihood* Prozedur wird nach jedem *trial* die der aktuellen Testreihe zugrunde liegende komplette psychometrische Funktion geschätzt und auf der Grundlage dieses Schätzwerts die nächstfolgende Faktorstufe berechnet. Als eventuell zu berechnender Schwellwert wird einfach der 50%-Punkt auf der nach dem letzten *trial* berechneten psychometrischen Funktion angenommen.⁶⁸

3.2.3 Staircase Verfahren

Diejenigen Verfahren, bei denen sich die Werte der dargebotenen Stimuli um einen gewissen Zielwert bewegen, also in diskreten Schritten auf- und absteigende Stimulussequenzen beinhalten, werden als *staircase* Verfahren bezeichnet.⁶⁹ Je nach Antwort der Versuchsperson wird der folgende Stimulus nach einer speziellen Regel ausgewählt. Vereinfachend könnte man

⁶⁵ *Trial*: kleinste Einheit im Experiment. Ein *trial* umfasst die Darbietung eines Stimulus oder einer Gruppe von Stimuli inklusive eventueller Pausen, sowie die Reaktion der Versuchsperson darauf.

⁶⁶ Vgl. Leek (2001), S. 1282.

⁶⁷ Vgl. Taylor und Creelman (1967), S. 782 ff.

⁶⁸ Vgl. Leek (2001), S. 1283.

⁶⁹ Vgl. ebd., S. 1281.

es so ausdrücken: nach (einer oder mehreren) richtigen Antworten folgt ein schwerer zu erkennender Stimulus, nach falschen Antworten wird die Merkmalsausprägung im folgenden Stimulus verstärkt.

Der entscheidende Vorteil der *staircase* Verfahren liegt in der geringen Anzahl von notwendigen Voraussetzungen, die für die valide Implementierung dieser Verfahren erfüllt sein müssen. Laut Levitt⁷⁰ stellt lediglich die monotone Steigung der psychometrischen Funktion das Ausschlusskriterium dar. Im Gegensatz zu den weiter oben besprochenen adaptiven Testverfahren müssen also keine zusätzlichen Annahmen über die zu Grunde liegende psychometrische Funktion getroffen werden. Dies ist in dem vorliegenden Versuch von entscheidendem Vorteil, da es in der Literatur keine Hinweise auf die Form der psychometrischen Funktion gibt, die der Wahrnehmung von kammfilterartigen linearen Verzerrungen zu Grunde liegt. Laut Leek⁷¹ sind vor allem Versuche, die zur Bestimmung eines bestimmten Punktes auf der psychometrischen Funktion dienen (wie z. B. der Hörschwelle) für die Einbeziehung von *staircase* Verfahren prädestiniert. In Kontradiktion dazu rät Leek zur Verwendung einer der oben beschriebenen adaptiven Prozeduren, wenn das Ziel des Experimentes die Rekonstruktion des gesamten Verlaufs der psychometrischen Funktion ist.

Ein großer Vorteil der *staircase* Verfahren ist die Einfachheit, mit der nach beendeter Testreihe der *threshold* ermittelt werden kann: eine einfache arithmetische Mittlung über die Werte der Merkmalsausprägung der Stimuli bei den letzten Umkehrungen ist nach Leek⁷² der effektivste Weg.

Die Regeln, nach denen der folgende Stimulus auf Grund einer aktuellen Antwort der Versuchsperson ausgewählt wird, teilen die *staircase* Verfahren in drei Gruppen ein.

Simple Up-Down Staircases

Diese "Urform" der *staircase* Verfahren folgt der Regel, dass nach jeder richtigen Antwort der Versuchsperson das untersuchte Merkmal im nächsten

⁷⁰ Vgl. Levitt (1971), S. 468.

⁷¹ Vgl. Leek (2001), S. 1280; Leek (1992), S. 255.

⁷² Vgl. Leek (2001), S. 1284.

Stimulus um eine Faktorstufe schwächer und nach jeder falschen Antwort stärker ausgeprägt ist (siehe Abbildung 8). Diese Form der adaptiven Untersuchungsmethode zielt auf den 50%- Punkt auf der psychometrischen Funktion: d. h. die Mehrzahl der dargebotenen Stimuli befinden sich im Umfeld dieses Punktes. Dies bedeutet somit auch, dass die Methode bei 50 Prozent korrekten Antworten konvergiert.

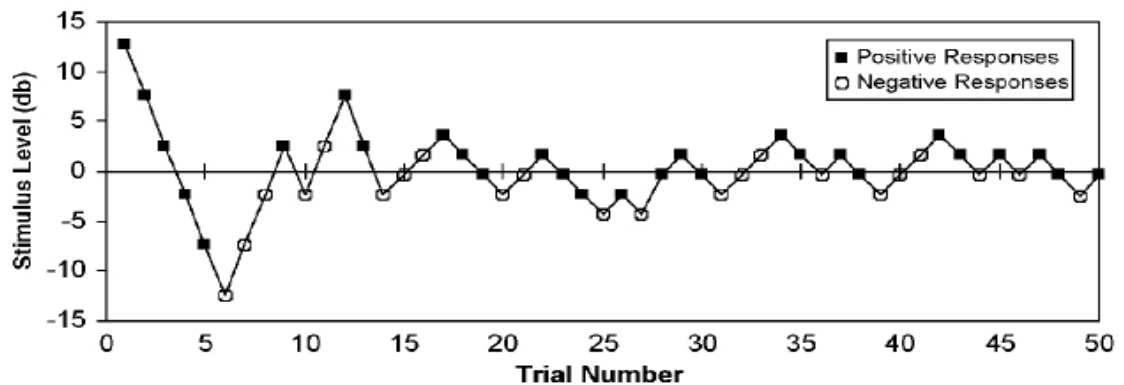


Abbildung 8: Simple Up-Down Staircase (aus: Leek (2001), S. 1285).

Für Yes/No-Detektionsaufgaben ist dieses Konvergenzlevel akzeptabel, für *multiple alternatives forced choice* Aufgaben liegt dieser Punkt viel zu tief. Da schon unter Verwendung einer *two alternative forced choice* Aufgabe der Mittelpunkt zwischen Ratewahrscheinlichkeit (50%) und perfekter Erkennung (100%) bei 75 Prozent liegt, muss der Zielpunkt des adaptiven Verfahrens verschoben werden. Dies ist durch eines der beiden folgenden Verfahren möglich.

Weighted Up-Down Staircases

Diese von Kaernbach⁷³ erstmals vorgestellte Methode bietet den Vorteil, jeden Punkt auf der psychometrischen Funktion als Zielpunkt auswählen zu können. Zur wirklichkeitsnahen Rekonstruktion einer kompletten psychometrischen Funktion ist diese Methode wohl bestens geeignet. Allerdings setzt der immens große Versuchsaufwand auf Grund der hohen Anzahl nötiger Versuche praktische Grenzen.

⁷³ Vgl. Kaernbach (1991), S. 227 ff.

Der Ansatz Kaernbachs ist, dass sich die Merkmalsausprägung, im Grunde ähnlich dem einfachen *up-down* Verfahren, nach jeder Antwort in die eine oder andere Richtung ändert. Der entscheidende Unterschied zum traditionellen *up-down* Verfahren ist, dass das Verhältnis der *step size* für eine Veränderung der Merkmalsausprägung nach „oben“ oder „unten“ nicht konstant=1 ist, sondern je nach gewünschtem Zielpunkt auf der psychometrischen Funktion frei gewählt werden kann. Das Kriterium „richtige Antworten“ für die Konvergenz der Prozedur lässt sich nach

Entry	Response Sequences		p
	Up	Down	
1	-	+	0.500
2	--+ or +-	-++ or -++	0.550
3	+- or -	+--+ or ++	0.597
4	+++ or +-	+++ or -+	0.648
5	+- or -	++	0.707
6	+++ or +++ or +-	+++ or +++	0.749
7	+++ or +-	+++	0.794
8	+++ or +++ or +- -	++++	0.841

Tabelle 2: Gruppierung verschiedener Antwortmuster bei transformierten Up- Down-Verfahren nach der Wahrscheinlichkeit p korrekter Antworten am Konvergenzpunkt (nach: Brown (1996), S. 2).

$$S_{up} * p = S_{down} * (1 - p) \tag{7}$$

S =stepsize
 p = Wahrscheinlichkeit richtiger Antworten

berechnen⁷⁴. So führt etwa die Wahl einer *up-stepsize* von „3“ und einer *down-stepsize* von „1“ zu einer Konvergenz der Prozedur bei 75 Prozent richtigen Antworten.

Transformed Up-Down Staircases

Die ebenfalls von Levitt erstmals vorgestellten transformierten *up-down staircases* legen den nächsten zu präsentierenden Stimulus nicht nur auf Basis der aktuellen Antwort der Versuchsperson fest, sondern beziehen noch weitere vorangegangene Antworten der Versuchsperson mit ein.⁷⁵ Das

⁷⁴ Vgl. Kaernbach (1991), S. 227.

⁷⁵ Vgl. Levitt (1971), S. 472 ff.

Regelsystem ist theoretisch unendlich erweiterbar – die gängigsten Regeln, zusammen mit den Zielpunkten auf der psychometrischen Funktion, sind in Tabelle 2, S. 25, zusammengefasst. Abbildung 9 zeigt die häufig verwendete 3-Down-1-Up Prozedur im Beispiel. In diesem Fall wird immer dann die Ausprägung des untersuchten Merkmals um die *step size* verringert, wenn zwei korrekte Erkennungen vorausgegangen sind. Im Gegensatz dazu wird bei jeder Nichterkennung im folgenden *trial* der Stimulus mit einer um die *step size* höheren Merkmalsausprägung präsentiert. Die Wahrscheinlichkeit einer korrekten Antwort am Konvergenzpunkt der Prozeduren⁷⁶ lässt sich wie folgt berechnen:

$$[P(X)]^n = 0,5 \quad (8)$$

n=Anzahl der notwendigen korrekten Antworten für eine Abschwächung des untersuchten Merkmals.

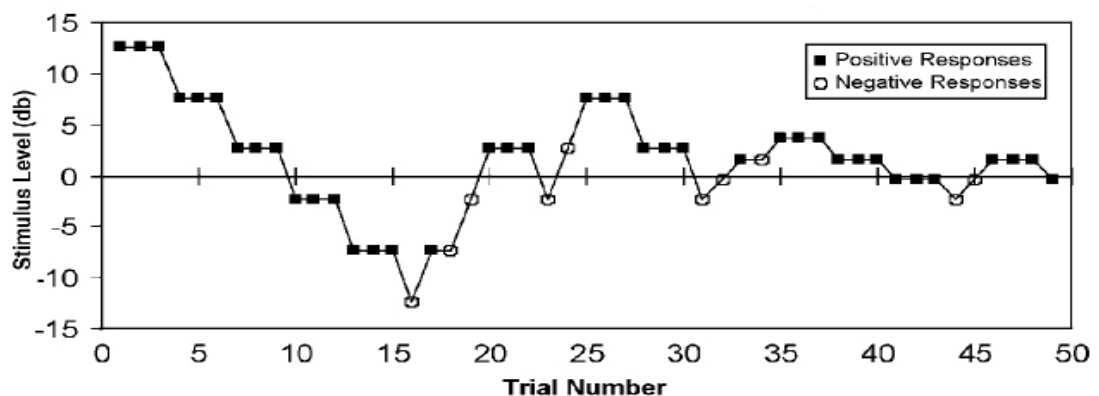


Abbildung 9: 3-Up-1-Down Staircase (aus: Leek 2001, S.1285).

⁷⁶ Sprich: am Zielpunkt auf der psychometrischen Funktion

4 Untersuchung

4.1 Quellsignale

4.1.1 Auswahl und Vorbereitung

Der vorliegende Hörversuch soll eine möglichst gute Approximation des realen Verlaufes der Hörschwelle für in der Praxis des Schallaufnahme- und Verarbeitungsprozesses vorkommende lineare kammfilterartige Verzerrungen liefern. Dazu scheint es erstrebenswert, eine Bandbreite an Signalen entsprechend der tatsächlich bei der Tonaufnahme vorkommenden Verzerrungen zu simulieren. Diese Prämisse bezieht sich vor allem auch auf die Auswahl der Quellsignale zur späteren Generierung der Stimuli. Dabei wird Wert auf eine gewisse Diversität der spektralen und temporären Struktur der Signale gelegt. Um nicht die praktische Durchführbarkeit des Hörversuches durch zu großen Zeitaufwand für die Versuchspersonen zu gefährden, beschränkt sich die Quellauswahl auf eine Klavierphrase, eine eingesprochene Wortgruppe eines männlichen Sprechers und einen Wirbel auf einer Snare. Dabei weist das Klavier eine charakteristische harmonische Obertonstruktur auf (Abbildung 10, S. 28), die Snare ein eher rauschhaftes, fast kontinuierliches Spektrum (Abbildung 11, S. 29) und im Sprachsignal finden sich rauschhafte und tonale Anteile (Abbildung 12, S. 29). Da der Kammfiltereffekt in einer Veränderung des Spektrums resultiert, steht zu erwarten, dass sich die Ergebnisse für diese drei Signalgruppen unterscheiden.

Die Quellsignale stammen von der auf der European Broadcasting Union kostenfrei zur Verfügung stehenden SQUAM-CD (*Super Quality Audio Material* CD).⁷⁷ Aus den in einer Auflösung von 44,1 kHz bei 16 Bit vorliegenden Samples wurden nun mit einem Sample-Editor Stücke mit Längen von approximativ einer, zwei, drei und vier Sekunden herausgeschnitten, wobei darauf geachtet wurde, sinnentstellende Schnitte

⁷⁷ Die auf der CD enthaltenen Samples stehen auch auf der EBU-Homepage gratis zum Download zur Verfügung -

http://www.ebu.ch/en/technical/publications/tech3000_series/tech3253/index.php .

zu vermeiden. So wurde im Fall des Sprechers nicht mitten im Wort und beim Klavier nicht inmitten eines Akkordes geschnitten.

4.1.2 Spektrale Struktur

Die drei Kategorien der Quellsignale unterscheiden sich stark in ihrer spektralen Struktur.

Das Klaviersample weist deutliche *peaks* an den Stellen der Grundfrequenz der jeweils gespielten Note⁷⁸, sowie an den Stellen der jeweiligen harmonischen Obertöne auf. Dabei fällt die Amplitude der Obertöne mit steigender Frequenz ab. Die auffälligsten *peaks* finden sich bei 334 Hz, 392 Hz, sowie bei 80 Hz, 220 Hz und 500 Hz, sowie den entsprechenden Obertönen. Ein aussagekräftiger Ausschnitt des Leistungsdichtespektrums für das unverzerrte Klaviersignal ist in Abbildung 10 dargestellt.

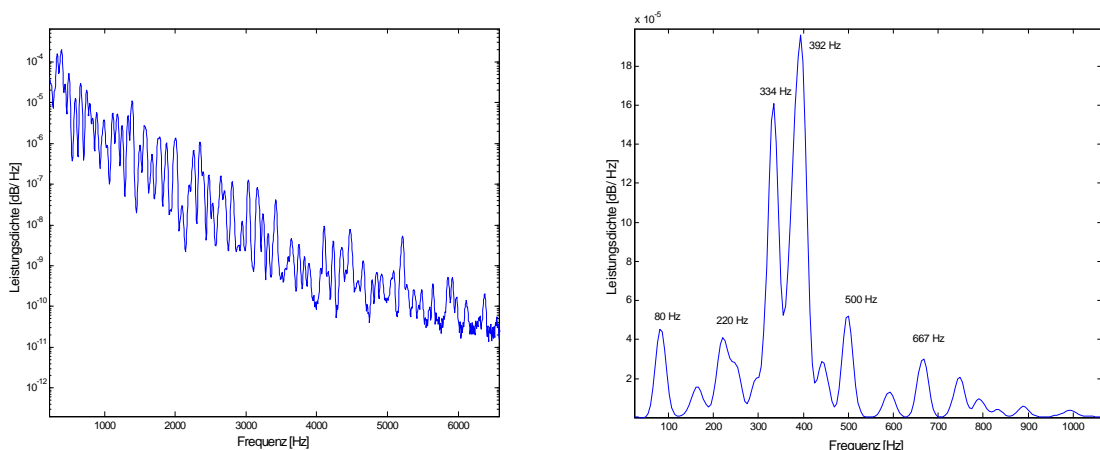


Abbildung 10: Leistungsdichtespektrum des unverzerrten Pianosignals; links: Gesamtstruktur mit logarithmischer y-Achsenkalierung; rechts: Ausschnitt mit linearer y-Achsenkalierung.

Das Snare-Signal weist ein sehr obertonreiches, rauschhaftes Spektrum auf. Auffällig ist dabei lediglich ein *peak* bei 231 Hz. Ein Ausschnitt des Leistungsdichtespektrums ist in Abbildung 11, S. 29, dargestellt.

Im Spektrum des Sprachsignals finden sich deutliche *peaks* an den Stellen der Formanten der Vokale und stimmhaften Konsonanten. Das

⁷⁸ D. h. der Resonanzfrequenz der jeweils angeschlagenen Saite(n)

Vorhandensein von Frikativen im Sprachsample führt zu rauschhaften Anteilen im Spektrum (Abbildung 12). Auffällig sind die Senken im Leistungsdichtespektrum, bei denen es sich um eine bereits im Originalsignal vorhandene kammfilterartige Verzerrung handeln könnte. Dies stellt für die Untersuchung zunächst kein Problem dar, da das Referenzsignal trotzdem stets das selbe ist.

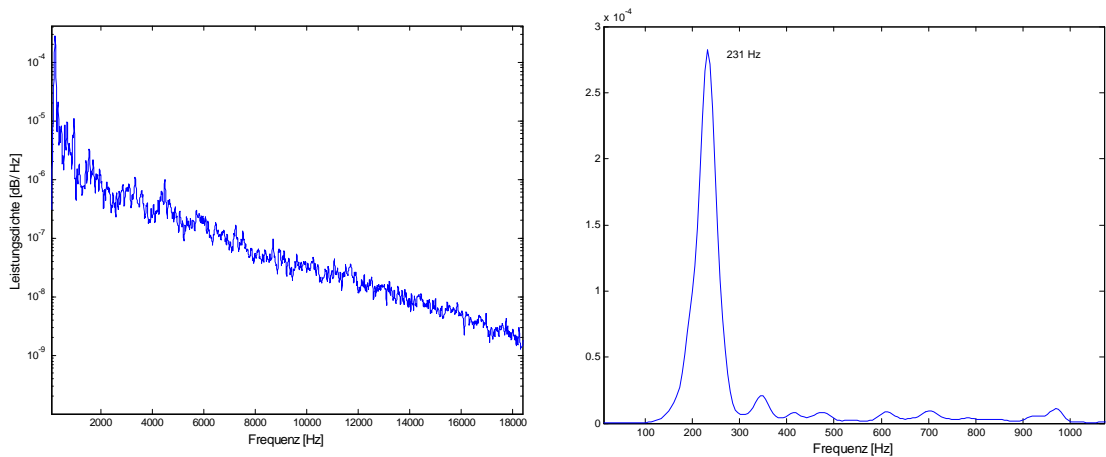


Abbildung 11: Leistungsdichtespektrum des unverzerrten Snaresignals; links: Gesamtstruktur mit logarithmischer y-Achsenkalierung; rechts: Ausschnitt mit linearer y-Achsenkalierung.

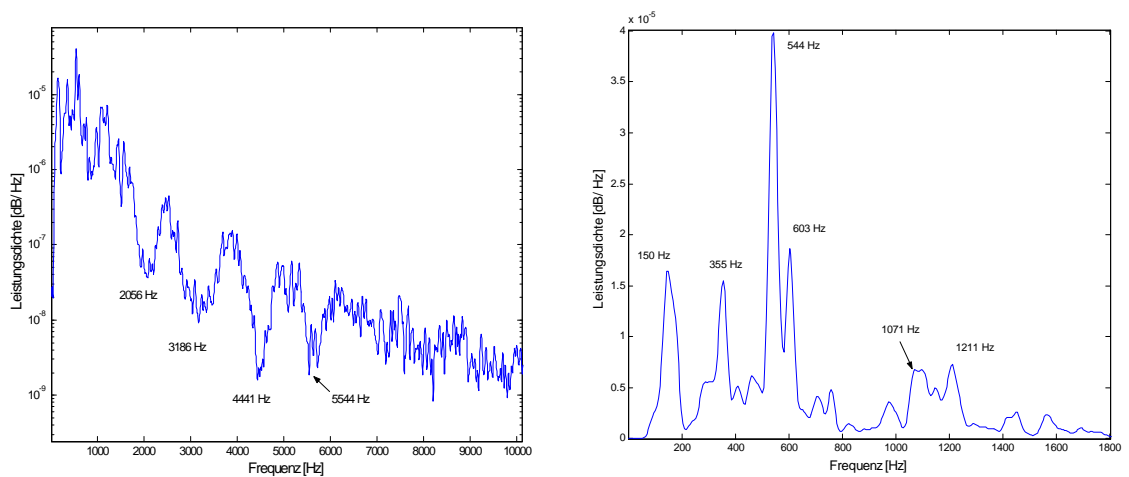


Abbildung 12: Leistungsdichtespektrum des unverzerrten Sprachsignals; links: Gesamtstruktur mit logarithmischer y-Achsenkalierung; rechts: Ausschnitt mit linearer y-Achsenkalierung.

4.2 Generierung der Stimuli

Eine MATLAB-Routine berechnet und erzeugt die Stimuli. Dabei wird das als Vektor vorliegende Originalsignal nach der Schaffung von ausreichendem Headroom zu einer um n Samples verschobenen und mit dem Faktor $L = 1$ multiplizierten Kopie addiert. Die Variable n wird so gewählt, dass Stimuli mit den Verzögerungszeiten $T = 0.1, 0.5, 0.8, 1, 1.2, 1.5, 2, 3, 4, 6, 10$ und 15 ms entstehen. Diese Werte orientieren sich an Erkenntnissen von Müller, der „durch subjektive Hörvergleiche“ feststellt, dass „die im Raum beobachteten Klangverfärbungen (...) vollständig durch elektroakustisch simulierte Reflexionen mit einer Verzögerungszeit von 0,5 bis 15 ms nachgebildet werden können.“⁷⁹ Laut Dickreiter „führen Reflexionen (...) mit einer Verzögerungszeit zwischen etwa 0 und 15 ms zu unangenehmen Klangverfärbungen.“⁸⁰ Tabelle 3 setzt die für diesen Hörversuch untersuchten Verzögerungszeiten mit den theoretisch zu Grunde liegenden Schallwegdifferenzen nach Gleichung 6 in Verbindung.

Die Variation von L resultiert in einer unterschiedlich starken Ausprägung des Kammfiltereffektes und berechnet sich nach

$$L = 10^{\frac{dB_{plus}}{20}} - 1, \quad (9)$$

wobei dB_{plus} im Erzeugungsalgorithmus die verschiedenen Stufen der Überhöhungen des Kammfiltersignals an den *peaks* in dB repräsentiert.

Dabei entstehen 24 Abstufungen, die mittels des Grades der Anhebung des Pegels an den Stellen der Phasengleichheit von Original- und Verzögerungssignal quantifiziert werden können. So entstehen Anhebungen

Verzögerungszeit [ms]	0,1	0,5	0,8	1	1,2	1,5	2	3	4	6	10	15
Schallwegdifferenz [m]	0,03	0,17	0,27	0,34	0,41	0,51	0,69	1,03	1,37	2,06	3,43	5,15

Tabelle 3: Zusammenhang zwischen Verzögerungszeiten und Schallwegdifferenz der überlagerten Signale.

⁷⁹ Müller (1968), S. E-61.

⁸⁰ Dickreiter (2003), S. 14 ff.

von 0,25 bis 6 dB in 0,25 dB-Schritten.⁸¹ Zur Verdeutlichung: Ein Faktor $L=1$ ⁸² und eine Verschiebung um 44 Samples simulieren ein (theoretisches) Signal, welches entstehen würde, wenn ein Quellsignal mit einer Verzögerung von 0,1 ms sich selbst ungedämpft überlagert würde. Die Samplingrate und Bitauflösung wurde während des Generierungsprozesses nicht verändert.⁸³

4.3 Anpassung der Lautheiten

Die so entstandenen Stimuli unterscheiden sich hinsichtlich ihrer Lautheit zum Teil beträchtlich. Da jedoch die Detektionsaufgabe von den Versuchspersonen nicht mittels der Detektion eines Lautheitsunterschiedes bewerkstelligt werden soll, müssen die Lautheiten der Stimuli angepasst werden. Dies geschieht mittels einer MATLAB-Routine, welche zuerst die Lautheit des betreffenden Stimulus ermittelt. Die Messung an sich erfolgt mittels eines MATLAB-Scripts von Aaron Hastings⁸⁴ und richtet sich nach DIN 45631⁸⁵. Die ermittelte Lautheit wird mit der Lautheit des Referenzsignals verglichen. Der Vektor des anzupassenden Stimulus wird in einer Feedbackschleife so lange mit einem Faktor multipliziert, bis die Differenz der Lautheiten 0,5 sone nicht mehr überschreitet. Bei gemessenen Lautheiten von 25-35 sone liegt der Quotient L / L ⁸⁶ bei maximal 0,02, deutlich unter dem von Ekman angegebenen Schwellwert für die Wahrnehmung von Lautheitsunterschieden von 0,03 (Ekman Law⁸⁷). Die

⁸¹ Bei der Generierung der Stimuli entstehen selbstredend nicht ausschließlich *peaks*, sondern ein Signal, welches die in Abschnitt 2.1.1 beschriebene Struktur aufweist.

⁸² Dies entspricht 6 dB Anhebung an einer *peak*-Frequenz und maximaler Auslöschung an der Stelle eines *notches*.

⁸³ Es galt weiterhin: 44,1 kHz Samplingrate und 16 Bit Auflösung.

⁸⁴ Das MATLAB-Script von Aaron Hastings kann unter <http://www.auditory.org/mhonarc/2000/msg00498.html> heruntergeladen werden.

⁸⁵ DIN 45631: Berechnung des Lautstärkepegels und der Lautheit aus dem Geräuschspektrum. Verfahren nach E. Zwicker.

⁸⁶ L bezeichnet hier die Lautheit.

⁸⁷ Vgl. Ekman, G. (1959), S. 343 ff. zitiert nach: Allen, Jont (2000), S. 222.

spektrale Struktur der Signale bleibt von der Lautheitsanpassung unberührt, da lediglich der Gesamtpegel des Signals angehoben oder abgesenkt wird.

4.4 Design des Experimentes

4.4.1 Grundlegendes Design

Das vorliegende psychometrische Experiment wurde als Detektionsexperiment ausgeführt. Detektionsexperimente können als eine spezielle Form des Diskriminationsexperiments angesehen werden. Bei letzteren handelt es sich um Versuche, die die Untersuchung der Fähigkeit der Probanden, zwei Stimuli zu unterscheiden, zum Gegenstand haben. Das hier am häufigsten verwendete Testdesign ist der klassische „Ja/Nein-Test“, der von der Versuchsperson eine Entscheidung erwartet, ob sie beispielsweise einen Ton gehört oder nicht gehört hat. Bei komplexeren Veränderungen im Stimulusmaterial wird häufig auf die sog. *Forced Choice* Experimente zurück gegriffen. Hier wird das Audiomaterial in zwei oder mehr zeitlich getrennten Alternativen dargeboten. Eine der Alternativen beinhaltet die „heiße Probe“. Die Versuchsperson soll nun entscheiden, bei welcher der Alternativen es sich um den Zielstimulus handelte.

Im vorliegenden Experiment wurde sich nach dem ersten Vorversuch⁸⁸ für ein *Three Alternative Forced Choice* Design entschieden. Die Ratewahrscheinlichkeit (P_{guess}) liegt bei dieser Methode mit

$$P_{\text{Guess}} = \frac{1}{n} \quad (10)$$

n=Anzahl der Alternativen

bei 33,3%. Im Unterschied zu der 2AFC-Methode liefert das 3AFC-Verfahren laut einer vergleichenden Untersuchung von Shelton und Scarrow auch bei einer geringen Anzahl an *trials*⁸⁹ stabile Daten für die Ermittlung einer Hörschwelle.⁹⁰

⁸⁸ Siehe auch: Abschnitt 4.5.1 & Abschnitt 4.5.2

⁸⁹ Die Anzahl der *trials* lag in den Hauptversuchen der im Rahmen dieser Arbeit vorgenommenen Untersuchung bei 30-40.

⁹⁰ Vgl. Shelton & Scarrow (1984), S. 391.

4.4.2 Weiterführendes Design

Um zu gewährleisten, dass die dargebotenen Stimuli möglichst in der Nähe der zu ermittelnden Schwelle liegen, wird ein transformiertes *staircase* Verfahren angewendet. Diese in Abschnitt 3.2.3 eingehend beschriebenen Verfahren beinhalten ein Set an Regeln, nach denen der im nächsten *run* zu detektierende Stimulus ausgewählt wird. Die zu präsentierenden Stimuli sollten in der Nähe der zu erwartenden Hörschwelle liegen. Dies wird durch drei Maßnahmen erreicht:

Erstens wird der Ausgangswert (*initial value*) jeder Versuchsreihe auf Grundlage der Ergebnisse der Versuchspersonen im ersten Vorversuch im Rahmen der zwei Sekunden langen Testsamples ausgewählt. Der *initial value* liegt immer zwei Stufen (das entspricht 0,5 dB Anhebung in den Peaks) über dem in diesem Vorversuch ermittelten *threshold*⁹¹. Da sich nach dem zweiten Vorversuch drei Sekunden als optimale Samplelänge herausstellten, ist zu erwarten, dass der Initialwert deutlich über der späteren Hörschwelle liegt. Die Wahl dieses *initial value* folgt u. a. einer Anmerkung Kleins, der einen der Vorteile der adaptiven *staircase* Verfahren darin sieht, dass sie den Probanden bereits per se mit dem zu detektierenden Stimulus vertraut machen⁹² (im Gegensatz z. B. zu den *maximum likelihood* Verfahren). Deshalb scheint es ratsam, den Anfangswert einer Reihe ein Stück über dem zu erwartenden Schwellwert zu platzieren. Dies ist auch ein Grund, weshalb die Trainingsphase im Hauptversuch relativ kurz gehalten werden kann. Im Grunde sind die ersten *trials* jeder neuen Reihe bis zum jeweils ersten *reversal* eine auf den jeweilig zu detektierenden Stimulus abgestimmte Trainingsphase.

Zweitens werden die Stimuli nach der 3-Down-1-Up-Regel (3D1U-Regel) ausgewählt und präsentiert. Die Regel besagt, dass erst die positive

⁹¹ Siehe Tabelle 10, S. 60.

⁹² Vgl. Klein (2001), S. 1437.

Erkennung von drei aufeinander folgenden Stimuli bewirkt, dass im folgenden *trial* ein schwerer zu erkennender Stimulus ausgewählt wird. Die 3D1U-Regel zielt auf den 79,1-Prozent-Punkt auf der psychometrischen Funktion. Das Verfahren konvergiert demnach, wenn 79,1 Prozent der Stimuli richtig erkannt werden. Im Gegensatz dazu konvergiert das häufig benutzte 2-Down-1-Up-Verfahren bei 70,4 Prozent richtiger Antworten. Es wird das 3D1U-Verfahren benutzt, da die Platzierung der Stimuli in der Nähe des *sweetpoints*, also des optimalen Wertes für das Präsentationslevel der Stimuli, eine geringere Standardabweichung der Ergebnisse, sowie eine kürzere Versuchsdauer verspricht.⁹³ Green legt sich nach Simulationen sowie Hörversuchen mit realen Versuchspersonen auf einen Wert von 84% bis 94% für diesen *sweetpoint* fest⁹⁴ - entgegen der häufig in der Literatur angegebenen 75% korrekter Antworten. Man mag einwenden, dass sich bei einem 4-Down-1-Up-Verfahren, welches auf den 84,1%-Punkt zielt, die präsentierten Stimuli noch näher am vermuteten *sweetpoint* befunden hätten. Gegen die Verwendung dieser Regel spricht erstens, dass sie den Versuch über eine vernünftige Dauer hinaus verlängert hätte, wodurch zusätzliches Bias durch Müdigkeit und nachlassende Konzentration zum Tragen gekommen wäre und dass zweitens die Verwendung der 3D1U-Regel in Kombination mit einer *Three Alternative Forced Choice* Detektionsaufgabe laut Kollmeier⁹⁵ und Leek⁹⁶ empirisch und mathematisch verifiziert die effizienteste Methodenkombination für adaptive Hörschwellenbestimmung darstellt, während Green seine Simulationen unter Verwendung einer 2AFC-Methode durchführte⁹⁷.

Zum Dritten wird die *stepsize* sukzessive verringert. Die *stepsize* beschreibt, um welchen Wert sich das aktuelle Präsentationslevel vom vorherigen im Falle einer Veränderung unterscheidet. Dies ermöglicht eine optimale

⁹³ Vgl. Leek (2001), S. 1290.

⁹⁴ Vgl. Green (1990), S. 2662.

⁹⁵ Vgl. Kollmeier (1988), S. 1861.

⁹⁶ Vgl. Leek (2001), S. 1290.

⁹⁷ Vgl. Green (1990), S. 2662.

Abfolge der Stimuli: zuerst wird grob das Gebiet eingegrenzt, in dem der *sweetpoint* ungefähr liegt, um dann mit geringeren Werten der *step size* die Stimuli in unmittelbarer Nähe der Hörschwelle zu platzieren. Die *step size* wird nach der ersten Umkehrung von 0,5 dB auf 0,25 dB verringert.⁹⁸

4.4.3 Software

Zur Steuerung der Hörtests sowie zur Erfassung der Antworten der Versuchspersonen und Auswertung der Versuche wurde eine MATLAB-Software erarbeitet und programmiert. Die dokumentierten Quellcodes sind auf der beiliegenden CD enthalten.⁹⁹ Das Programm ist in verschiedene Module unterteilt. Der Programmablauf während einer Versuchssession wird mittels Abbildung 13 verdeutlicht.

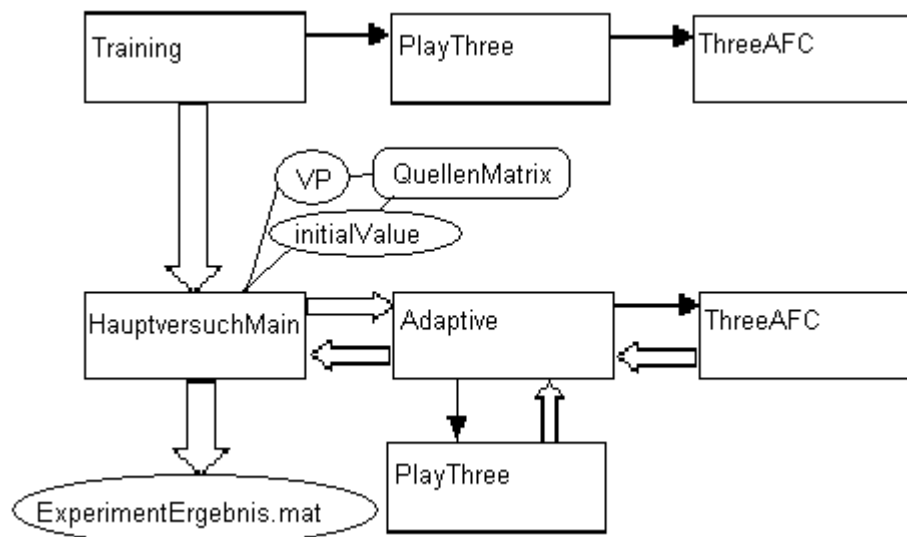


Abbildung 13: Schema des Ablaufs der MATLAB-Routine. Die Namen der Blöcke entsprechen den Funktions- und Scriptnamen.

4.5 Vorversuche

Um den Hauptversuch zu optimieren wurden zwei Vorversuche durchgeführt. Der erste Vorversuch sollte dabei ein allgemeines Feedback für die Durchführbarkeit des Experimentes nach der gewählten Methode geben,

⁹⁸ Vgl. Levitt (1971), S. 471.

⁹⁹ Eine Inhaltsbeschreibung der CD, sowie eine Kurzanleitung für die Software findet sich in Anhang D, S. 68.

Daten für die Festlegung der *initial values* liefern sowie den Wertebereich der Variable *Verzögerungsdauer* für den zweiten Vorversuch eingrenzen. In diesem sollte die optimale Länge der Stimuli ermittelt werden, da während des ersten Vorversuches keine eindeutige Tendenz festzustellen war und diese Variable im Hauptversuch aus forschungspraktischen Gründen¹⁰⁰ konstant zu halten war.

4.5.1 Erster Vorversuch

Stimuli

Es wurden Stimuli der Kategorie „Piano“, „Sprache“ und „Snare“ mit einer Länge von einer und zwei Sekunden präsentiert und die Verzögerungszeiten $T = 0,1; 0,5; 0,8; 1; 1,2; 1,5; 2; 3; 4; 6; 10$ und 15 ms getestet.

Versuchspersonen

Als Versuchspersonen stellten sich eine Studentin und sieben Studenten der TU Berlin im Alter von 21 bis 37 Jahren zur Verfügung. Auf einer Skala von 1 bis 6 (1: sehr gut, 6: sehr schlecht) schätzten die Versuchspersonen ihr Hörvermögen im Bereich 1 bis 3 ein. Die Versuchspersonen nahmen freiwillig an dem Experiment teil und erhielten keine Bezahlung. Für den Versuch wurden die Probanden in zwei Gruppen eingeteilt. Für jede Gruppe wurde aus Zeitgründen lediglich ein Teil der Verzögerungszeiten getestet (Gruppe A: 0,1; 0,8; 1,2; 2; 4 und 10 ms; Gruppe B: 0,5; 1; 1,5; 3; 6 und 15 ms).

Versuchsaufbau

Der Hörversuch lief automatisch softwaregesteuert ab. Die Stimuli wurden den Versuchspersonen über elektrostatische Kopfhörer der Firma STAX präsentiert. Diese waren an den Line Out Port der Soundkarte eines Windows-Computers angeschlossen, der laut Herstellerinformationen einen

¹⁰⁰ Die Versuchsdauer im Hauptversuch hätte bereits bei zwei verschiedenen zu testenden Längen der Samples eine Stunde weit überschritten.

linearen Frequenzgang im relevanten Bereich aufweist. Die im Mono- Format vorliegenden Samples wurden binaural präsentiert.

Als Bedienoberfläche diente ein in MATLAB programmiertes Graphical User Interface (siehe Abbildung 14).¹⁰¹

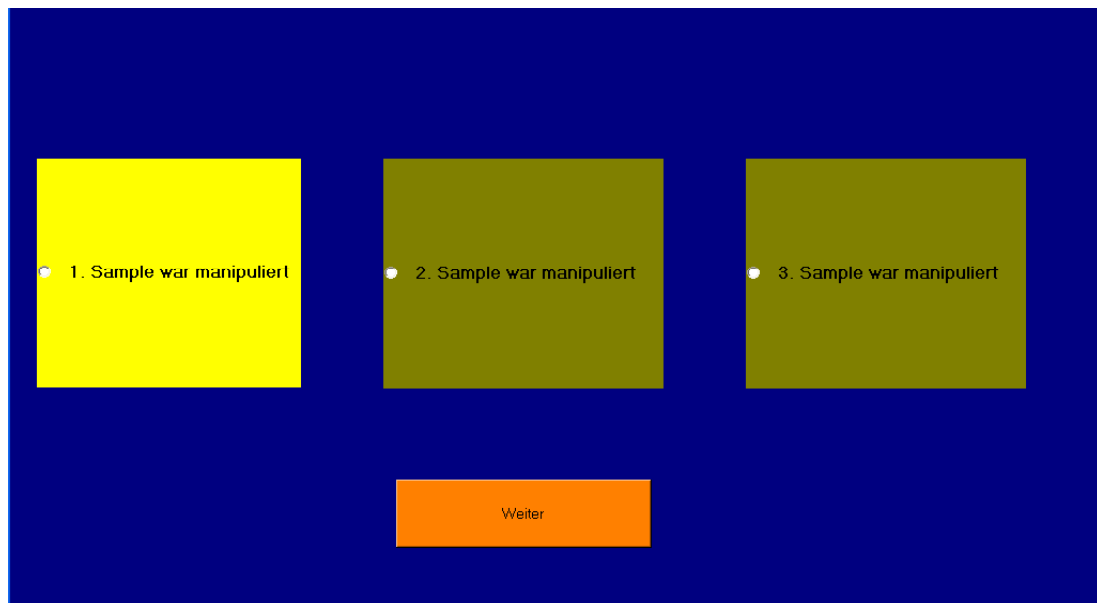


Abbildung 14: Screenshot während einer Versuchsreihe aus dem Hauptversuch. Gerade wird das erste Sample abgespielt.

Versuchsablauf

Nach einer Instruktion begann der Vorversuch mit einer Trainingsphase, während derer der Proband Extrembeispiele der später zu detektierenden verzerrten Stimuli zu hören bekam. Die Versuchsperson konnte sich die Stimuli im Vergleich mit einem unverzerrten Signal so oft anhören wie sie wollte. Nach 18 Durchläufen mit Signalen aus allen drei Kategorien (Piano, Snare, Sprache) folgte die zweite Trainingsphase, in der mit den identischen Stimuli der eigentliche Versuchsablauf simuliert wurde, um die Versuchsperson auf das Procedere einzustellen. Insgesamt nahm die Trainingsphase ca. 10 Minuten in Anspruch.

Der eigentliche Versuch wurde als *Two-Interval-Two-Alternatives Forced Choice* Test durchgeführt. Der Proband hörte nacheinander vier

¹⁰¹ Die obige Abbildung zeigt einen Screenshot des Hauptversuches, da das Design nach den Vorversuchen noch optimiert wurde.

Stimuli in zwei Paaren. Die Paare waren durch eine Pause von 0,5 Sekunden getrennt, die Stimuli innerhalb eines Paares durch eine Pause von 0,2 Sekunden. Die Aufgabe der Versuchsperson war es laut Instruktion nun, durch Klicken auf die entsprechende Schaltfläche zu entscheiden, in welchem Paar sie einen Unterschied zwischen den Stimuli wahrgenommen hat. Durch einen weiteren Klick auf die Schaltfläche „Weiter“ wurde zum nächsten Stimulus übergegangen. Getestet wurden Stimuli aller drei Kategorien mit Verzögerungszeiten entsprechend der Gruppenzugehörigkeit des Probanden. Im Anschluss an den Versuch wurden die Versuchspersonen gebeten, einen kurzen Fragebogen auszufüllen, sowie persönliche Anmerkungen zum Versuchsablauf bzw. Verbesserungsvorschläge zu machen.

Ergebnisse und Interpretation

Abbildung 15 fasst die Ergebnisse des ersten Vorversuches zusammen. Bereits durch die visuelle Auswertung ist erkennbar, dass die Snare besser

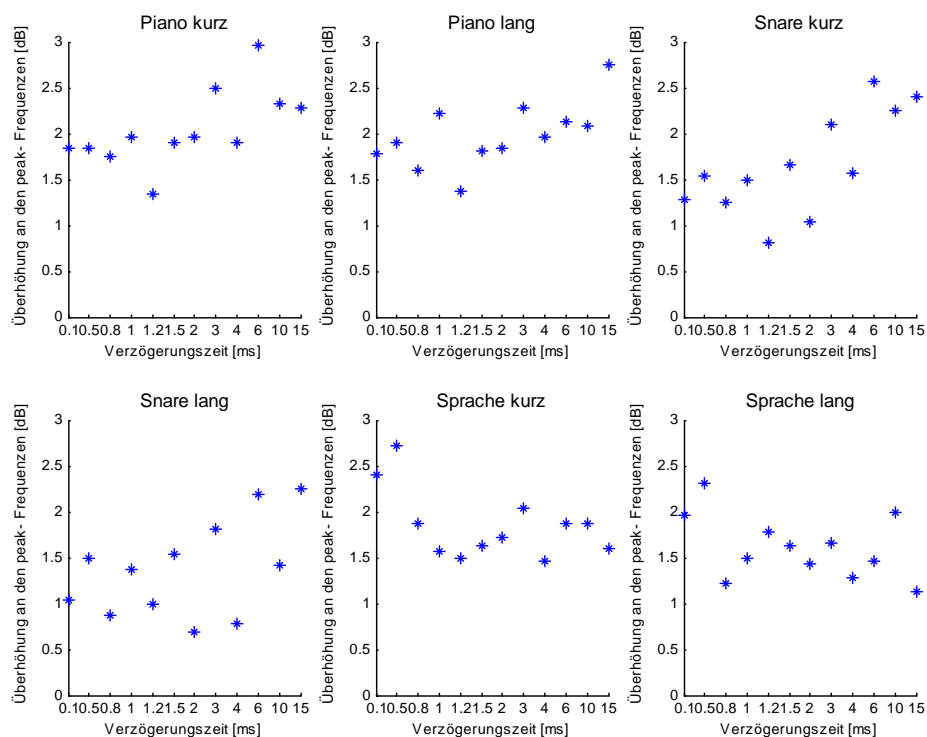


Abbildung 15: Ergebnisse des ersten Vorversuches. Dargestellt sind die peak – Überhöhungen an der für die jeweilige Verzögerungszeit ermittelten Hörschwelle.

wahrgenommen wurde als Sprache und Piano. Außerdem ist allgemein eine höhere Sensibilität der Versuchspersonen für zeitliche Verschiebungen von 0,8 bis 1,2 ms zu bemerken. Die geringe Stichprobengröße sowie die Aufteilung des Wertebereichs der zu testenden Primärvariable auf die zwei Gruppen von Probanden sind wahrscheinliche Ursachen für die erkennbaren Sprünge im Kurvenverlauf. In Tabelle 10 auf S. 63 sind die Ergebnisse detailliert aufgeführt. Die in diesem Vorversuch ermittelten Hörschwellen dienten als Grundlage für die Initialwerte des Hauptversuches.

4.5.2 Zweiter Vorversuch

Stimuli

Ausgehend vom ersten Vorversuch wurden lediglich Stimuli der Kategorie „Snare“ mit Längen von einer, zwei, drei und vier Sekunden getestet. Um den Versuch so kurz wie möglich zu halten, wurden die zu testenden *delay times* auf 0,8 und 1,2 ms reduziert, da dieser Bereich im ersten Vorversuch die besten Erkennungsraten lieferte.

Versuchspersonen

Als Versuchspersonen stellten sich eine Studentin, vier Studenten und ein wissenschaftlicher Mitarbeiter der Technischen Universität Berlin zur Verfügung. Die Versuchspersonen nahmen freiwillig an dem Experiment teil und erhielten keine Bezahlung.

Versuchsaufbau

Der Versuchsaufbau war identisch mit dem des ersten Vorversuches.

Versuchsablauf

Nach einer Instruktion begann der Vorversuch mit einer Trainingsphase während derer der Proband Extrembeispiele der später zu detektierenden verzerrten Stimuli in der Darbietungsform des späteren Versuches zu hören bekam. Die Trainingsphase umfasste 15 Durchläufe.

Der eigentliche Versuch wurde als *Three-Alternative-Forced-Choice* Test durchgeführt. Da in den Interviews des ersten Vorversuches sehr deutlich

geworden war, dass die Versuchspersonen nicht nur das korrekte Paar, sondern auch explizit die Position des verzerrten Stimulus zu bestimmen in der Lage waren, konnte das für den vorliegenden Untersuchungsgegenstand optimale Testverfahren angewendet werden. Dieses ist nun aber nach Leek eben das 3AFC-Verfahren.¹⁰² Vor allem die geringere Ratewahrscheinlichkeit von 33% gegenüber 50% beim 2AFC-Test sprechen für 3AFC. Die Bedienung unterschied sich nicht vom ersten Vorversuch. Die Probanden hörten nun drei zeitlich durch Pausen von 0,3 Sekunden separierte Samples, von denen sich eines von den beiden anderen unterschied. Die Detektionsentscheidung wurde wieder mit Klick auf den entsprechenden Button manifestiert. Während des Experimentes wurden nach jeder erfolgreich beendeten Testreihe (d. h. nach jeder vollständig getesteten Verzögerungszeit) automatisch Pausen annonciert. Dabei konnte die Versuchsperson selbst über die Dauer der Pause entscheiden, wurde jedoch vom Versuchsleiter angehalten, sich mindestens zwei längere Pausen (5-10 Minuten) und drei kürzere Pausen (ca. 2 Minuten) zu gestatten.

Ergebnisse und Interpretation

Der Versuch resultierte im Ergebnis in einer erkennbaren Tendenz von drei Sekunden Stimulilänge hinsichtlich der besten Erkennungsraten. Um im Hauptversuch tatsächlich unter bestmöglichen Bedingungen eine Hörschwelle zu ermitteln, wurde folgerichtig die Länge der dort zu testenden Stimuli auf drei Sekunden festgelegt. Die Versuchsergebnisse sind in Abbildung 16, S. 41, und in Tabelle 11, S. 63, dargestellt. Aus den im Nachhinein geführten Interviews wurden zusätzlich für den Hauptversuch noch diverse Anregungen der Versuchspersonen umgesetzt. So wurde eine Bedienung des Versuches per Tastatur ermöglicht, die Farbgebung der Oberfläche nochmals verändert, sowie die Möglichkeit der optischen Orientierung durch eine farbliche Aufhellung des zum gerade präsentierten Stimulus gehörenden Buttons eingeräumt.

¹⁰² Vgl. Leek (2001), S. 1290.

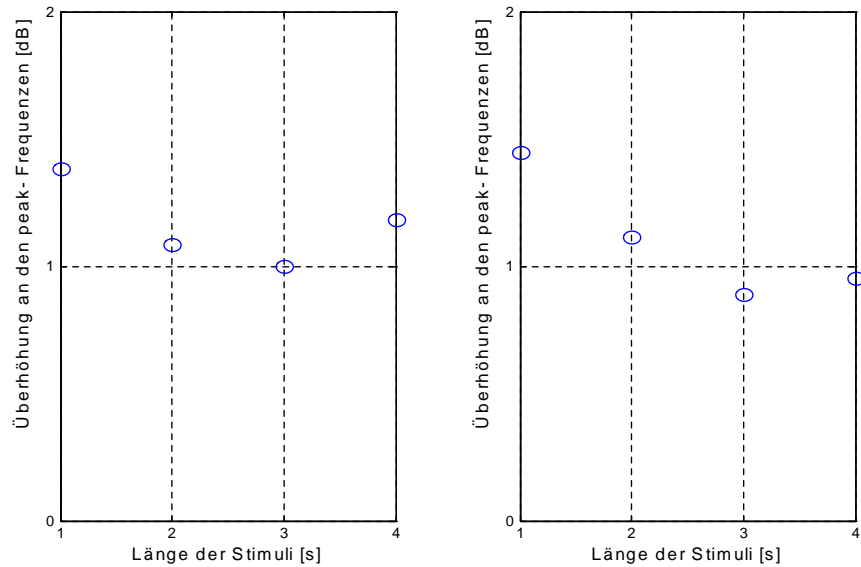


Abbildung 16: Ergebnisse des zweiten Vorversuches; links: Die Hörschwelle bei 0,8 ms Verzögerung; rechts: Die Hörschwelle bei 1,2 ms Verzögerung.

4.6 Hauptversuch

4.6.1 Stimuli

Im Hinblick auf die Ergebnisse der Vorversuche wurden Stimuli der Kategorie „Piano“, „Speech“ und „Snare“ mit einer Länge von jeweils drei Sekunden präsentiert. Es wurden die Verzögerungszeiten $T = 0,1; 0,5; 0,8; 1; 1,5; 3; 6$ und 15 ms getestet. Der Pool der Verzögerungszeiten wurde im Vergleich zum ersten Vorversuch reduziert, um die Belastung für die Versuchspersonen zu verringern. Die getesteten Verzögerungszeiten konzentrierten sich nun noch mehr um den (nach Erkenntnissen aus dem ersten Vorversuch) Bereich mit der niedrigsten zu erwartenden Hörschwelle bei 1 ms Verzögerungszeit, um einen möglichst repräsentativen Verlauf der Wahrnehmungsschwelle dokumentieren zu können.

4.6.2 Versuchspersonen

Als Versuchspersonen stellten sich drei Tonmeister, zehn TonmeisterstudentInnen der UdK Berlin, neunzehn Studierende der Kommunikationswissenschaften der TU Berlin¹⁰³ sowie zwei Musiker und zwei Studierende anderer Fachrichtungen zur Verfügung. Die Probanden waren 20 bis 62 Jahre, im Durchschnitt 28 Jahre alt. Unter den Versuchspersonen waren drei Frauen und dreiunddreißig Männer. Ihr eigenes Hörvermögen schätzten die Probanden auf einer Skala von 1 bis 6 (1: sehr gut, 6: sehr schlecht) im Bereich 1 bis 3¹⁰⁴ ein. Ein Proband, der auf diese Frage mit „fünf“ geantwortet hatte, brach den Versuch vorzeitig ab. Die bis dahin von ihm erhobenen Daten gingen nicht mit in die Untersuchung ein. Die Probanden erhielten keine Bezahlung.

4.6.3 Versuchsaufbau

Der Hörversuch lief automatisch softwaregesteuert ab. Die Stimuli wurden den Versuchspersonen über die elektrostatischen Kopfhörer SR-202 der Firma STAX, angeschlossen an den zugehörigen Kopfhörervorverstärker SRM 212, präsentiert. Als Interface diente ein MOTU 828 Firewire Interface. Die im Mono-Format vorliegenden Samples wurden binaural präsentiert. Als Versuchsraum diente das kleine elektronische Studio der TU Berlin. Um die Lüftergeräusche des Laptops als Störvariable auszuschalten, wurden nach dem Durchführen von fünf Experimenten ein externer Monitor, eine externe Maus und ein externes Keyboard angeschlossen. Die Daten der ersten fünf Versuchspersonen wiesen keine signifikanten Auffälligkeiten auf, so dass sie weiterhin in die Untersuchung eingingen.

4.6.4 Versuchsablauf

Nach einer Instruktion¹⁰⁵, welche sich in den Grundzügen nach Empfehlungen der Nordtest Gesellschaft zur Durchführung von

¹⁰³ Davon neun mit Neben- oder Hauptfach Musikwissenschaften

¹⁰⁴ Durchschnittlich mit 2,1

¹⁰⁵ Siehe Anhang B, S. 66.

Hörversuchen¹⁰⁶ richtete und dem Ausfüllen des Fragebogens¹⁰⁷ begann der Vorversuch mit einer Trainingsphase. Diese entsprach im Ablauf der eigentlichen Testphase, allerdings ohne die Speicherung der Antworten der Versuchsperson. Als Stimuli wurden Extrembeispiele aus der auch im eigentlichen Versuch zu detektierenden Kategorie (Piano, Snare oder Sprache) in zufälliger Reihenfolge ausgewählt. Die Trainingsphase umfasste zehn Durchläufe. Dabei erhielt die Versuchsperson nach jeder getroffenen Entscheidung ein optisches Feedback in Form eines roten oder grünen Farbfeldes (falsche Detektion: rot, richtige Detektion: grün). Die Trainingsphase diente mehr dem Einstimmen der Versuchspersonen auf den Ablauf des Experimentes denn der Sensibilisierung für die Art der später im Experiment zu detektierenden Stimuli. Letzteres war eher zweitrangig, da jeder Anfang einer neuen Testreihe durch den hohen Initialwert einen Trainingseffekt mit sich brachte.¹⁰⁸

4.6.5 Ergebnisse

Die Versuchsergebnisse werden nacheinander für jede einzelne Kategorie tabellarisch und grafisch dargestellt.¹⁰⁹ Die Analyse und Interpretation der Ergebnisse erfolgen in Kapitel 5. Die Bedeutung der einzelnen Diagramme wird im folgenden Abschnitt exemplarisch anhand der Kategorie Piano erläutert und gilt auch für die anderen beiden Kategorien.

Piano

Tabelle 4, S. 44, zeigt die Ergebnisse des Hörversuches für die Kategorie Piano. Dabei wird die Faktorstufe, die im Versuch als Schwelle ermittelt worden war, zahlenmäßig nicht extra aufgeführt, sondern direkt die sich ergebende *peak*- Anhebung an der Hörschwelle aufgezeigt. Die Ermittlung der Werte erfolgte durch Bildung des arithmetischen Mittels aller zwölf

¹⁰⁶ Vgl. Nordtest, (2001), S. 9f.

¹⁰⁷ Siehe Anhang D, S. 67.

¹⁰⁸ Siehe Abschnitt 4.4.2.

¹⁰⁹ Der Autor hält eine ausführliche grafische Darstellung bereits an dieser Stelle für sinnvoll, da dem Leser ansonsten wichtige Bezugspunkte für die spätere Auswertung fehlen würden.

Versuchspersonen dieser Testgruppe. Die Einzelergebnisse sind in Anhang A, S. 64 f., aufgeführt.

Verzögerungszeit [ms]	0,1	0,5	0,8	1	1,5	3	6	15
peak- Anhebung an der ermittelten Hörschwelle [dB]	2,1	1,9	1,7	1,9	1,8	2,4	2,2	2,3
Standardabweichung	1,32	0,52	0,46	0,56	0,54	0,54	0,40	0,68

Tabelle 4: Ermittelte Hörschwelle in der Kategorie Piano.

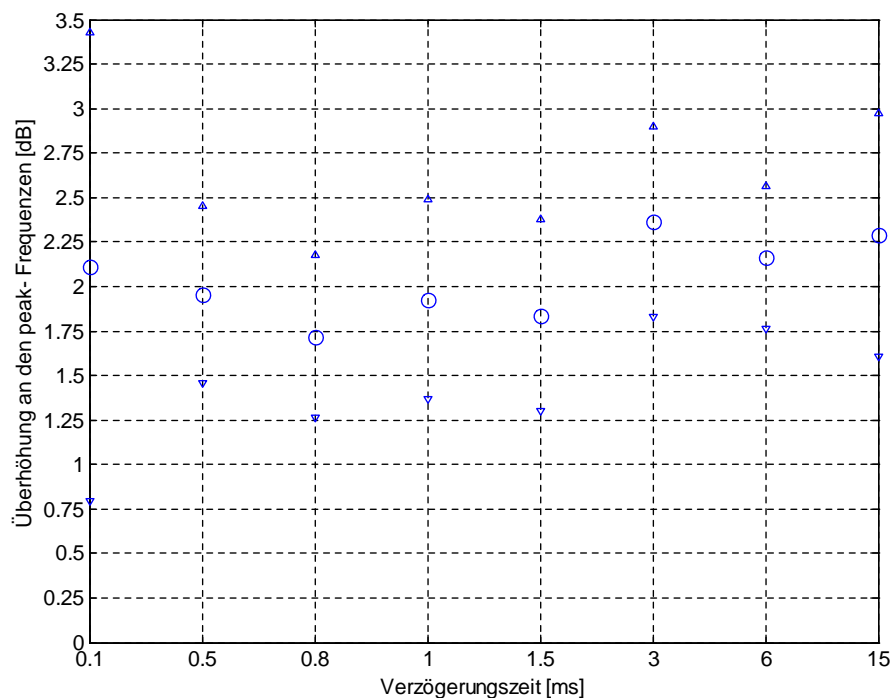


Abbildung 17: Diagramm der ermittelten Hörschwelle für die Kategorie Piano. Die Standardabweichung ist mittels der Dreiecke dargestellt.

In Abbildung 17 werden die Versuchsergebnisse in einem Diagramm zusammengefasst.

Da es durch Extremwerte („Ausreißer“)¹¹⁰ bei geringen Stichprobengrößen zu starken Verschiebungen des arithmetischen Mittelwertes kommen kann, wird für jeden Datensatz auch der Median berechnet, wie in Abbildung 18, S. 45,

¹¹⁰ Ausreißer (engl.: *outliers*) können durch Messfehler, Unkonzentrationen der Versuchsperson oder anderes Bias entstehen.

im Rahmen eines Boxplots zum Vergleich mit den Mittelwerten dargestellt. Diese Plots zeigen neben dem Median der Werte aller Versuchspersonen bei jeweils einer Verzögerungszeit auch die *interquartile range*¹¹¹ (IQR). Des Weiteren ist die Verteilung der Daten außerhalb der IQR durch horizontale Linien (*whiskers*) markiert. Die Länge der *whiskers* beträgt maximal das Anderthalbfache der IQR. Als „Ausreißer“ werden diejenigen Werte definiert, die um mehr als das anderthalbfache der IQR vom Median entfernt und damit auch außerhalb der *whiskers* liegen.¹¹²

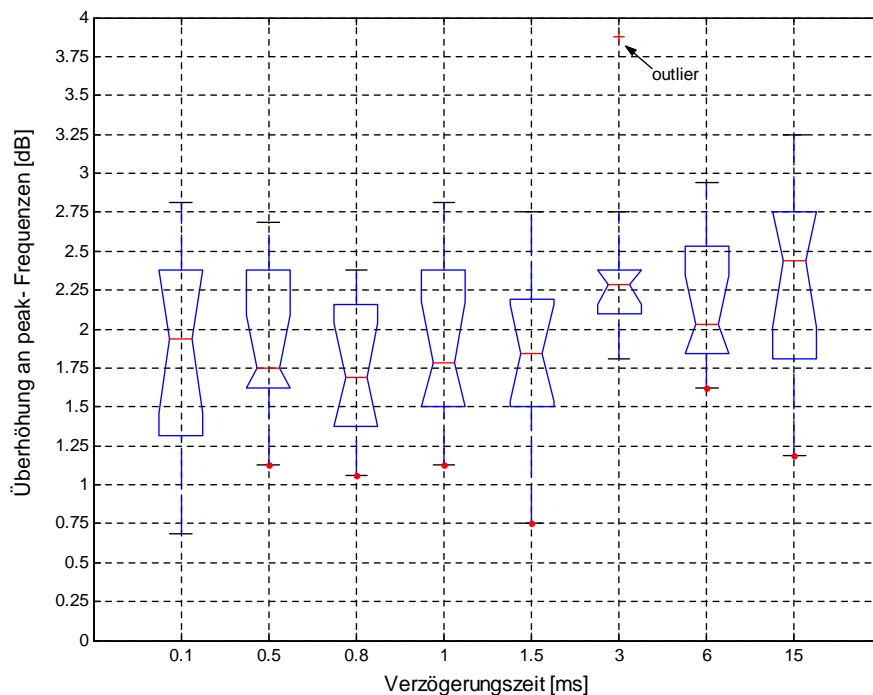


Abbildung 18: Boxplot der Werte an der Hörschwelle für die Kategorie Piano.

In Abschnitt 2.1.2 wurde auf die Bedeutung des Verhältnisses der Pegel der überlagernden Signale eingegangen. Tabelle 5 und Abbildung 19, S. 46,

¹¹¹ Die *interquartile range* umfasst alle Daten zwischen dem 25. und dem 75. Perzentil des Datensatzes.

¹¹² Im Fall der Kategorie Piano liegen zwei Extremwerte vor: bei 3 ms Verzögerungszeit und bei 0,1 ms. Der Wert des zweitgenannten *outliers* beträgt 5,8 dB und ist deshalb nicht mehr in der Grafik (Abbildung 18) enthalten, da die y-Achse zur besseren Darstellung der Werte auf 0 bis 4 dB begrenzt wurde.

stellen den Wert der Pegeldifferenzen der Signale an der Hörschwelle nach Gleichung 9 über der jeweiligen Verzögerungszeit dar.

Verzögerungszeit [ms]	0,1	0,5	0,8	1	1,5	3	6	15
Pegeldifferenz der überlagerten Signale an der Hörschwelle	-11,2	-12,1	-13,2	-12,1	-12,6	-10,1	-11	-10,4

Tabelle 5: Pegeldifferenzen an der ermittelten Hörschwelle für die Kategorie Piano.

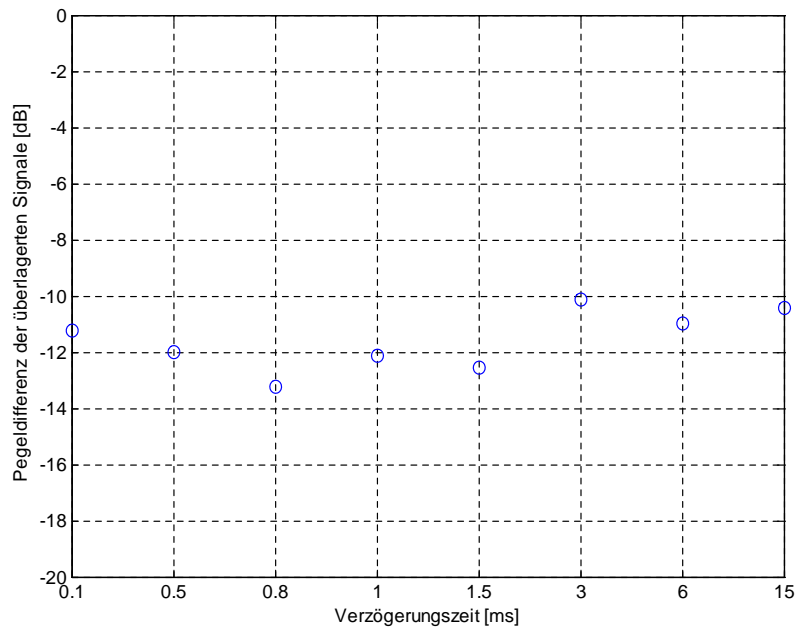


Abbildung 19: Diagramm der Pegeldifferenzen an der ermittelten Hörschwelle für die Kategorie Piano.

Snare

Tabelle 6 fasst die Ergebnisse des Hörversuches für die Kategorie Snare zusammen.

Verzögerungszeit [ms]	0,1	0,5	0,8	1	1,5	3	6	15
<i>peak</i> -Anhebung an der ermittelten Hörschwelle [dB]	1,82	1,07	1,01	1,14	1,21	1,11	1,82	2,36
Standardabweichung	0,74	0,51	0,43	0,64	0,40	0,44	0,37	0,86

Tabelle 6: Ermittelte Hörschwelle in der Kategorie Snare.

Abbildung 20 zeigt den ermittelten Verlauf der Hörschwelle.

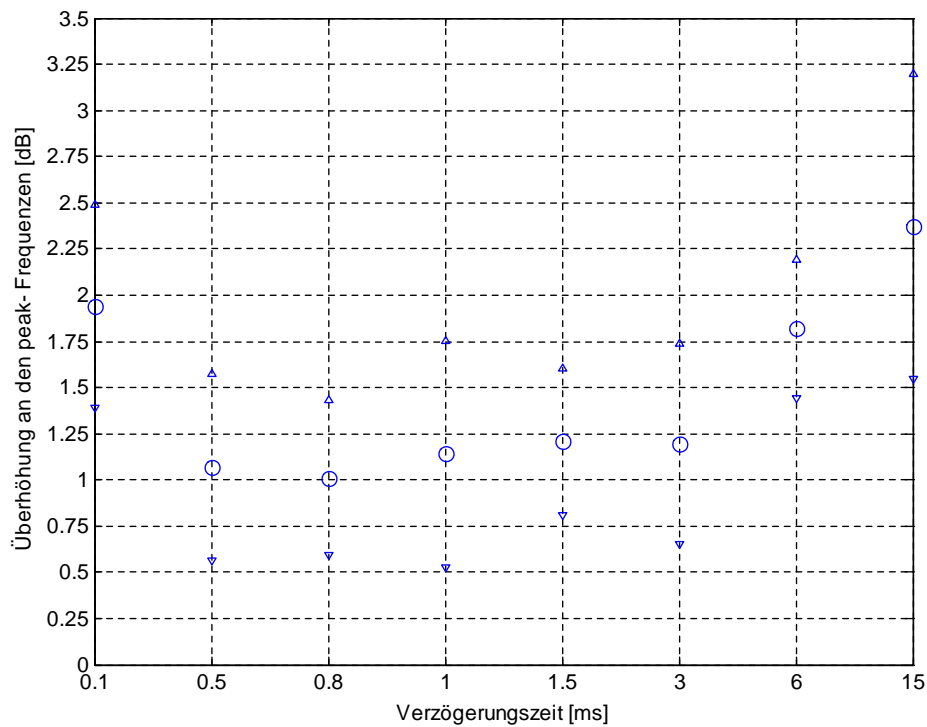


Abbildung 20: Diagramm der ermittelten Hörschwelle für die Kategorie Snare. Die Standardabweichung ist mittels der Dreiecke dargestellt.

Abbildung 21 stellt die Daten als Boxplot dar.

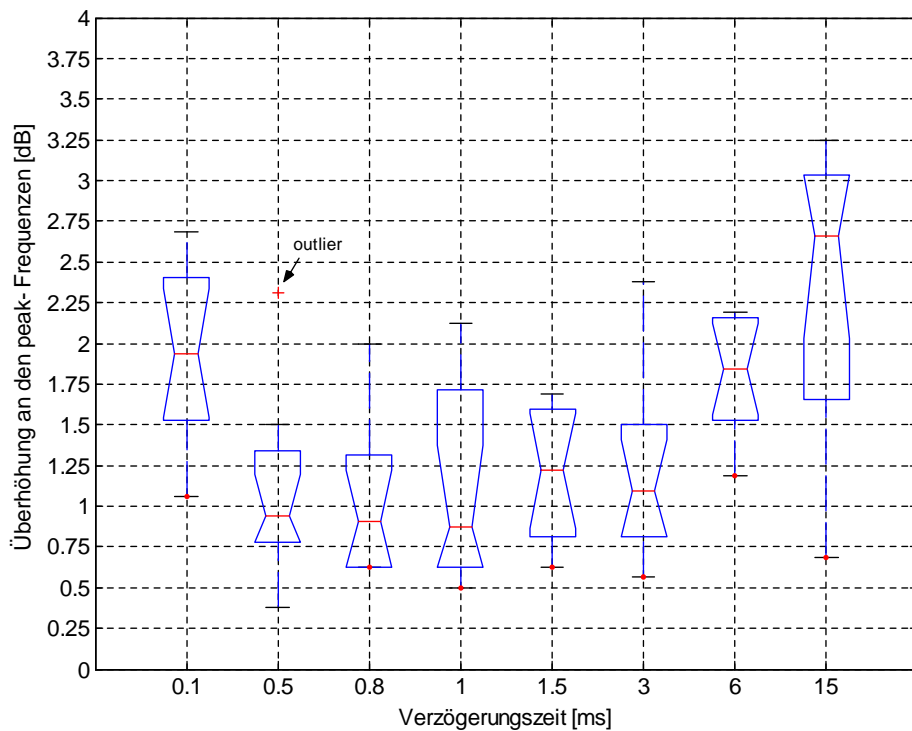


Abbildung 21: Boxplot der Werte an der Hörschwelle für die Kategorie Snare.

Tabelle 7 und Abbildung 22 veranschaulichen den Verlauf der Pegeldifferenzen der Signale an der Hörschwelle für die jeweilige Verzögerungszeit.

Verzögerungszeit [ms]	0,1	0,5	0,8	1	1,5	3	6	15
Pegeldifferenz der überlagerten Signale an der Hörschwelle	-12,7	-17,7	-18,2	-17,1	-16,5	-17,3	-12,7	-10,1

Tabelle 7: Pegeldifferenzen an der ermittelten Hörschwelle für die Kategorie Snare.

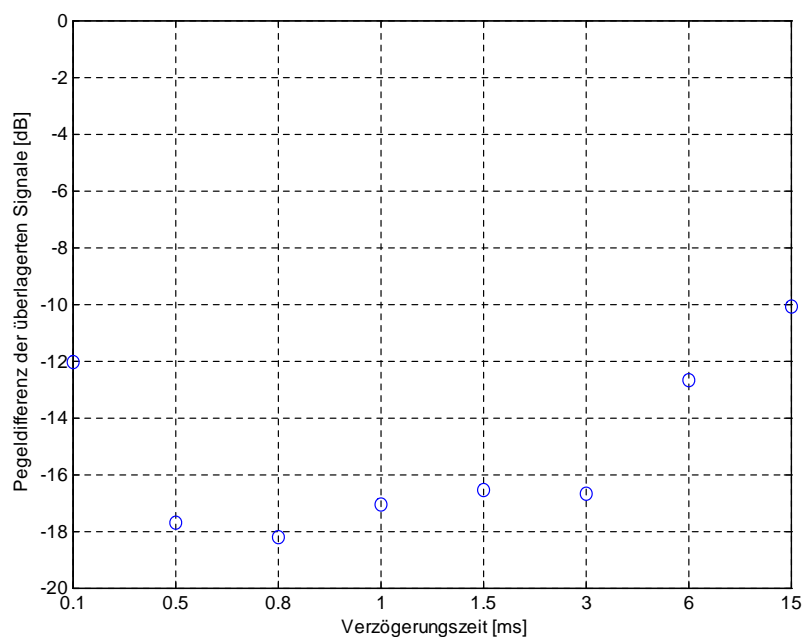


Abbildung 22: Diagramm der Pegeldifferenzen an der ermittelten Hörschwelle für die Kategorie Snare.

Sprache

Tabelle 8 fasst die Ergebnisse des Hörversuches für die Kategorie Sprache zusammen.

Verzögerungszeit [ms]	0,1	0,5	0,8	1	1,5	3	6	15
<i>peak</i> -Anhebung an der ermittelten Hörschwelle [dB]	2,80	2,10	1,94	1,88	1,80	1,30	1,40	1,20
Standardabweichung	0,78	0,76	0,86	0,84	0,60	0,26	0,44	0,44

Tabelle 8: Ermittelte Hörschwelle in der Kategorie Sprache.

Abbildung 23 zeigt den ermittelten Verlauf der Hörschwelle.

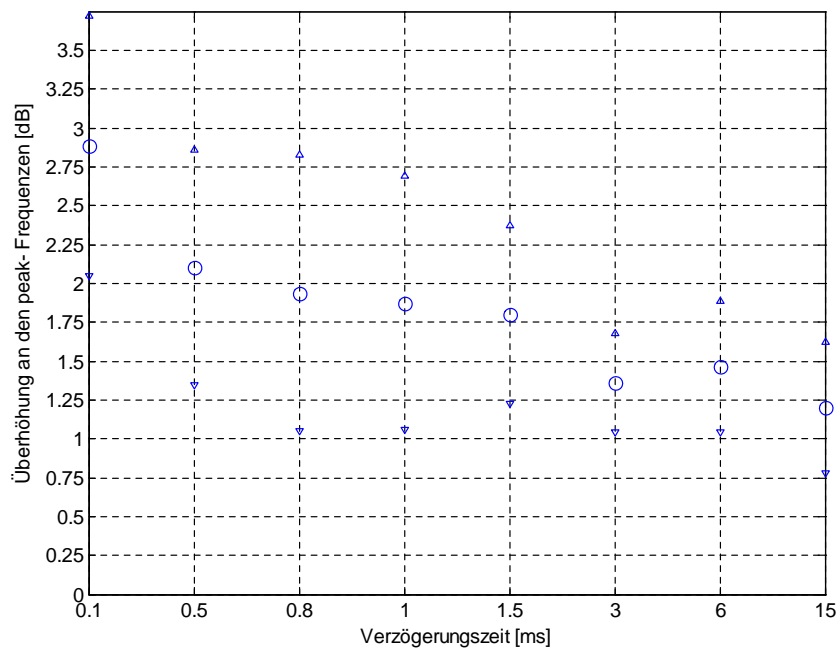


Abbildung 23: Diagramm der ermittelten Hörschwelle für die Kategorie Sprache. Die Standardabweichung ist mittels der Dreiecke dargestellt.

Abbildung 24 stellt die Daten als Boxplot dar.

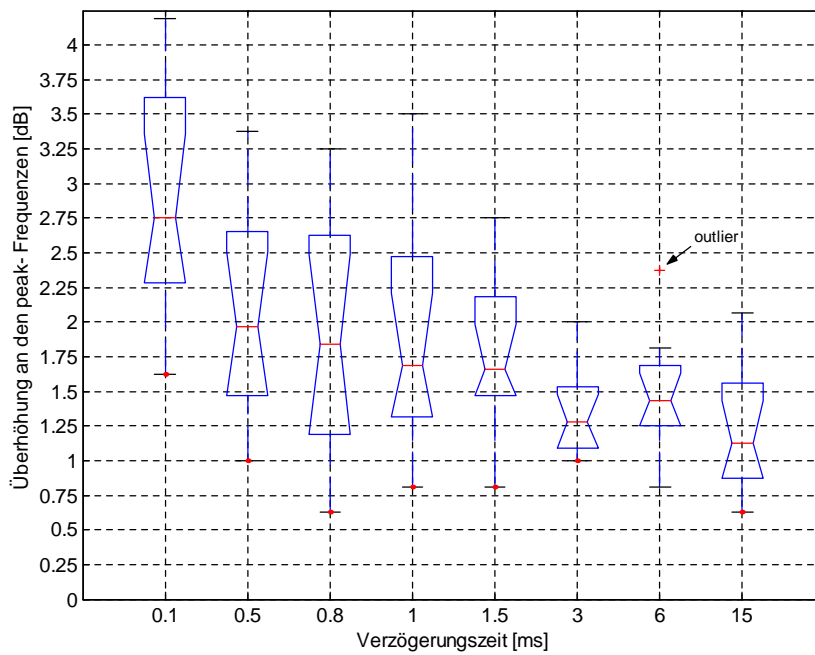


Abbildung 24: Boxplot der Werte an der Hörschwelle für die Kategorie Sprache.

Verzögerungszeit [ms]	0,1	0,5	0,8	1	1,5	3	6	15
Pegeldifferenz der überlagerten Signale an der Hörschwelle	-8,4	-11,2	-12,0	-12,4	-12,8	-15,9	-15,1	-16,6

Tabelle 9: Pegeldifferenzen an der ermittelten Hörschwelle für die Kategorie Sprache.

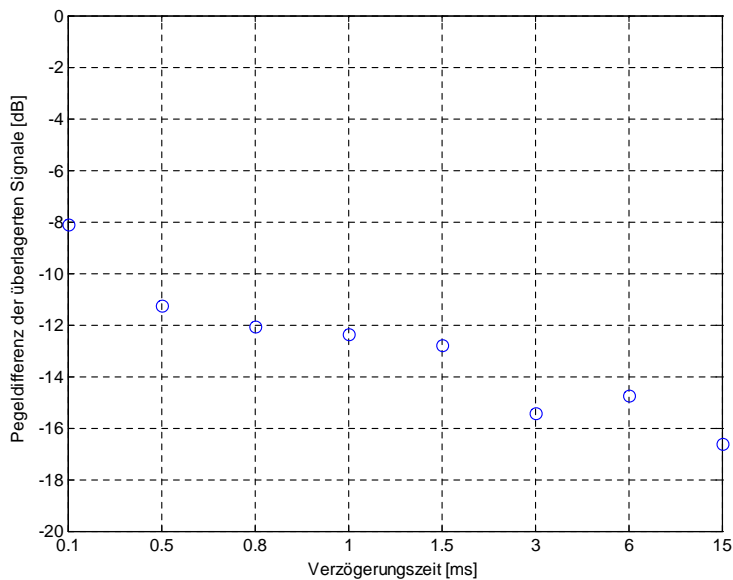


Abbildung 25: Diagramm der Pegeldifferenzen an der ermittelten Hörschwelle für die Kategorie Sprache.

Tabelle 9 und Abbildung 25 veranschaulichen den Verlauf der Pegeldifferenzen der Signale an der Hörschwelle für die jeweilige Verzögerungszeit. Abbildung 26 zeigt zum Vergleich nochmals die Versuchsergebnisse für alle drei Kategorien zusammen.

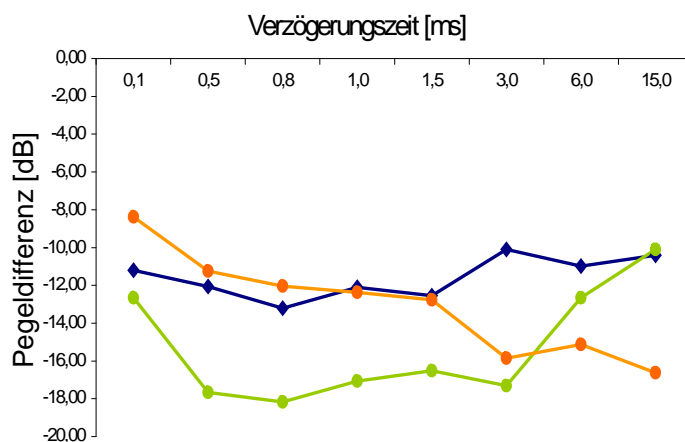


Abbildung 26: Pegeldifferenzen der überlagernden Signale an der ermittelten Hörschwelle für alle drei Kategorien (blau: Piano, orange: Sprache, grün: Snare).

5 Auswertung

5.1 Interpretation der Ergebnisse

5.1.1 Piano

Die Hörschwelle für die Kategorie Piano weist ein lokales Minimum bei einer Verzögerungszeit von 0,8 ms auf. An dieser Stelle wird ein im Vergleich zum unverzerrten Signal um 1,7 dB erhöhter *peak* gehört. Der erste *peak* des verzerrten Signals tritt bei dieser Verzögerungszeit bei 1250 Hz auf. Zu den oberen und unteren Begrenzungen des Messbereichs steigt die Kurve leicht an, d. h. Verzerrungen, welche durch längere oder kürzere Verzögerungszeiten als 0,8 ms entstanden sind, werden schlechter gehört. Eine deutliche Ausnahme bildet hier ein *peak* bei 3 ms Verzögerung, der auch dann nachweisbar ist, wenn der in Abb. 18, S. 45, dargestellte *outlier* unberücksichtigt bleibt. Dieser Wert entspricht einem spektralen *peak* bei 333 Hz. Berücksichtigt man nun die Tatsache, dass das Originalsignal bereits einen *peak* bei 334 Hz aufweist (siehe Abbildung 10, S. 28), so besteht Anlass zu der Vermutung, dass eine bloße Überhöhung eines bereits vorhandenen spektralen *peaks* durch einen Kammfiltereffekt schlechter detektiert wird als eine Anhebung eines Frequenzbereiches, welcher im Originalsignal kein Maximum aufweist. Diese These könnte in weiteren Untersuchungen verifiziert werden. Für die Praxis der Schallaufnahme ist sicherlich die Pegeldifferenz der überlagerten Signale an der ermittelten Hörschwelle am bedeutendsten. So können laut den Versuchsergebnissen Pegeldifferenzen von bis zu 13,2 dB zwischen superpositionierten Signalen gehört werden. Konkret auf die Praxis angewendet bedeutet dieses Versuchsergebnis, dass die Reflexion eines Signals bei einer Schallwegdifferenz von 27 Zentimetern selbst dann noch eine hörbare kammfilterartige Verzerrung verursachen würde, wenn sie bereits um 13,2 dB abgeschwächt ist.

Ein interessanter Aspekt ergab sich aus dem offenen Feedback der Versuchspersonen nach Durchführung des Experiments. So gab ein Großteil der Probanden an, sich bei der Detektion der Verzerrung in der Kategorie Piano auf die Hörbarkeit eines Tiefpassfilters gestützt zu haben. In der Tat

entsprechen Kammfilter mit kleinen Verzögerungszeiten dann einem Tiefpassfilter, wenn lediglich die erste *peak*- und die zweite *notch*-Frequenz samt dazugehöriger Flanke in den Hörbereich fallen. Der *cue* für die bewusste Detektion ist jedoch nach der o. g. Aussage trotz alledem auch bei längeren Verzögerungszeiten eine Höhenabsenkung des Signals.

5.1.2 Snare

Auch der Verlauf der Hörschwelle in der Kategorie Snare verzeichnet ein Minimum an der Stelle der Verzögerungszeit von 0,8 ms. An dieser Stelle wird eine *peak*-Überhöhung des verzerrten Signals von 1 dB gehört.¹¹³ Zu den oberen und unteren Begrenzungen des Wertebereichs hin verläuft der Anstieg der Hörschwelle jedoch steiler als in der Kategorie Piano, so dass die Vermutung nahe liegt, tatsächlich ein globales Minimum detektiert zu haben. Der Pegelunterschied zwischen den summierten Signalen beträgt an der ermittelten Hörschwelle bis zu 18,2 dB. Die Probanden gaben als Feedback an, sich bei der Detektion des verzerrten Signals nach der Tonhaltigkeit des Stimulus gerichtet zu haben. Die Verzerrung war als Tonhöhe des ansonsten rauschhaften Signals zu hören. Je nach Verzögerungszeit änderte sich die durch die Versuchspersonen wahrgenommene Tonhöhe.

5.1.3 Sprache

Die ermittelte Hörschwelle dieser Kategorie unterscheidet sich im Verlauf erheblich von denen der vorangegangenen Kategorien. Das Minimum innerhalb des Messbereichs liegt bei 15 ms zeitlicher Verzögerung des superpositionierten Signals. Allerdings ist der Verlauf der Hörschwellenkurve nach visueller Approximation fast linear und in Richtung längerer Verzögerungszeiten fallend, so dass jede Aussage über den weiteren Verlauf der Hörschwelle rein spekulativ wäre. Allerdings ist gerade im Hinblick auf die konkrete Schallaufnahmesituation das Versuchsergebnis interessant. Die für eine Sprachaufnahme typischen Schallwegdifferenzen, etwa durch die

¹¹³ Nimmt man den Median als Berechnungsgrundlage, so liegt das Minimum bei 1 ms.

Reflexionen vom Sprechertisch oder vom Textblatt, betragen um die 50 cm. Dies entspricht einer Laufzeitverzögerung von ca. 1,5 ms. Bei dieser Verzögerungszeit wird selbst bei 12,8 dB Pegeldifferenz zwischen Direktschall und Reflexion noch der Kammfiltereffekt gehört. Laut dem Feedback der Versuchspersonen wurden die verzerrten Stimuli anhand der Hörbarkeit einer Räumlichkeit identifiziert.

5.2 Diskussion

Die Versuchsergebnisse lassen darauf schließen, dass kammfilterartige Verzerrungen dann am besten detektiert werden, wenn es zur Entstehung eines Residualtones im empfindlichsten Hörbereich kommt. Anlass zu diesem Schluss geben die deutlich besseren Versuchsergebnisse in der Kategorie Snare bei Verzögerungszeiten von 0,5 bis 3 ms, die nach der Superposition zu Residualtönen von 333 Hz bis 2 kHz führen. Es ist zu vermuten, dass die sehr gute Detektierbarkeit in der Kategorie Sprache bei längeren Verzögerungszeiten mit der zunehmenden Wahrnehmbarkeit einer unnatürlichen Räumlichkeit zusammenhängt.

Der Vergleich der hier ermittelten Hörschwelle mit den aus früheren Untersuchungen vorliegenden Werten¹¹⁴ kann auf Grund der fehlenden Informationen bezüglich des Zustandekommens dieser Werte und der teilweise verschiedenen Versuchsparameter lediglich grob quantitativ erfolgen.

Die von Müller postulierten 12 dB Pegeldifferenz der überlagernden Signale an der Hörschwelle bei Sprache¹¹⁵ finden sich im hier vorliegenden Untersuchungsergebnis bei Verzögerungszeiten von 0,5 bis 1,5 ms Verzögerungszeit in Pegeldifferenzen von 11,2 bis 12,8 dB Pegeldifferenz wieder.¹¹⁶

¹¹⁴ Vgl. Abschnitt 2.3.

¹¹⁵ Vgl. Müller (1968), S. E61 ff.

¹¹⁶ Siehe Tabelle 9, S. 50.

Das von Kuhl untersuchte rosa Rauschen ist spektral eng mit dem hier untersuchten Snaresignal verwandt. Da in der hier vorliegenden Untersuchung 2 ms Verzögerungszeit nicht getestet wurden, könnte Kuhls 17 dB Pegelunterschied¹¹⁷ lediglich mit den Werten von 1,5 ms (16,5 dB) und 3 ms (17,3 dB) verglichen werden. Die Werte für Verzögerungszeiten von 5 bis 15 ms liegen bei Kuhl mit konstanten 13 dB Pegeldifferenz etwas höher als die Ergebnisse dieser Untersuchung, welche eher eine stetige Abnahme der Detektierbarkeit mit zunehmender Verzögerungszeit nahe legen.

Die Faustregel von Dickreiter, 10 dB Pegeldifferenz einzuhalten, damit der Kammfiltereffekt unbedeutend bleibt¹¹⁸, kann nicht bestätigt werden. So wurden beispielsweise im Fall der Snare Kammfiltereffekte, die durch die Überlagerung zweier Signale mit 18,2 dB Pegelunterschied entstanden sind, gehört.

Burroughs' 3:1-Regel zur Vermeidung von hörbaren Kammfiltereffekten sollte, betrachtet man die vorliegenden Versuchsergebnisse, ebenso mit Vorsicht angewendet werden, da die Ausprägung des Kammfiltereffektes vor allem von der Pegeldifferenz der superpositionierten Signale abhängt. Diese kann jedoch in einem nach den vorliegenden Untersuchungsergebnissen kritischen Bereich liegen, wenn die 3:1- Regel eingehalten wird. Unabhängig davon ist eine ausreichende räumliche Trennung von Haupt- und Stützmikrofonen in jedem Fall zu empfehlen.

5.3 Kritik

Die vorliegende Untersuchung wurde nach einer Methode, welche nach dem heutigen Stand der Forschung das effektivste Verfahren zur Ermittlung einer Hörschwelle darstellt, geplant und durchgeführt. Durch die vollautomatische Steuerung des Versuchs über eine explizit für dieses Experiment gestaltete

¹¹⁷ Vgl. Kuhl (1969), S. 206.

¹¹⁸ Vgl. Dickreiter (2003), S. 157.

Software war die Konstanz des Versuchsablaufs gewährleistet. Die softwaregesteuerte Generierung der Stimuli und die Anpassung der Lautheiten stellen ein weiteres Indiz für die interne Validität der Untersuchung dar. Durch die Durchführung zweier Vorversuche konnten die Bedienbarkeit der Benutzeroberfläche, sowie diverse andere Versuchsparameter optimiert werden.

Das größte Manko der vorliegenden Untersuchung ist die geringe Größe der Stichprobe. Mit 36 Personen, die in drei Gruppen aufgeteilt wurden, konnte nur ein kleiner Teil einer Population erfasst werden. Die Generalisierbarkeit der ermittelten Werte muss daher in Frage gestellt werden. Da die Probanden Expertenhörer waren, lässt das Ergebnis nicht auf die Hörschwelle der Grundgesamtheit aller Hörer schließen. Ein weiterer Kritikpunkt ist die Kürze der Versuchsreihen. Eine größere Anzahl *trials* wäre in jedem Fall wünschenswert gewesen. Andererseits wäre hierdurch wiederum die Belastung für die Versuchspersonen viel zu groß gewesen. Bei adaptiven Testverfahren kann es zu Anpassungseffekten kommen. Diese hätten durch das *interleaving* zweier unabhängiger Versuchsreihen vermieden werden können. Dies hätte allerdings zu einer Verdopplung der Versuchsdauer geführt.

Somit ist zu konstatieren, dass eine Beschränkung der Untersuchung auf eine Kategorie zu testender Quellsignale sinnvoll gewesen wäre, um die Validität zu gewährleisten.

Ein weiterer Kritikpunkt ist die Auswahl der Quellsignale. Am deutlichsten wird dies anhand der Kategorie Piano. Es ist in Anbetracht der Versuchsergebnisse fraglich, ob eine kurze Phrase repräsentativ für das Instrument Klavier ist. Beispielsweise sind klangliche Besonderheiten des Fortissimo- oder Pianissimospiels nicht in dieser Untersuchung repräsentiert. Ebenso ist eine Generalisierbarkeit des hier untersuchten eingesprochenen Satzes als repräsentativ für „Sprache“ in Frage zu stellen, da allein die spektralen Unterschiede zwischen männlicher und weiblicher Stimme gravierend sind. Eine Untersuchung mit standardisierten Signalen (Sinus-

oder Rauschsignale) ist in dieser Beziehung sicherlich unverfänglicher. Dies kollidierte in dieser Arbeit jedoch mit dem Praxisbezug, der hier im Vordergrund stehen sollte.

Die Untersuchung beschränkt sich auf wenige Variablen zur Kategorisierung der Verzerrung. Parameter wie der Diffusfeldanteil oder Verdeckungs- und Adaptionprozesse werden ausgespart und bedürfen weiterer Untersuchung.

6 Zusammenfassung

Kammfilterartige lineare Verzerrungen können im Verlauf einer Schallaufnahme an diversen Positionen auftreten und dabei störende spektrale Manipulationen hervorrufen. In komprimierter Form wurde ein Überblick über die zu diesem Thema bereits publizierten Arbeiten gegeben. Aus der Evaluation dieser Arbeiten resultierte die Planung und Durchführung von Hörversuchen, welche die Ermittlung einer Hörschwelle für diese Art der Verzerrung zum Gegenstand hatten. Hierbei stand die Aufzeichnung und spätere Auswertung der Reaktionen der Probanden auf die Variation der zwei wichtigsten physikalischen Parameter, welche die Ausprägung des Kammfiltereffektes beeinflussen - die Verzögerungszeit und die Pegeldifferenz eines superpositionierten Signals im Vergleich zum Ursprungssignal - im Vordergrund. Um einen Überblick über die Variabilität der Wahrnehmung von Kammfiltereffekten zu geben, wurden drei Kategorien an Quellsignalen zur Generierung der Stimuli benutzt: eine Klavierphrase, ein Snare-Roll und ein gesprochener deutscher Satz. Im Unterschied zu früheren Untersuchungen zur Wahrnehmbarkeit von Kammfiltereffekten fand eine Lautheitsanpassung der Stimuli statt, so dass davon ausgegangen werden kann, dass die Detektion des veränderten Signals tatsächlich anhand der Parameter *Verzögerungszeit* und *Grad der Ausprägung des Effektes* erfolgte. Die Methodik der Hörversuche richtete sich nach dem aktuellen Stand der psychometrischen Forschung und manifestierte sich in der Programmierung einer Versuchssoftware, welche den gleichbleibenden Ablauf der Hörversuche absicherte. Des Weiteren wurden die Methodik, sowie die Auswahl der Teststimuli in Vorversuchen erprobt und den Versuchsergebnissen entsprechend angepasst.

Überlegungen nach Auswertung der Versuchsergebnisse, die zu allgemeinen Aussagen betreffs der Wahrnehmung der Verzerrungen führen, haben hypothetischen Charakter und seien hier als Thesen zur weiteren Untersuchung formuliert:

1. Die Wahrnehmbarkeit kammfilterartiger linearer Verzerrungen hängt von der Rauschhaftigkeit des Spektrums des untersuchten Signals ab. Dabei gilt: je rauschhafter das Spektrum eines Signals, desto stärker bilden sich bei der Superposition tonale Anteile heraus, welche den wichtigsten *cue* für die Detektion der Verzerrung liefern.
2. Wird durch die kammfilterartige Verzerrung eine bereits im Spektrum des Originalsignals vorhandene Spitze angehoben, so ist diese Überhöhung schlechter zu detektieren, als eine Anhebung eines spektralen lokalen Minimums.
3. Mit zunehmender Verzögerungszeit zwischen den superpositionierten Signalen wächst der Eindruck einer unnatürlichen Räumlichkeit, welcher sich besonders bei Sprachsignalen störend auswirkt und die Detektierbarkeit des Kammfiltereffektes erleichtert.

Nach Auswertung der Versuchsergebnisse lassen sich konkrete Aussagen über die Hörbarkeit kammfilterartiger Verzerrungen bei Kenntnis der Variablen Schallweg- und Pegeldifferenz zweier superpositionierter Signale der hier untersuchten Kategorien treffen. Die ermittelte Hörschwelle resultiert in der allgemeinen Aussage, dass Schallreflexionen mit einer Schallwegdifferenz von drei Zentimetern bis zu fünf Metern, die mit weniger als 19 dB Pegeldifferenz einem spektral kohärenten Signal überlagert werden, zu hörbaren kammfilterartigen Verzerrungen führen können.

7 Literatur

Ahnert, Wolfgang (1988): Comb-filter Distortions and their Perception in Sound Reinforcement Systems, in: Proceedings of the 84th AES Convention, AES Preprint 2565 (A-4).

Allen, Jont (2000): The intensity JND comes from Poisson neural noise: Implications for image coding, in: Rogowitz, B.E. and Pappas, T.N. (Hg.): Proceedings Proc. of SPIE, Human Vision and Electronic Imaging Vol. 3959, S. 222-233.

Anazawa, Takeaki; Takahashi, Yukio; Clegg, Almon H. (1987) : Digital Time- Coherent Recording Technique, in: Proceedings of the 83th AES Convention, Preprint 2493.

Bilsen, F.A.; Ritsma, R.J. (1969): Repetition Pitch and its Implication for Hearing Theory, in: Acustica, Vol. 22, 1969/70, Heft 2, S. 63-73.

Brown, Lawrence G. (1996): Additional rules for the transformed up-down method in psychophysics, in: Perception & Psychophysics, 1996, 58 (6), S. 959-962.

Bücklein, R. (1981): The Audibility of Frequency Response Irregularities, in: Journal of the Audio Engineering Society, Vol. 29, No.3, 1981, S. 126-131.

Burroughs, Lou (1974): Microphones: Design and Application, 1. Aufl., Sagamore Publishing Company, New York, 1974.

Buser, Pierre und Imbert, Michael (1992): Audition, MIT Press. Cambridge, Massachusetts, 1992.

Dai, Huanping (1995): On measuring psychometric functions: A comparison of the constant-stimulus and adaptive up-down methods, in: Journal of the Acoustical Society of America 98 (6), 1995, S. 3135-3139.

Dickreiter, Michael (1997): Handbuch der Tonstudientechnik, Bd.1, 6., verb. Aufl., K.G. Saur Verlag KG, München, 1997.

Dickreiter, Michael (2003): Mikrofonaufnahmetechnik, 3. Aufl., S. Hirzel-Verlag Stuttgart, Leipzig 2003.

DIN 45631: Berechnung des Lautstärkepegels und der Lautheit aus dem Geräuschspektrum, Verfahren nach E. Zwicker, 1991.

Ekman, Györgi (1959): Weber's law and related functions, in: Journal of Psychology, 47, S. 343-352.

Green, D. M. (1990): Stimulus selection in adaptive psychophysical procedures, in: Journal of the Acoustical Society of America, 87, S. 2662-2674.

Hartmann, W. M. (1978): Flanging and Phasers, in: Journal of the Audio Engineering Society, 1978, Vol. 26, No.6, S. 439-443.

Hartmann, William M. (1998): Signals, Sound and Sensation, Springer- Verlag New York, Inc., 1998.

Huygens, Christiaan (1693): Correspondence with de la Hire, in: Oeuvres Complètes, Vol. 10., Bosscha, J. (Hg., 1905), Royal Holland Society of Sciences and Humanities, Haarlem, S. 570.

Jennings, A.R. (2005): On mechanism for the analysis of spectral and temporal envelope shape in the human auditory system, Thesis, Supervisor: Timothy Griffith, Newcastle, 2005.

Kaernbach, Christian (1991): Simple adaptive testing with the weighted up-down method, in: Perception & Psychophysics, 49, S. 227-229.

Klein, Stanley A. (2001): Measuring, estimating and understanding the psychometric function: A commentary, in: Perception & Psychophysics, 63 (8), S. 1421-1455.

Kollmeier, Birger; Gilkey, Robert H.; Sieben, Ulrich K. (1988): Adaptive staircase techniques in psychoacoustics: A comparison of human data and a mathematical model, in: Journal of the Acoustical Society of America, 83, S. 1852-1861.

Kuhl, Walter (1965): Das Zusammenwirken von direktem Schall, ersten Reflexionen und Nachhall bei der Hörsamkeit von Räumen und Schallaufnahmen, in: Rundfunktechnische Mitteilungen, Band 9, H. 3, 1965, S. 170-183.

Kuhl, Walter (1969): Unterschiedliche Bedingungen beim Hören in einem Raum und bei elektroakustischen Übertragungen, in: Rundfunktechnische Mitteilungen, Bd. 13, 1969, Nr.5, S. 205-208.

Leek, M. R.; Hanna, T. E.; Marshall, L. (1992): Estimation of psychometric functions from adaptive tracking procedures, in: Perception & Psychophysics, 51, 1992, S. 247-256.

Leek, M. (2001): Adaptive procedures in psychophysical research, in: Perception & Psychophysics, 2001, 63 (8), S. 1279-1292.

Levitt, H. (1971): Transformed up-down methods in psychoacoustics, in: Journal of the Acoustical Society of America, 49, S. 467-477.

Luce, R. Duncan (1993): Sound & Hearing. A Conceptual Introduction, Lawrence Erlbaum Associates, Inc., Hillsdale, New Jersey, 1993.

Meyer, Norbert (1954): Über die Hörbarkeit von linearen Verzerrungen bei natürlichen Klängen, in: Funk und Ton, Nummer 1, Jahrgang 1954, S. 1-6.

Müller, L. (1968): Zur Klangfärbung durch kurze Zeitreflexionen bei Rauschen, Sprache und Musik, Vortrag auf dem 6. Internationalen Kongress für Akustik, Tokio, 1968, S. E-61 – E-64.

Preis, D. (1976): Linear Distortion, in: Journal of the Audio Engineering Society, Vol. 24, No. 5, 1976, S. 346-367.

Ritter, Steffen (2005): Magnetoencephalographic Study on the Representation of Huygens Noise in the Auditory Cortex of Humans, Dissertation, Heidelberg, 2005.

Shelton, B. R.; Scarrow, I. (1984): Two- alternative versus three- alternative procedures for threshold estimation, in: Perception & Psychophysics, 1984, 35 (4), S. 385-392.

Taylor, M.M.; Creelman, C.D. (1967): PEST: Efficient estimates on probability functions, in: Journal of the Audio Engineering Society, 41, S.782-787.

Theile, Günther (1980): Über die Lokalisation im überlagerten Schallfeld, Dissertation, TU Berlin, 1980.

Treutwein, B. (1995): Adaptive psychophysical procedures, in: Vision Research, 35, S. 2503-2522.

Völker, E.-J. (1974): Beeinflussung der Mikrophoneaufnahme durch Schallreflexion, Klangfärbung und Aufnahmetechnik, in: Fernseh- und Kinotechnik, Nr. 8, 1974, S. 223-226.

Webers, Johannes (1999): Das Handbuch der Tonstudioteknik, 7. Aufl., Franzis´ Verlag GmbH, Poing, 1999.

Internetquellen:

Hastings, Aaron (2000): Matlabcode zur Berechnung der Lautheit nach Zwicker, verfügbar unter:

<http://www.auditory.org/mhonarc/2000/msg00498.html> .

Zugriff: März 2006.

Nordtest (2001): Acoustics: human sound perception – guidelines for listening tests, NT ACOU 111, Sapoo (Finland), 2001, verfügbar unter:

<http://www.nordicinnovation.net/nordtestfiler/acou111.pdf> .

Zugriff: März 2006.

Sengpiel, Eberhard (2000): Kammfiltereffekt bei der Aufnahme, verfügbar unter:

<http://www.sengpielaudio.com/KammfiltereffektBeiDerAufnahme.pdf> .

Zugriff: März 2006.

Anhang A – Versuchsergebnisse

Ergebnisse des ersten Vorversuches

<i>peak</i> -Überhöhung an der Hörschwelle [dB] für:						
Verzögerungszeit [ms]	Piano 1s	Piano 2s	Snare 1s	Snare 2s	Sprache 1s	Sprache 2s
0,1	1,84	1,78	1,28	1,03	2,41	1,97
0,5	1,84	1,91	1,53	1,50	2,72	2,31
0,8	1,75	1,59	1,25	0,88	1,88	1,22
1	2,00	2,22	1,50	1,38	1,56	1,50
1,2	1,34	1,38	0,81	1,00	1,50	1,78
1,5	1,91	1,81	1,66	1,53	1,63	1,63
2	2,00	1,84	1,03	0,69	1,72	1,44
3	2,50	2,28	2,09	1,81	2,03	1,66
4	1,91	2,00	1,50	0,78	1,47	1,28
6	3,00	2,13	2,56	2,19	1,88	1,47
10	2,33	2,08	2,25	1,42	1,88	2,00
15	2,28	2,75	2,41	2,25	1,60	1,13

Tabelle 10: Erster Vorversuch. Überhöhungen an den *peak*-Frequenzen an der ermittelten Hörschwelle [dB].

Ergebnisse des zweiten Vorversuches

Länge des Stimulus [s]	<i>peak</i> -Überhöhung [dB] an der Hörschwelle für	
	T=0,8 ms	T=1,2 ms
1	1,38	1,44
2	1,08	1,11
3	1	0,89
4	1,18	0,95

Tabelle 11: Zweiter Vorversuch. Überhöhungen an den *peak*-Frequenzen an der ermittelten Hörschwelle [dB].

Hauptversuch

PIANO	Verzögerungszeit [ms]							
Versuchsperson	0,10	0,50	0,80	1,00	1,50	3,00	6,00	15,00
1	1,25	1,13	1,44	1,63	1,88	2,75	1,63	2,00
4	0,69	1,69	1,13	1,13	1,38	2,38	2,63	1,19
7	1,63	2,19	1,88	1,50	1,94	3,88	2,56	2,75
10	1,38	1,31	1,06	1,38	1,81	2,13	1,81	1,19
13	1,06	1,63	2,38	2,56	2,31	1,81	1,81	2,13
16	2,19	2,50	2,13	2,75	2,75	2,38	2,94	3,06
19	2,13	1,81	1,63	2,13	1,63	2,38	2,00	2,69
22	1,88	1,69	1,31	2,19	2,06	1,88	2,50	2,75
25	5,81	2,25	1,44	1,50	0,75	2,06	1,94	3,25
28	2,00	2,69	2,19	1,63	1,81	2,19	1,88	2,38
31	2,81	2,69	2,31	2,81	2,44	2,31	2,19	2,50
34	2,56	1,63	1,75	1,94	1,31	2,25	2,06	1,63

Tabelle 12: Ergebnisse des Hauptversuches in der Kategorie Piano. Angegeben sind die *peak*-Überhöhungen [dB] an der Hörschwelle.

SNARE	Verzögerungszeit [ms]							
Versuchsperson	0,10	0,50	0,80	1,00	1,50	3,00	6,00	15,00
2	2,69	1,00	0,63	0,63	1,63	1,06	1,81	3,25
5	1,06	1,31	1,00	1,00	1,31	1,13	1,69	2,94
8	1,81	2,31	2,00	1,56	1,13	1,19	2,19	2,44
11	2,00	0,88	0,94	0,75	0,69	0,56	1,81	1,56
14	1,44	0,69	0,63	0,69	0,81	0,94	1,88	2,25
17	1,25	0,81	0,63	0,50	0,63	0,94	1,38	1,75
20	2,06	0,75	1,31	2,00	1,69	1,81	2,13	3,19
23	2,13	1,00	0,63	2,13	1,63	1,88	2,13	2,88
26	0,13	1,38	1,31	1,88	1,50	1,31	2,19	3,13
29	2,69	0,81	0,81	0,63	1,13	0,69	1,19	0,63
32	1,88	0,38	0,88	0,56	0,81	0,56	1,25	1,50
35	2,69	1,50	1,38	1,38	1,56	1,25	2,19	2,88

Tabelle 13: Ergebnisse des Hauptversuches in der Kategorie Snare. Angegeben sind die *peak*-Überhöhungen [dB] an der Hörschwelle.

SPRACHE	Verzögerungszeit [ms]							
Versuchsperson	0,10	0,50	0,80	1,00	1,50	3,00	6,00	15,00
3	2,19	2,06	3,13	2,00	2,50	1,50	1,63	0,88
6	1,63	1,00	0,63	1,38	0,81	1,13	0,88	0,63
9	2,69	2,56	1,63	1,38	1,38	1,88	0,81	1,63
12	2,88	2,50	2,06	2,44	1,63	1,31	1,69	1,63
15	3,56	3,13	2,63	2,50	1,69	1,00	1,69	1,13
18	3,94	2,75	2,63	2,81	1,69	1,25	1,06	2,06
21	4,19	3,38	3,25	3,50	2,75	1,56	2,38	1,50
24	2,63	1,31	0,88	1,00	1,50	1,25	1,25	0,81
27	3,31	1,88	1,50	0,81	1,44	1,06	1,25	0,88
30	2,38	1,75	2,44	2,06	2,69	1,25	1,38	0,94
33	2,38	1,56	1,63	1,25	1,88	1,06	1,50	1,14
36	1,88	1,38	0,88	1,38	1,63	1,31	1,31	1,13

Tabelle 14: Ergebnisse des Hauptversuches in der Kategorie Sprache. Angegeben sind die peak-Überhöhungen [dB] an der Hörschwelle.

Anhang B – Instruktionen für den Hauptversuch

Instruktionen für den Hörtest

Herzlich Willkommen und vielen Dank, dass Sie sich die Zeit nehmen, an diesem Hörtest teilzunehmen!

Dieser Hörtest ist Bestandteil meiner Magisterarbeit, die sich mit der Wahrnehmbarkeit von kammfilterartigen Verzerrungen beschäftigt. Dieser Hauptversuch soll Aufschluss darüber geben, wie die Veränderung diverser Parameter die Hörbarkeit von kammfilterartigen Verzerrungen beeinflusst. Diese Verzerrungen äußern sich in unter anderem einer Verfärbung des Audiosignals.

Insgesamt besteht der Versuch aus einer Trainingsphase und mehreren Testphasen. Sie können zwischen diesen Phasen, welche jeweils ca. 5 bis 6 Minuten dauern, kurze Pausen machen.

Sie werden nachher dort vor dem Monitor sitzen und, wenn Sie bereit sind und auf den Button „START“ geklickt haben, über Kopfhörer die Testsignale hören. Es handelt sich bei jeder Darbietung um drei kurze Signale identischer Länge. Eines dieser Signale unterscheidet sich klanglich von den beiden anderen. Die Reihenfolge der Signale ist zufällig. Ihre Aufgabe ist es nun, zu hören, welches Sample sich Ihrer Meinung nach von den anderen unterscheidet. Klicken Sie nach dem Anhören aller drei Samples bitte die entsprechende Box und dann den Button „Weiter“ an. Es gibt in diesem Hörversuch keine richtigen oder falschen Antworten – es gibt nur Ihre Antwort. Der Test endet automatisch und Pausen werden annonciert. Ich stelle fest, dass Sie freiwillig an diesem Test teilnehmen und keinerlei Bezahlung erhalten.

Ich würde Sie jetzt noch bitten, mir diesen Fragebogen auszufüllen.

Vielen Dank nochmals für Ihre Mithilfe und jetzt

Viel Spaß!

Anhang C – Fragebogen

Fragebogen zum Hörtest

Alter:

Geschlecht:

Beruf:

Wie schätzen Sie selbst Ihr Hörvermögen ein?
(auf einer Skala von 1 bis 6; 1:sehr gut; 6: sehr schlecht)

Wie groß ist Ihre Erfahrung im Erkennen von Klangverfärbungen?
(auf einer Skala von 1 bis 6; 1:sehr viel Erfahrung; 6: keine Erfahrung)

Spielen Sie ein oder mehrere Instrumente? Wenn ja: welche(s) und wie lange schon?

Ich spiele kein Instrument.

Ich spiele seit ____Jahren _____

Wissen Sie, wie sich Kammfiltereffekte anhören? Wenn ja, bitte beschreiben Sie kurz den Klang einer kammfilterartigen Verzerrung!

Ich weiß nicht, wie eine derartige Verzerrung klingt.

Ich würde den Klang einer derartigen Verzerrung wie folgt beschreiben:

Anhang D – Inhalt der CD und Bedienungsanleitung

Auf der beigefügten CD sind ein Teil der für den Hörversuch benutzten Software sowie alle Stimuli enthalten. Mittels der Software kann der gesamte Hauptversuch nachvollzogen werden. Dazu sollte der Inhalt der CD in ein Unterverzeichnis des *work*-Ordners von MATLAB kopiert, sowie die entsprechenden Pfade gesetzt werden. Gestartet wird das Programm mit der Eingabe „HauptversuchMain“ in der *command line*. Dann muss lediglich noch eine Zahl (Nummer der Versuchsperson) eingegeben und den Bildschirmanweisungen gefolgt werden. Pausen und das Ende des Versuchs werden automatisch angezeigt. Dieses Programm ist für eine Bildschirmauflösung von 1024 x 768 Bildpunkten optimiert. Der Code wurde unter MATLAB 6.3 unter Verwendung diverser Toolboxen erstellt.

Das Ergebnis einer Versuchssession ist ein Datenfile mit folgender Namensstruktur:

ExperimentErgebnisHauptDelayTimeSignaltyp.mat

Dabei bezeichnet *DelayTime* die Verzögerungszeiten in Millisekunden multipliziert mit 100, so dass 1 ms der Zahl 150 und 0,5 ms der Zahl 50 entsprechen. *Signaltyp* ist entweder *PianoDrei*, *SnareRollDrei* oder *SpeechDrei*, so dass ein Ergebnisfilename beispielsweise lauten könnte: ExperimentErgebnisHaupt1500SpeechDrei.mat .

Das Ergebnisfile enthält eine *cell*-Struktur namens *ExperimentErgebnis* mit folgendem Aufbau:

ExperimentErgebnis{Versuchsperson,1} Ergebnis der Mittlung aus den letzten 4 Reversals
ExperimentErgebnis{Versuchsperson,2} detaillierter Verlauf der Antworten --- 0=falsch
ExperimentErgebnis{Versuchsperson,3} Dummy für Lauf B bei *interleaved* Methode
ExperimentErgebnis{Versuchsperson,4} detaillierter Verlauf der Levels
ExperimentErgebnis{Versuchsperson,5} Dummy für Lauf B bei *interleaved* Methode
ExperimentErgebnis{Versuchsperson,6} Reversals Lauf A

ExperimentErgebnis{Versuchsperson,7} Dummy für Lauf B bei *interleaved* Methode
ExperimentErgebnis{Versuchsperson,8} in welchem Trial Reversal A
ExperimentErgebnis{Versuchsperson,9} Dummy für Lauf B bei *interleaved* Methode
ExperimentErgebnis{Versuchsperson,10} benötigte Zeit

Bei der Auswertung ist unbedingt darauf zu achten, dass in dem File die Ergebnisse als *Faktorstufen* abgelegt sind, nicht als konkrete Zahlenwerte. Auf Grund des Erzeugungsalgorithmus´ repräsentiert der Zahlenwert einer Faktorstufe das Vierfache der Überhöhung des Signals an den *peak*-Frequenzen in Dezibel.

Anhang E – Quellcodes

An dieser Stelle seien die Quellcodes der wichtigsten verwendeten Funktionen und Scripte dokumentiert. Die Bedeutung der Programmteile lässt sich aus Abbildung 13 auf Seite 35 ersehen.

Quellcode des Ausgangsscripts des Hauptversuches

HauptversuchMain.m

Quellcode der Funktion *AdaptiveTest.m*

Quellcode der Funktion *PlayThreeA.m*

Quellcode der Funktion *ThreeAFC.m*

